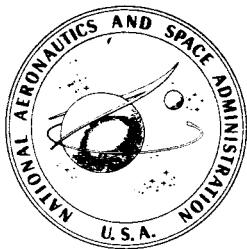


PUBLICATIONS OF GODDARD SPACE FLIGHT CENTER 1964



**Volume II
SPACE TECHNOLOGY**

For sale by the Superintendent of Documents, U. S. Government Printing Office
Washington, D. C. 20402 Price \$7.75

FOREWORD

This publication is a collection of articles, papers, talks, and reports generated by the scientific and engineering staff of Goddard Space Flight Center in the year 1964. Many of these articles were originally published in scientific or engineering Journals or as official NASA technical publications, while others are documents of a more informal nature. All are reprinted here as nearly verbatim as typography and format will permit.

These articles are grouped into broad subject categories, but no detailed subdivision has been made. Within each category, the articles are arranged alphabetically by author. An overall author index is given in the back of the volume.

The years 1963, 1964, and 1965 are being published as whole-year issues, and the resulting size dictates the use of two volumes; the first volume is titled Space Sciences, and the second Space Technology. It is anticipated, however, that future issues will be quarterly single volumes.

CONTENTS

I. PROJECTS AND PROGRAMS

"A SATELLITE SYSTEM FOR RADIO ASTRONOMICAL MEASUREMENTS AT LOW FREQUENCIES", <i>by J. K. ALEXANDER and R. G. STONE</i>	1
"INTERIM STATUS REPORT INTERPLANETARY MONITORING PLATFORM IMP I—EXPLORER XVIII", <i>by F. A. CARR</i>	7
"ADVANCED RESEARCH FOR COMMUNICATIONS SATELLITES", <i>by R. J. MACKEY, JR.</i>	20

II. SPACE DYNAMICS AND CONTROL SYSTEMS

"A LIGHTWEIGHT FOLDING BOOM FOR ACCURATELY PLACING SCIENTIFIC EXPERIMENTS UP TO 25 FEET FROM LARGE SPACECRAFT", <i>by J. P. BAUERNSCHUB, JR.</i>	25
"VIBRATION ISOLATION OF SATELLITE TAPE RECORDERS", <i>by J. H. CONN</i>	31
"ROLLING ELEMENT SLIP RINGS FOR VACUUM APPLICATION", <i>by E. J. DEVINE</i>	36
"ENERGY ABSORBER FOR THE ARIEL I INSTRUMENT BOOMS", <i>by T. L. ENG</i>	45
"EXPLORER XVII (1963 9A) REAL TIME PCM TELEMETRY DATA PROCESSING AND DISPLAY TEST STAND", <i>by M. M. GRANT, C. C. STEPHANIDES, and W. N. STEWART</i>	54
"AN EFFICIENT PCM ERROR CORRECTION AND SYNCHRONIZATION CODE", <i>by M. S. MAXWELL and R. L. KUTZ</i>	98
"DEVELOPMENT OF A 1200 FOOT ENDLESS-LOOP TAPE TRANSPORT FOR SATELLITE APPLICATIONS", <i>by K. W. STARK</i>	107
"SURVEY OF CONTINUOUS LOOP RECORDERS DEVELOPED FOR AND FLOWN ON METEOROLOGICAL SATELLITES", <i>by K. W. STARK and A. F. WHITE, JR.</i>	119
"DEVELOPMENT OF A BRUSHLESS DC MOTOR FOR SATELLITE APPLICATION", <i>by P. A. STUDER</i>	131
"A MEDIUM-DATA-RATE DIGITAL TELEMETRY SYSTEM", <i>by M. R. TOWNSEND, P. M. FEINBERG, and J. G. LESKO, JR.</i>	139

III. SPACECRAFT AND SUBSYSTEMS

"DIRECT MICROWAVE TO MICROWAVE TRANSPONDERS FOR COMMUNICATION SATELLITES", <i>by W. K. ALLEN</i>	149
"ELECTRONICALLY STEERABLE ANTENNAS FOR COMMUNICATION SATELLITES", <i>by D. L. BACKUS</i>	153
"ORBITAL PLACEMENT AND CONTROL OF THE SYNCOM II SPACECRAFT", <i>by J. L. BAKER</i>	163
"APPLICATION OF INTEGRATED CIRCUITS TO TELEMETRY SYSTEMS", <i>by J. M. BALDERSTON</i>	185
"THE USE OF SOLID CIRCUITS IN SATELLITE INSTRUMENTATION", <i>by E. G. BUSH</i>	193
"THE EPE-D TWO YEAR TIMER", <i>by T. D. CLEM</i>	200
"ANALYSIS OF NON-LINEAR NOISE IN FDM TELEPHONY TRANSMISSION OVER AN SSB-PM SATELLITE COMMUNICATION SYSTEM", <i>by P. J. HEFFERNAN</i>	208
"THE AUTOMATIC PICTURE TRANSMISSION SYSTEM ON TIROS VIII", <i>by D. W. HOLMES and C. M. HUNTER</i>	228
"TELEMETRY INSTRUMENTATION OF THE ECHO II PASSIVE COMMUNICATIONS SATELLITE", <i>by H. S. HORIUCHI</i>	234

"EXPERIMENTS TO DETERMINE COMMUNICATION CAPABILITY OF THE ECHO II SATELLITE", <i>by</i> W. C. NYBERG, R. L. KAISER, <i>and</i> W. E. LEAVITT.....	242
"SLOPE DETECTION AS A METHOD OF DETERMINING THE PEAK POWER POINT OF SOLAR ARRAYS", <i>by</i> J. PAULKOVICH.....	246
"PULSE-FREQUENCY-MODULATION TELEMETRY", <i>by</i> R. W. ROCHELLE.....	252
"A SOLID STATE SATELLITE SEPARATION SEQUENCE TIMER", <i>by</i> J. C. SCHAFFERT <i>and</i> T. D. CLEM.....	292
"TRANSISTORIZED VHF TRANSMITTER DESIGN FOR SPACECRAFT APPLICATIONS", <i>by</i> C. R. SOMERLOCK.....	298
"ON THE DESIGN OF PFM TELEMETRY ENCODERS", <i>by</i> H. D. WHITE, JR.....	318

IV. SENSORS

"THE TIROS LOW RESOLUTION RADIOMETER", <i>by</i> F. BARTKO, V. KUNDE, C. CATOE, <i>and</i> M. HALEV.....	339
"ABSOLUTE CALIBRATION AND USE OF A SOFT X-RAY SOURCE OF CHARACTERISTIC CARBON RADIATION", <i>by</i> A. J. CARUSO <i>and</i> W. M. NEUPERT.....	366
"A MULTIPLE ELEMENT SOFT X-RAY SOURCE", <i>by</i> A. J. CARUSO <i>and</i> W. M. NEUPERT.....	375
"ULTRAVIOLET PHOTOGRAPHY AND SPECTROSCOPY USING A SPECTRALLY SELECTIVE IMAGE CONVERTER", <i>by</i> L. DUNKELMAN <i>and</i> J. P. HENNES.....	377
"LOW ENERGY CHARGED PARTICLE DETECTION USING THE CONTINUOUS CHANNEL ELECTRON MULTIPLIER", <i>by</i> D. S. EVANS.....	383
"AN ULTRAVIOLET SPECTROPHOTOMETER FOR SATELLITE ASTRONOMY", <i>by</i> K. HALLAM <i>and</i> J. MANGUS.....	394
"HORIZON SENSING FOR ATTITUDE DETERMINATION", <i>by</i> B. K. LUNDE.....	398
"RESPONSE OF MODIFIED REDHEAD MAGNETRON AND BAYARD-ALPERT VACUUM GAUGES ABOARD EXPLORER XVII", <i>by</i> G. P. NEWTON, D. T. PELZ, G. E. MILLER, LTJG, USN, <i>and</i> R. HOROWITZ.....	405
"AN APPLICATION OF PRESSURE DIVISION FOR CALIBRATING IONIZATION GAUGES", <i>by</i> J. R. POYNTER <i>and</i> S. SAITO.....	412
"ANALOG TO DIGITAL CONVERTER FOR THE S-57 GSFC ION-CHAMBER EXPERIMENT", <i>by</i> J. C. THORNWALL.....	419
"THE DETERMINATION OF ABSOLUTE PHOTON FLUXES AND APPLICATIONS TO CALIBRATION PROCEDURES IN THE 100A TO 300A RANGE", <i>by</i> D. H. TOMBOULIAN.....	430
"CORRELATION OF ULTRAVIOLET EXCITED LUMINESCENCE AND INTERBAND TRANSITIONS IN HEXAGONAL ZINC SULFIDE", <i>by</i> W. C. WALKER.....	435
"PILE-OF-PLATES POLARIZER FOR THE VACUUM ULTRAVIOLET", <i>by</i> W. C. WALKER.....	441
"POLARIZATION DEPENDENCE OF THE FAR ULTRAVIOLET OPTICAL TRANSITIONS IN CdS", <i>by</i> W. C. WALKER.....	444
"AN ULTRAVIOLET ROCKET STELLAR SPECTROMETER", <i>by</i> D. U. WRIGHT, JR.....	446

V. GENERAL ELECTRONICS

"DEMODULATION OF WIDE-BAND FREQUENCY MODULATION BY A PHASE-LOCK TECHNIQUE", <i>by</i> R. C. BOOTON, JR.....	453
"HIGH EFFICIENCY COMPLIMENTARY ASTABLE PRODUCES PULSES OF EITHER POLARITY", <i>by</i> S. CANTARANO.....	463
"VOLTAGE AND CURRENT MONITOR FOR SCIENTIFIC SATELLITES", <i>by</i> W. R. CROCKETT.....	464
"A HYBRID RANGING SYSTEM FOR SPACECRAFT", <i>by</i> R. T. FITZGERALD, P. ENGELS, W. SHAFFER, E. HABIB, <i>and</i> M. MITCHKO.....	473
"PHOTOEFFECTS AND RELATED PROPERTIES OF SEMICONDUCTING DIAMONDS", <i>by</i> C. C. JOHNSON, H. STEIN, T. YOUNG, J. WAYLAND, <i>and</i> W. LEIVO.....	490
"A RUTILE TRAVELING WAVE MASER", <i>by</i> C. C. JOHNSON.....	499
"WELDED ELECTRONIC MODULE FABRICATION", <i>by</i> J. C. LYONS <i>and</i> D. R. DARGO.....	506

"WIDE RANGE PHASE DETECTOR", by G. B. ROBINSON.....	514
"THE DIGITAL TAPE RECORDER ANALYZER", by J. A. SCIULLI.....	531
"DESIGNING TOROIDAL INDUCTORS WITH DC BIAS", by G. D. SMITH.....	542
"A VERTICAL TEST RANGE FOR ANTENNA RADIATION MEASUREMENTS", by J. STECKEL and R. KORVIN.....	549

VI. ENVIRONMENTAL TESTING

"SOLAR SIMULATION TESTING OF AN EARTH SATELLITE AT GODDARD SPACE FLIGHT CENTER", by R. E. BERNIER, R. H. HOFFMAN, A. R. TIMMINS, and E. I. POWERS.....	567
"CORIOLIS EFFECTS IN COMBINED ENVIRONMENT TESTING", by J. C. BOYLE.....	577
"PROBLEMS IN THE CONSTRUCTION OF A SPACE ENVIRONMENT SIMU- LATOR", by D. S. COPE.....	609
"POST IRRADIATION ROOM TEMPERATURE ELECTRICAL CHARACTERIS- TICS OF N-P SILICON SOLAR CELLS", by B. T. CUNNINGHAM and E. MOSS.....	618
"AN ANALYSIS OF ERRORS IN THE GEOGRAPHIC REFERENCING OF NIM- BUS CLOUD PICTURES", by E. M. DARLING, JR.....	625
"A GEOGRAPHICAL GRID FOR NIMBUS CLOUD PICTURES", by E. M. DARLING, JR.....	642
"HIGH SPEED VACUUM PERFORMANCE OF MINIATURE BALL BEARINGS LUBRICATED WITH COMBINATIONS OF BARIUM, GOLD, AND SILVER FILMS", by T. W. FLATLEY.....	652
"ANALYSIS OF DISSIMILAR SATELLITE APPENDAGES DURING ERECTION", by R. W. FORSYTHE.....	659
"A METHOD FOR SIMULATING ZERO-GRAVITY ERECTION OF SATELLITE APPENDAGES", by R. W. FORSYTHE.....	669
"MODELLING OF SPACECRAFT UNDER RANDOM LOADING", by J. E. GREEN- SPON.....	679
"ENVIRONMENTAL TEST PROGRAM FOR ARIEL I", by W. H. HORD, JR.....	684
"CREATION OF SPACE VACUUM UTILIZING CRYOGENIC TECHNIQUES", by J. E. A. JOHN and W. F. HARDGROVE.....	725
"A BIFILAR MOMENT OF INERTIA FACILITY", by G. H. JONES.....	731
"TRIAXIAL BALANCING TECHNIQUES (A STUDY OF SPACECRAFT BALANCE WITH RESPECT TO MULTIPLE AXES)", by W. E. LANG.....	742
"SPACE VACUUM INVESTIGATIONS", by H. MAURER, JR.....	765
"LAUNCH ENVIRONMENT PROFILES FOR SOUNDING ROCKETS AND SPACECRAFT", by W. J. NEFF and R. A. MONTES DE OCA.....	777
"THREE AND FOUR COIL SYSTEMS FOR HOMOGENEOUS MAGNETIC FIELDS", by M. E. PITTMAN and D. L. WALDELICH.....	801
"ADAPTATION OF AN MoS ₂ "IN SITO" PROCESS FOR LUBRICATING SPACE- CRAFT MECHANICAL COMPONENTS", by C. E. VEST.....	812
"AN INTEGRATED SYSTEM FOR COLLECTING AND ANALYZING ENVIRON- MENTAL TEST DATA", by P. YAFFEE, F. STARBUCK, J. W. BAILEY, C. W. SCOTT, A. C. JOHNSON, C. F. RILEY, JR., W. E. FLOWERS, and R. K. HOVEY.....	823

VII. TRACKING SYSTEMS

"SOLAR CELL RADIATION DAMAGE STUDIES WITH 1 MEV ELECTRONS AND 4.6 MEV PROTONS", by W. R. CHERRY and L. W. SLIFER.....	835
"CODING AN ANALOG VARIABLE FOR CONSTANT PERCENTAGE ERROR", by R. A. CLIFF.....	854
"THE ELECTRICAL CHARACTERISTICS OF IRRADIATED SILICON SOLAR CELLS AS A FUNCTION OF TEMPERATURE", by B. T. CUNNINGHAM, R. L. SHARP, and L. W. SLIFER.....	866
"EXPERIMENT. TIME SYNCHRONIZATION OF REMOTE STATIONS USING SYNCOM SATELLITE", by R. L. GRANATA, P. D. ENGLS, and P. F. McCAUL.....	877
"AN AUTOMATIC MICROWAVE POLARIZATION ANALYZER", by J. A. KAISER and K. L. HANLIN.....	879

"PHASE COMPARISON PROPERTIES OF THE DUAL MODE SPIRAL WITH APPLICATION TO DIRECTION FINDING", <i>by J. A. KAISER, and J. H. LITTLE</i>	884
"DEVELOPMENT OF OPTIMUM FRAME SYNCHRONIZATION CODES FOR GODDARD SPACE FLIGHT CENTER PCM TELEMETRY STANDARDS", <i>by J. L. MAURY, JR., and F. J. STYLES</i>	889
"TIMING ERRORS—THEIR DETECTION AND CORRECTION IN THE IMP INFORMATION PROCESSING SYSTEM", <i>by W. H. MISH</i>	900
"GEOS-1 LASER RETROREFLECTOR DESIGN AND PRELIMINARY SIGNAL CALCULATIONS", <i>by H. H. PLOTKIN</i>	926
"THERMAL RADIATION TO A FLAT SURFACE ROTATING ABOUT AN ARBITRARY AXIS IN AN ELLIPTICAL EARTH ORBIT: APPLICATION TO SPIN-STABILIZED SATELLITES", <i>by E. I. POWERS</i>	931
"APPLICATION OF EDDY CURRENT CLUTCHES TO TRACKING ANTENNA DRIVE SYSTEMS", <i>by N. A. RAUMANN</i>	945
"A METHOD FOR COMPARING THE REFLECTING POWER OF CONCAVE GRATINGS IN THE SOFT X-RAY REGION", <i>by D. H. TOMBOULIAN and W. E. BEHRING</i>	966
"A NARROW-BAND TRACKING FILTER", <i>by R. WOODMAN</i>	972

VIII. ORBIT DETERMINATION

"DIGITAL OHMMETER", <i>by J. SEMYAN</i>	979
"LOCAL ERROR CONTROL IN NUMERICAL INTEGRATION THROUGH OPTIMIZING THE ORDER OF THE INTEGRATOR", <i>by W. S. SOAR</i>	988
"NASCOM REAL-TIME SYSTEM, GSFC 490 ELECTRONIC DATA SWITCHING CENTER INSTALLATION", <i>by D. L. SOLLARS</i>	995
"THE DRIFT OF A 24-HOUR EQUATORIAL SATELLITE DUE TO AN EARTH GRAVITY FIELD THROUGH 4TH ORDER", <i>by C. A. WAGNER</i>	1008

IX. GENERAL

"RESONANCE SCATTERING AT LYMAN-ALPHA BY AN ATOMIC HYDROGEN CELL", <i>by J. E. BLAMONT, P. DELACHE, and A. K. STOBER</i>	1037
"FREQUENCY AND TIME SYNCHRONIZATION AT NASA TRACKING STATIONS", <i>by A. R. CHI</i>	1041
"PRELIMINARY INVESTIGATION OF A RADIOISOTOPE FUELED THERMIONIC DIODE POWER SYSTEM FOR THE 0.5 TO 3.0 KW POWER RANGE REQUIREMENTS ASSOCIATED WITH ELECTRIC PROPULSION DEVICES", <i>by D. F. HASSON</i>	1051
"FACILITY DESCRIPTION. GODDARD ANTENNA TEST RANGE", <i>by P. A. LANTZ</i>	1071
"HANDBOOK OF NASA/GSFC TRACKING, DATA-ACQUISITION, AND COMMUNICATIONS ANTENNAS", <i>by P. A. LANTZ</i>	1095
"FLIGHT VIBRATION DATA OF THE AEROBEE 150A SOUNDING ROCKET", <i>by J. A. NAGY and G. L. COBLE, JR.</i>	1145
"PROPOSED MECHANISM FOR THERMOPHOTOTROPIC BEHAVIOR IN PEROVSKITE-STRUCTURED TITANATES", <i>by S. M. NEUDER</i>	1160
"A PHENOMENOLOGICAL MODEL OF THERMOPHOTOTROPISM", <i>by J. B. SCHUTT</i>	1168
"THERMOPHOTOTROPISM IN CORTICATED ALKALINE EARTH TITANATES", <i>by J. B. SCHUTT, J. A. COLONY, and D. R. LEPP</i>	1179
AUTHOR INDEX.....	1183

I. PROJECTS AND PROGRAMS

A SATELLITE SYSTEM FOR RADIO ASTRONOMICAL MEASUREMENTS AT LOW FREQUENCIES*

J. K. ALEXANDER AND R. G. STONE

A study program was initiated at the Goddard Space Flight Center in 1962 to survey the astronomical observations most readily and profitably performed by an orbiting radio observatory, to analyze the antenna configurations most appropriate for such a mission, and to develop the radiometers and other supporting instrumentation required for a radio astronomy satellite. This study has culminated in the conceptual design of a spacecraft intended solely for radio astronomical observations at wavelengths beyond the ionospheric cut-off. By utilizing long (250m) antenna elements with a terminating resistance placed an odd number of quarter-wavelengths from the end, a V-antenna can be formed having a beamwidth less than 30° and an 18db front-to-back ratio over a band of frequencies near 5 Mc/s. Proper combinations of four such long elements in a double-V or X configuration, furthermore, can be made to yield some gain over a dipole for observations as low as 0.3 Mc/s. A step-frequency, Ryle-Vonberg radiometer which employs an all solid-state component, comparison noise source has been developed to perform over a 70db dynamic range with a relative accuracy of $\pm .5$ db. A simple, precise instrument to measure both the resistive and reactive components of antenna impedance has also been developed for satellite use. Observations with this system from a gravity-gradient stabilized spacecraft in a 6000 km orbit would afford significant contributions to galactic studies, solar physics, and planetary astronomy by extending radio observations to frequencies in the range 0.3 to 10 Mc/s.

INTRODUCTION

During the first five years of the exploration of space by artificial satellites major scientific effort was directed largely towards the study of the near-earth environment. Due, in part, to the special technological problems involved, astronomical observations from satellites have become an established sub-discipline only in the last few years. In one important area, radio-astronomical observations at wavelengths beyond the limits of the ionospheric window, the pioneering work was confined to a very small group among whom were F. T. Haddock of the University of Michigan, A. E. Lilley of Harvard University, J. H. Chapman of DRTE, and F. G. Smith of Cambridge University. Stimulated by the efforts of these workers and discussions with other astronomers, the Goddard Space Flight Center began to study the development of a satellite radio telescope in 1962. This program included a survey of the astronomical observations most readily and profit-

ably performed by an orbiting radio observatory, analysis of the antenna configurations most appropriate for such a mission, and the study of radiometers and other supporting instrumentation required for a radio astronomy satellite. The study culminated in the conceptual design of a spacecraft intended solely for radio astronomical observations at wavelengths beyond the ionospheric cut-off.

Extrapolation of theoretical models and Alouette satellite measurements of the ionospheric electron density distribution below 1000 km have been combined with experimental results obtained by the incoherent scatter technique at the geomagnetic equator to predict the expected variations of electron plasma frequency with altitude and time (R. E. Bourdeau, private communication). At 6000 km the plasma frequency is expected to range from about 0.15 to 0.3 Mc/s. Although this is only an approximate picture, there is sufficient experimental evidence to conclude that, at this altitude, radio astronomical measurements can be made down to 0.3 Mc/s independent of latitude and diurnal time.

*Published as Goddard Space Flight Center Document X-615-64-220, August 1964.

This report describes the major features of the proposed design which provides for measurement of cosmic radio noise below 10 Mc/s with a directive antenna. In the first section we briefly summarize the scientific objectives of the satellite experiment. The second and third sections are devoted to discussion of the antenna concepts and electronics, respectively. The salient features of the proposed radio astronomy satellite are then summarized in the final section.

SCIENTIFIC OBJECTIVES

Galactic Studies

By mapping the distribution of cosmic radio noise at frequencies below 10 Mc/s, it should be possible to estimate the emissivity of source regions in the galaxy and thereby infer information regarding the galactic magnetic fields and the relativistic electron spectrum. If the relativistic electrons are produced in the formation of cosmic rays, we may have a method of studying the problem of cosmic ray formation in the galaxy. From observations of the change in the synchrotron emission spectrum due to free-free absorption in HII regions we may estimate distances to the HII regions as well as electron densities and/or temperatures for assumed galactic models by measuring optical depth. Since other loss mechanisms such as Bremsstrahlung, inverse Compton effect, and ionization also may be expected to affect the observed radio noise spectrum, its interpretation will never be a completely clear-cut matter. Nevertheless, observations of the spectral and spatial distribution of low frequency cosmic noise will provide a significant new source of information regarding the composition of the galaxy.

Solar Physics

By making measurements over the frequency range of 0.3 to 20 Mc/s it will be possible to observe burst activity at altitudes ranging from about 3 to 20 solar radii in the corona. Observations of the flux spectra and low frequency limits of the different burst types may be expected to contribute to the adoption of a theory for the emission mechanism. Measurements of the dura-

tion and time decay of bursts of type II and III can be interpreted in terms of the coronal electron temperature and can facilitate the measurement of the temperature distribution in the corona for assumed electron density distributions. The measurements will therefore serve as a powerful tool for study of the interaction of solar corpuscular streams and the coronal plasma.

Planetary Studies

The sporadic decameter noise bursts from Jupiter have been the subject of considerable attention in recent years, and yet the mechanism of this emission is still not clearly understood. Among the major factors which have hampered the study of the Jovian emission have been the effects of the earth's ionosphere; however by observing from a spacecraft above the ionosphere, one should be able to determine the spectrum of the bursts at frequencies well below 10 Mc/s. Such data are of considerable importance in the determination or confirmation of a theory regarding the emission mechanism and its subsequent interpretation in terms of the structure of the Jovian magnetosphere. Similarly, observations of the dynamic spectra of the Jovian storms from above the ionosphere should simplify the problem of determining how much of the burst fine structure is due to the impulsive nature of the source and the effects of inhomogeneities in the interplanetary plasma.

Although similar noise storms have never been detected unambiguously from other planets, such a possibility cannot be entirely ruled out. The apparent similarity in the magnetospheres of the earth and Jupiter make the occurrence of terrestrial noise bursts quite likely, and such activity may already have been observed in earlier space experiments. The radio astronomy satellite will be able to detect and map magnetospheric radio noise activity above 300 kc/s.

ANTENNA CONCEPTS

In evaluating an antenna system for a first generation radio astronomy satellite, the following properties were considered:

1. The antenna must have sufficient directivity over a large enough band of fre-

quencies to permit accomplishment of the scientific objectives discussed above.

2. The system must be a mechanically simple structure which can be deployed from the spacecraft.
3. The electrical characteristics of the antenna shall not be influenced significantly by the space environment.

After a careful consideration of various antenna types it was decided that a long linear antenna configuration could meet all of these requirements with the least degree of compromise. The "V" antenna was selected over a single linear antenna because it has a solid main lobe and much better side lobe suppression. By insertion of terminating resistances a quarter wavelength from the end of the antenna, the pattern can be made unidirectional with 18db of front-to-back ratio over a band of frequencies. The basic antenna configuration consists of two V's back-to-back which allows a differential method to be utilized to attain a very high front-to-back ratio over the entire frequency range. This particular configuration also provides some additional advantages which will be discussed subsequently. Typical power patterns for the V antenna are shown in Figure 1 and 2.

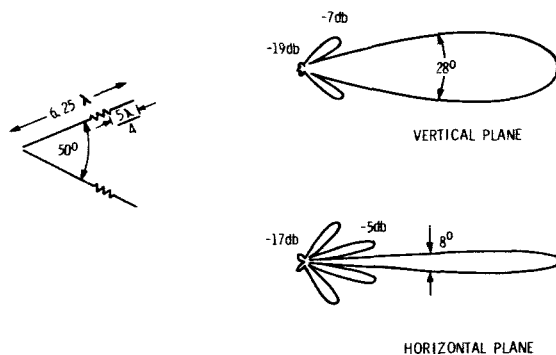


FIGURE 1.—V antenna power patterns for 6.25λ element length.

Considering the system of four conductors which forms the two V arrays, it is seen that several other configurations can be utilized for specific purposes. For example, we know that for a given length the patterns of V antennas with obtuse angles have greater directivity than V antennas with acute angles at the lower frequen-

cies (below 1.5 Mc/s). Consequently, we can switch the four elements of the antenna system so that V antennas with both acute and obtuse angles will be used on a time-sharing basis in order to obtain the maximum directivity at all frequencies.

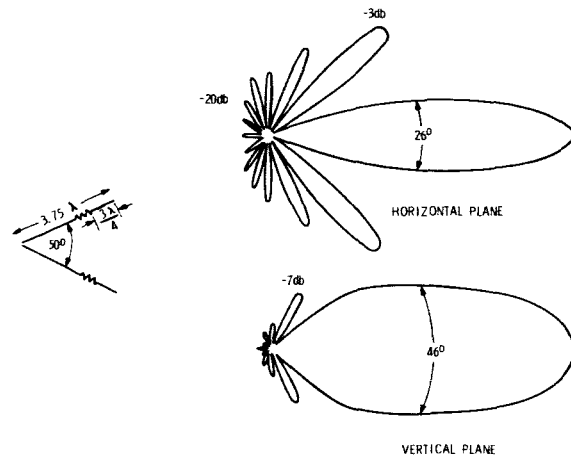


FIGURE 2.—V antenna power patterns for 3.75λ element length.

A desirable but not imperative feature of the radio astronomy experiment is gravity-gradient attitude stabilization. It is desirable principally because (a) it greatly facilitates the reduction of data and (b) it provides a convenient and important method of separating the celestial noise from terrestrial noise and from noise generated by the earth's radiation belt. Separation of extraterrestrial from terrestrial signals is made possible by the high front-to-back antenna gain ratios of the individual V antennas. Consider the principal mode of operation of the antenna elements where one acute angle V antenna points toward the celestial sphere, the other toward the earth, each attached to identical but separate receiving systems. The two V antennas will respond quite differently to the individual signals. Consequently two equations with two unknowns will be telemetered and as a result the extraterrestrial and terrestrial components separately extracted in the data reduction process.

The four 250-meter elements which comprise the double-V antenna system will be extended by use of a motor-driven system similar to that used for the 25-meter antenna elements on the Alouette satellite. Figure 3 shows the shape of the booms

tained at the level required to null the input signal from the antenna.

The all solid state noise source consists of an AGC-controlled, high-gain, wide-band amplifier which amplifies the thermal noise of the early stages, a narrow band filter centered on the operating frequency, and an output power amplifier stage. Since the noise source is capable of providing several milliwatts of output noise power, it is followed by a passive attenuator which reduces the noise to the required reference level within the 70db dynamic range of the system. In order to achieve optimum stability and accuracy, the noise source output (which is proportional to the power received by the antenna) is measured by a self-balancing thermistor bridge.

In order to relate the power delivered to the radiometer to the actual power incident on the antenna one must know the antenna impedance. Since, at the lower observing frequencies, the antenna impedance may be modified by magneto-ionic effects, considerable effort has been directed towards the development of a probe to measure both the real (R) and imaginary (X) components of antenna impedance (Z), where

$$|Z| = (R^2 + X^2)^{1/2}$$

One such probe, which has already been tested on a sounding rocket, is diagrammed in Figure 6. A

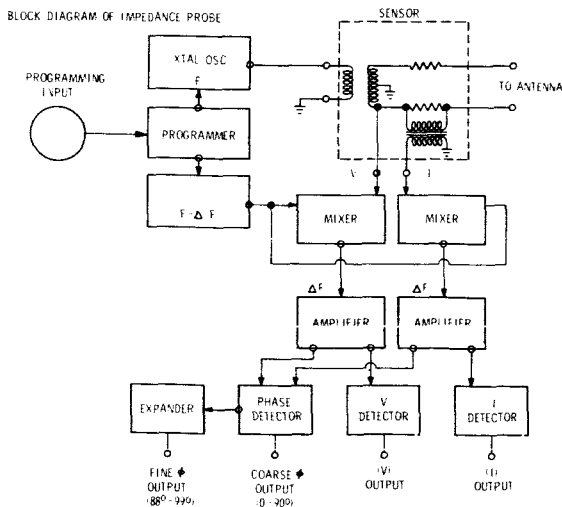


FIGURE 6.—Block diagram of satellite antenna impedance probe.

sensor picks off signals proportional to the antenna voltage, V , and current, I , at a frequency, f . These are mixed with a signal at frequency $f - \Delta f$, and the two resultant signals at frequency Δf are amplified and detected to provide outputs proportional to $|V|$ and $|I|$. The two signals are also compared through a phase-sensitive detector to provide a measurement of phase, ϕ . An expander is included to allow high precision measurements of phase angle.

The fast burst radiometer used in conjunction with the short dipole antenna is designed to measure the characteristics of impulsive noise bursts such as solar Type III bursts. It will consist of four fixed-frequency receivers—all of which will be open to reception at the same time until a burst is detected. The presence of a burst will be established by the detection of a signal above some pre-set threshold, in which case the receiver will store the detected signal and shut itself off. The telemetry system will sample the receivers serially, and each receiver will be reactivated upon reading out its measured data. Upon reactivation, the cycle will repeat. The time between readings on any one channel will be about 160 milliseconds and will require about 20 milliseconds to perform a reading. By locking up a channel after a signal is detected and measuring its time of occurrence and amplitude and continuing in this fashion after each channel is read out, it is possible to measure the intensity, duration, and frequency drift of a burst.

SUMMARY

The system described above will readily meet the objectives of a first generation orbiting radio telescope with maximum simplicity and reliability. Use of the acute angle V antenna will provide directive maps of the galactic noise background at several frequencies in the 1.5 to 10 Mc/s range. By properly utilizing switched combinations of the V booms and the "short" dipole antenna, less directive measurements and integrated noise values can be unambiguously obtained down to 0.3 Mc/s. Observations at a frequency near 20 Mc/s furthermore, will provide a "bridge" to the lowest frequency ground-based measurements. The salient features of the proposed antenna and radiometer are briefly listed in Table I.

TABLE I.

ANTENNA

Leg Length.....	250M
V Apex Angle.....	50°
3db Beamwidth.....	26°×46° at 4.5 Mc/s 8°×28° at 7.5 Mc/s
Front-to-back.....	18db

RADIOMETER

Frequency Range.....	0.3 to 20 Mc/s in 10 steps
Bandwidth.....	20 kc/s
Noise Figure.....	<8db
Dynamic Range.....	>70db
Relative Accuracy.....	±0.5db

Since the system under consideration is estimated to weigh 275 to 300 lb. (125 to 135 kg) it can be placed in a 6000 km altitude, 50° inclination, circular orbit with a relatively low cost launch vehicle. A gravity-gradient stabilized antenna in such an orbit would provide a scan of the sky at an angular rate of about 1.5°/min. while orbit precession would afford coverage of almost the entire celestial sphere in about one year. The experiment should provide valuable insight to the physics of the galaxy, the sun, and the planetary system and should lead to the development of more sophisticated systems for radio astronomical observations from space.

INTERIM STATUS REPORT INTERPLANETARY MONITORING PLATFORM IMP I—EXPLORER XVIII

FRANK A. CARR

The IMP I spacecraft, in orbit for nearly three months as of this date, has already provided over two thousand hours of excellent scientific data.

The spacecraft was launched on 26 November 1963 from AMR, achieving an orbit whose apogee is approximately 31 earth radii. Predictions indicate that the orbit lifetime will extend to mid-November 1965—nearly two years after launch.

The spin-stabilized satellite has exhibited no detectable precession (i.e., <0.1 deg), however, the spin rate unexpectedly decreased during the first orbit: thereafter a continuous increase has been observed. This is attributed to the effect of solar-radiation pressure acting on the spacecraft solar-celled paddles.

All scientific experiments and spacecraft subsystems continue to function properly, although only ~10 percent of the data from the Thermal Ion-Electron experiment is usable due to intermittent operation of the experiment's programmer.

Spacecraft operation thru the brief (~20 minutes) shadow period experienced during each of the first twelve orbits has been excellent. However, extended shadow periods (up to 8.2 hours) will be encountered beginning 160 days after launch (3 May 1964) and ending 5 days later. An 8 hour shadow is expected to reduce the temperature of the spacecraft electronics to approximately -55 to -80°C . The effect of such a condition will be watched with much interest (but little hope for survival).

Two abnormalities in spacecraft operation have been observed—first, an extremely high current drain with a subsequent loss of approximately 5 minutes of data. This occurred late in the first orbit: the spacecraft resumed completely nominal operation within about 10 minutes. Secondly, it has been noted that (some of) the performance parameters appear to be following an upward trend or drift. The cause of these abnormalities has not yet been determined; however, neither condition has caused any serious consequences to spacecraft operation.

Additional reports of this type will be issued periodically during the life of IMP I.

INTRODUCTION

The first Interplanetary Monitoring Platform was launched on 26 November 1963. The Delta launch vehicle marked its twentieth straight success by placing the IMP A spacecraft into a highly elliptical orbit. In general, the operation of the spacecraft has been excellent with over 2000 hours of scientific data recorded by ground stations during the first $2\frac{1}{2}$ months following launch.

This report summarizes some of the pertinent information and data concerning the launch, orbit, performance parameter data and spacecraft operation. Measured data, as well as predictions for the one-year period following launch are presented.

LAUNCH

The 138 lb. IMP A (prototype) spacecraft was launched November 26, 1963 from Pad 17B,

AMR by the Delta 21 vehicle (References 1 and 2). All spacecraft and vehicle checkouts during the countdown were satisfactory and the terminal count was begun at T-35 minutes (2055 hours EST). The spacecraft was turned on eight minutes later at 2103 hours. Lift-off occurred (2130:01.07 EST) only 1.07 seconds after the opening of the forty-five minute "launch window".

The flight sequence was very near nominal and no problems developed with any of the spacecraft functions (i.e., paddle and fluxgate boom erection, separation and magnetometer extension). These functions, telemetered on several performance parameters, were confirmed in real time; however, because of the quality of data recorded during this time period, it has not been possible to process it thru the IMP Information Processing System for precise determination of sequence timing.

The following event times* are given in seconds after lift-off:

Launch Sequence

Event	Actual	Nominal
Lift-off.....	2130:01.07 EST	2130 EST
MECO.....	T+147.0	T+148.6
Jettison Nose Fairing	+182.0	+182.6
SECO.....	+322.9	+317.2
Fire Spin Rockets.....	+337	+338.6
X-258 Ignition.....	+341.9	+344.6
X-258 Burn-out.....	+366.8	+366.6
Paddle Erection.....) Not available, but observed to be very close to nominal.	+401.6
Boom Erection.....		+403.6
Separation.....		+408.6
Tumble X-258.....		+410.6
Spacecraft Sunlight Entrance.....	+1234.4	-----

ORBIT

During and immediately following the launch it was predicted that the apogee of IMP I would be somewhat higher than planned. However, early Minitrack and range and range rate data

*These times are taken from Reference 1 and are preliminary pending publication of Reference 2.

indicated that the actual orbit achieved would be considerably lower than nominal-although within the 3 sigma limit. The reason(s) for the lower apogee has not been (and probably will not be) conclusively ascertained. However, sub-performance of the X-258 third stage motor is suspected to be the primary cause.

It is the opinion of the IMP experimenters that the accomplishment of the scientific mission objectives will not be compromised by the lower apogee.

A comparison of computed orbital elements with nominal values is given below:

	Computed Date		Nominal
	11/27/63†	1/16/64†	
Apogee (km).....	197,615.8	195,552.8	277,184
Perigee (km).....	191.8	197.0	190
Period (min).....	5666.0	5583.2	9164
Inclination (deg).....	33.338	33.336	33.003
Eccentricity.....	.9376	.9369	.9547

†GSFC Operational Control Reports, dtd 2 Dec. 63 and 27 Jan. 64, respectively.

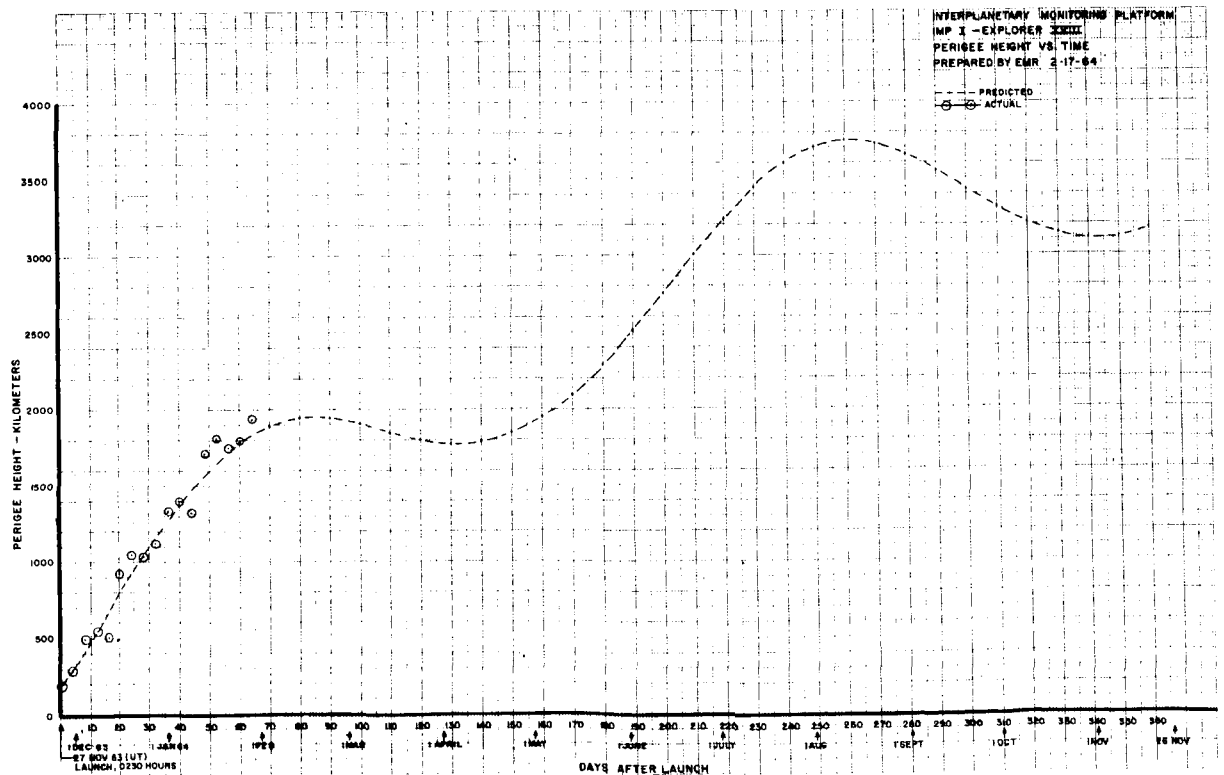


FIGURE 1.—Perigee height vs time after launch.

Computer predictions, using the first 30 days tracking data, of the variations of the orbital parameters for the two-year period following launch have been completed. Table 1 is a summary of these predictions for the first year; Figures 1, 2 and 3 are curves plotted from the data in Table 1 representing perigee height, apogee-sun angle, and spin axis-sun angle respec-

tively (all versus time after launch). The relationship of the IMP I orbit to the earth-sun system is shown in Figure 4. Also shown on Figures 1, 2 and 3 are data as measured for the first 2 months following launch.

The orbit lifetime study indicates that the perigee height will decrease to below 100 miles some 712 days after launch (8 November 1965).

TABLE 1.—IMP I—Summary of Orbital Data—One Year Predictions—Based on First 30 Days Data
(Computed 16 Jan 64)

Date	Days After Launch	Perigee* (Km)	Inclination (Degrees)	Shadow (Hrs/Orbit)	Apogee** Sun Angle (Degrees)	Spin-axis*** Sun Angle (Degrees)
11/27/63	0	197	33.34	0.37	25	110
12/07/63	10	484	33.35	0.37	34	116
12/17/63	20	788	33.26	0.37	44	123
12/27/63	30	1090	33.12	0.34	54	128
1/06/64	40	1369	32.99	0.27	64	133
1/16/64	50	1605	32.93	0 (Day 47)	74	136
1/26/64	60	1783	32.98	0	84	137
2/05/64	70	1896	33.16	0	94	136
2/15/64	80	1946	33.50	0	104	134
2/25/64	90	1943	33.99	0	114	130
3/06/64	100	1905	34.62	0	124	125
3/16/64	110	1851	35.33	0	134	119
3/26/64	120	1803	36.09	0	143	113
4/05/64	130	1778	36.82	0	153	106
4/15/64	140	1789	37.45	0	163	99
4/25/64	150	1845	37.93	0	172	92
5/05/64	160	1948	38.21	7.5	176	84
5/15/64	170	2098	38.29	0	168	77
5/25/64	180	2289	38.17	0	159	71
6/04/64	190	2512	37.90	0	149	64
6/14/64	200	2755	37.52	0	140	58
6/24/64	210	3002	37.11	0	130	53
7/04/64	220	3236	36.72	0	121	48
7/14/64	230	3441	36.40	0	111	45
7/24/64	240	3599	36.20	0	102	43
8/03/64	250	3701	36.16	0	92	43
8/13/64	260	3742	36.28	0	83	45
8/23/64	270	3723	36.56	0	73	48
9/02/64	280	3655	37.00	0	63	52
9/12/64	290	3553	37.54	0.1 (Day 295)	54	57
9/22/64	300	3433	38.15	0.21	44	63
10/02/64	310	3313	38.76	0.32	34	70
10/12/64	320	3212	39.31	0.37	24	77
10/22/64	330	3140	39.71	0.39	15	84
11/01/64	340	3108	39.93	0.41	5	91
11/11/64	350	3119	39.94	0.41	6	99
11/21/64	360	3175	39.72	0.39	15	106
12/01/64	370	3272	39.31	0.36	25	113

*See Figure 1

**See Figure 2

***See Figure 3

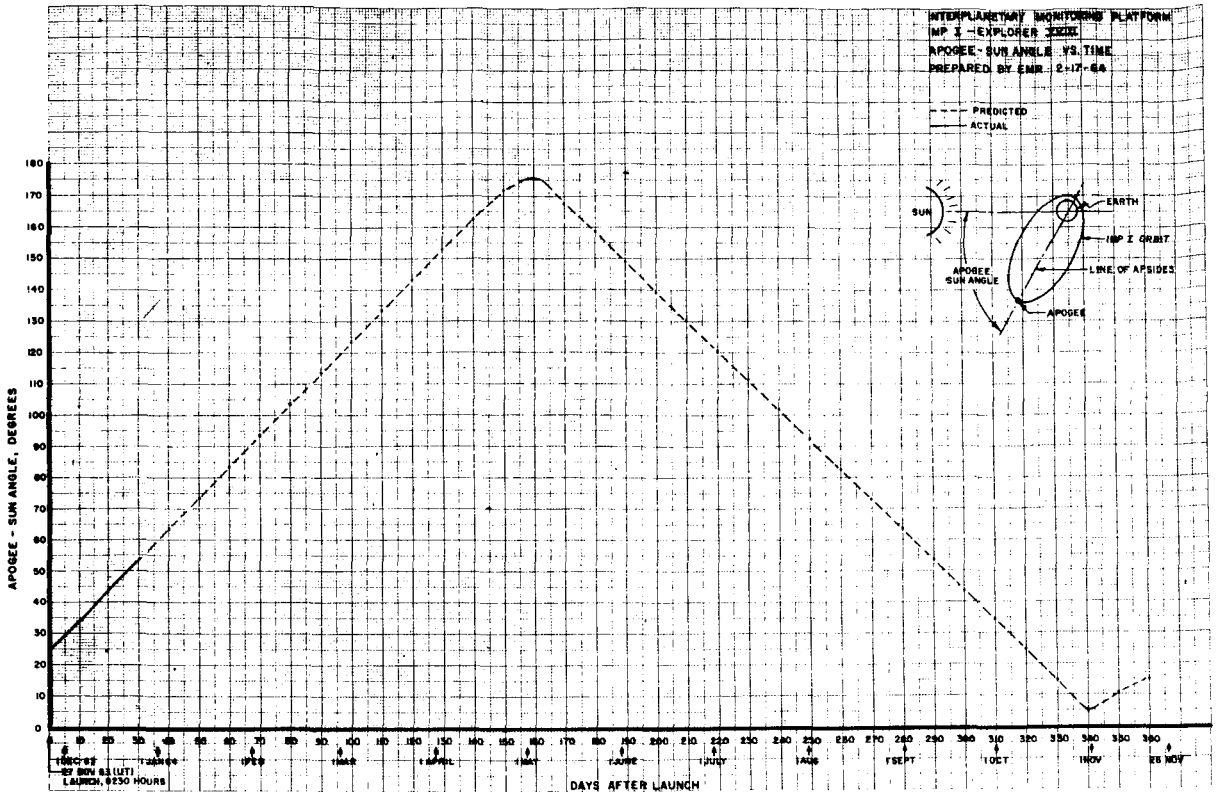


FIGURE 2.—Apogee-sun angle vs time after launch.

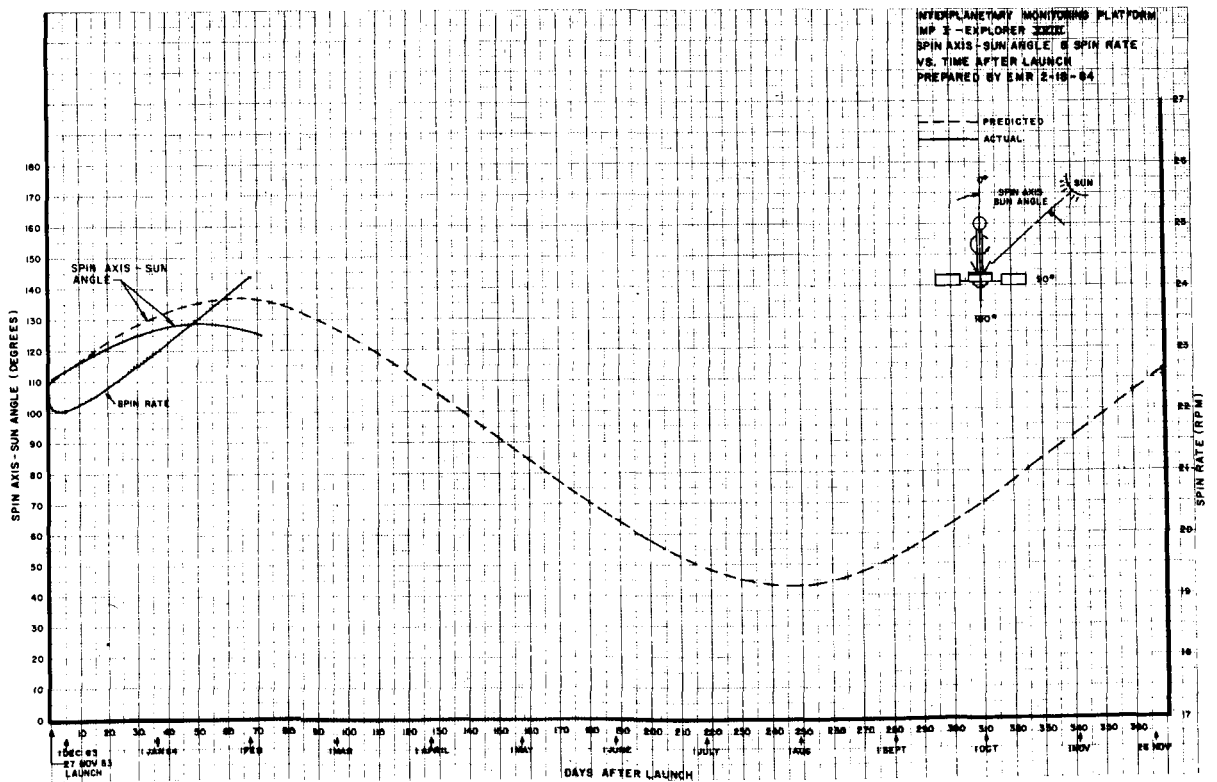


FIGURE 3.—Spin axis-sun angle and spin rate vs time after launch.

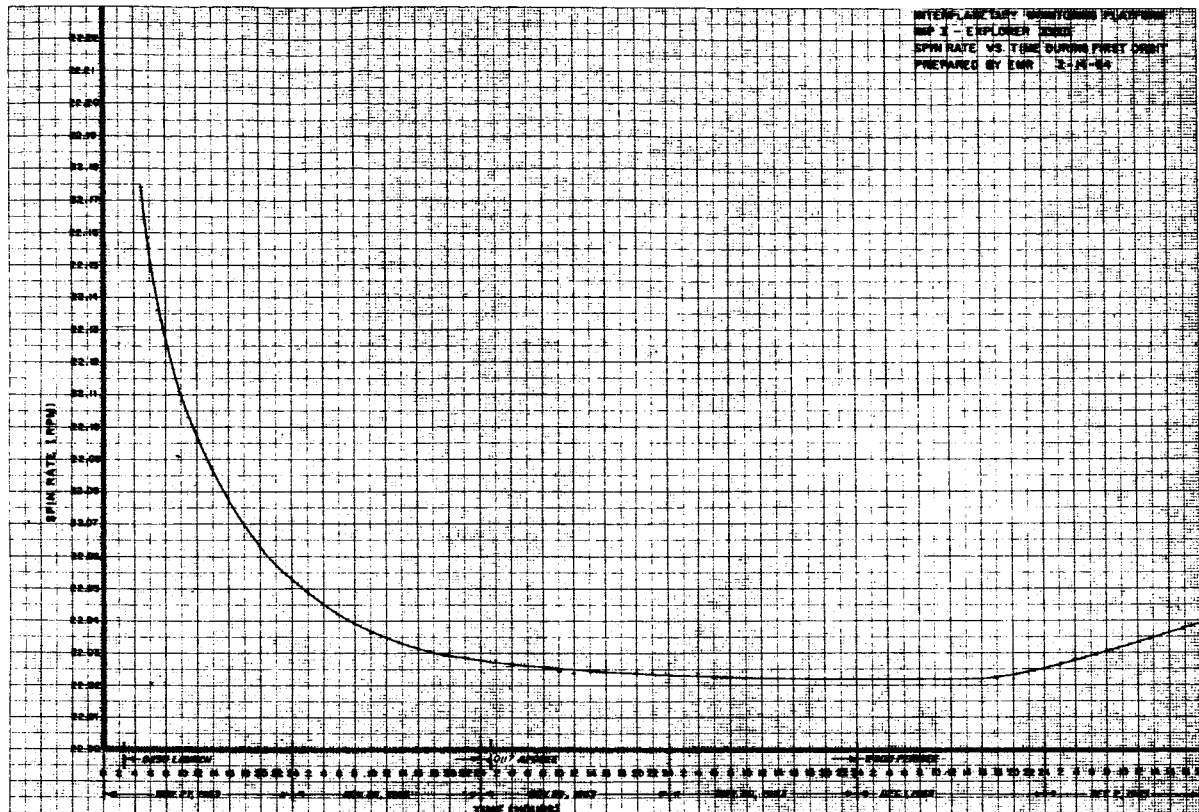


FIGURE 3a.—Spin rate-time after launch.

Reentry and disintegration should occur during one of the succeeding passes.

SPACECRAFT ATTITUDE

The spin axis-sun angle, as measured by the spacecraft optical aspect sensor, was 110° at launch (i.e., 20° below the spacecraft equator). For the first few weeks after launch, the sun has “moved” further away from the spacecraft equator reaching a maximum (total) angle of nearly 130° . After the second week of January this angle began to return toward the spacecraft equator.

The spin rate (Figure 3a), $2\frac{1}{2}$ hours after launch (earliest available data), was 22.173 rpm (nominal was 20 ± 2 rpm). During the first orbit, the spin rate decreased to 22.024 rpm ($93\frac{1}{2}$ hours after lift-off). Subsequent to the first pass, the spin rate has steadily increased, reaching 24.19 rpm some 68 days after launch (Figure 3).

The spin-up is attributed to solar radiation pressure acting on the solar paddles producing a

“windmill” effect. This phenomena was predicted and spin-up should continue as long as the sun is below the spacecraft equator. When the

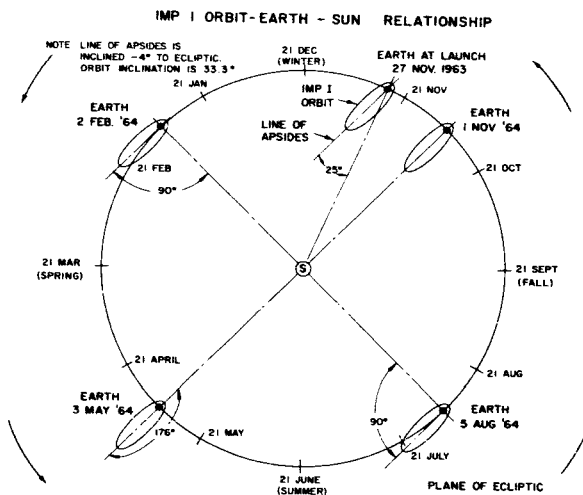


FIGURE 4.—IMP I orbit-earth-sun relationship.

sun begins to shine from above the spacecraft equator (~151 days after launch) a decrease of the spin rate should be observed.

The spin decay of >0.149 rpm during the first orbit was not anticipated and a rigorous theoretical explanation has not yet been retrofitted to the data (i.e., reasons unknown).

The right ascension and declination (calculated for the first orbit) of the spin axis, as measured

on the celestial sphere, was 115.3° and -24.8° respectively.

Data from the spacecraft indicates that there is *no* detectable precession or abnormal motion of the spin axis. Precession greater than $\sim 0.1^\circ$ could be observed with the IMP optical aspect system.

The following is a summary of the change of spacecraft temperatures due to the shadow:

Parameter	Just Prior to Shadow Entrance (°C)	At Shadow Exit (°C)	Ultimate Level (Due to Thermal Lag) (°C)	Gross Change (°C)
PP5 Skin Temp #1.....	+24.4	+18	+15.4	- 9.0
PP11 Skin Temp #2.....	+22.9	-3.1	- 3.1	-25
PP10 Solar Paddle.....	+11.4	-30	-33.3	-44.7
PP7 Battery.....	+42.1	+44.5*	+40*	- 2.1
PP14 Prime Converter.....	+36.1	+33.6	+31.8	- 4.3
PP15 Transmitter.....	+48.5	+43.1	+42.6	- 5.9

The curves of Figure 11 indicate a slight temperature rise (about 2-3 degrees at external locations) prior to shadow entry. This is probably explained by the fact that the spacecraft is rapidly approaching the earth at this time, hence, the thermal contribution from the earth's albedo increases, and results in this temperature rise.

Computations for the one-year period following launch indicate that an average of about 22 minutes shadow per orbit will be experienced during each of the first 12 orbits. Thereafter, 100 percent sunlight will exist thru 3 May 1964 (158 days after launch). The predictions for days 159 thru 163 are as follows:

Date	Day	Shadow Hours Per Orbit
4 May 64.....	159	5.3
5 May 64.....	160	7.5
6 May 64.....	161	7.8
7 May 64.....	162	8.2
8 May 64.....	163	3.8

Since the period of IMP's orbit is nearly four days, at least one, but not more than two, deep shadow periods will be encountered.

Thermal predictions for an eight-hour shadow period have been computed based on the following assumptions:

Spin axis-sun angle:.....	90°
Spacecraft recycle period:.....	8 hours
Spacecraft operating time in shade:.....	1½ hours
Non-operating time in shade:.....	6½ hours
Time in sunlight before first attempted spacecraft reactivation:.....	1½ hours

EXPERIMENT OPERATION

Each of the experiments on IMP I have provided abundant and excellent scientific data. With one exception, the experiments have continued successful operation throughout the first two months after launch.

The Thermal Ion-Electron experiment functioned properly for the first 20 hours following launch, producing excellent data. Thereafter, the mechanical programmer (which changes the experiment's mode of operation) began to function

*Due to high current discharge from the battery during shadow, the internal temperature increased $2\frac{1}{4}^\circ\text{C}$, but over-all spacecraft cooling decreased the battery temperature after this high discharge was removed (i.e., after shadow).

erratically, and only intermittent data was obtained for the next 3 days. The programmer then developed a somewhat repetitive pattern allowing the experimenter to recover approximately 10 percent of the data. This pattern has continued thru the present date.

At a scientific symposium, to be held at the Goddard Space Flight Center in March 1964, the IMP experimenters will present the first papers on the scientific results from IMP I.

SPACECRAFT PERFORMANCE

The performance of the IMP I spacecraft sub-systems has been excellent throughout the first two months of its life. The performance parameter data, with but two exceptions, have consistently remained at or near nominal values. (The two exceptions are discussed later in this section.)

The performance parameters monitored during Frame 2 of the telemetry format include four voltages, three currents and eight temperatures (Reference 3). Table 2 lists performance parameter and other data recorded at 32 minutes, 1 hour, 36 hours, 25 days and 60 days after launch. Figures 5 and 6 show the long term variation of the monitored voltages, currents and temperatures.

These parameters are also plotted for the complete first orbit (Figures 7 and 8) and for the eighth orbit (Figures 9 and 10).

Shadow Data

A shadow period (varying from approximately 10 to 22 minutes in duration, occurring near perigee) was experienced on each of the first several orbits. On the eighth pass, complete telemetry coverage was obtained throughout a 20 $\frac{3}{4}$ minute shadow beginning just before and

TABLE 2.—IMP I—Spacecraft Performance Parameter Data

Item	Time After Launch				
	32 Min.	1 Hour	1½ Days	25 Days	60 Days
Date (UT).....	11/27/63	11/27/63	11/28/63	12/22/63	1/26/64
Time (UT).....	0302	0350	1554	1033	0805
Tape No.....	16-002	16-002	20-004	001-005	01-026
Spin Rate, rpm.....		~22.18	22.03	22.593	23.91
Spin Axis-Sun Angle, deg.....	~110	~110	111	123	~129
PP1 System Voltage (19.6v nom).....	19.6	19.5	19.6	19.7	19.8
PP2 Prime Conv, +50v.....	50.2	50.2	50.2	50.9	51.2
PP3 Battery Charge, ma.....	600	267	16.2	18.5	30.5
PP4 Spacecraft Current, amps.....	1.94	1.96	1.85	1.81	1.84
PP5 Skin Temp #1 (Top), °C.....	+12.7	+14.5	+18.2	+21.1	+21.9
PP6 Rb Gas Cell Temp, °C.....	+35.0	+39.6	+45.5	+48.2	+49.0
PP7 Battery Temp, °C.....	+27.0	+28.0	+34.4	+40.9	+44.1
PP8 Prime Conv, +12v.....	12.0	12.1	12.1	12.3	12.3
PP9 Ave. Paddle Current, amps.....	3.86	3.83	3.70	3.48	3.10
PP10 Paddle Temp, °C.....	-5.8	+8.5	+9.3	+6.0	+4.5
PP11 Skin Temp #2 (Side), °C.....	+12	+14.1	+16.9	+19.2	+19.7
PP12 Multi-Conv, +7v.....	7.1	7.1	7.1	7.2	7.2
PP13 Rb Lamp Temp, °C.....	+93	+110	+109	+110	+109
PP14 Prime Conv Temp, °C.....	+28.9	+27.4	+31.9	+33.5	+34.7
PP15 Transmitter Temp, °C.....	+31.5	+32.9	+37.5	+46.9	+50.3
Average Paddle Power, watts.....	75.7	74.7	72.6	68.4	61.3
Average Spacecraft Load, watts.....	38.0	38.2	36.2	35.5	36.4

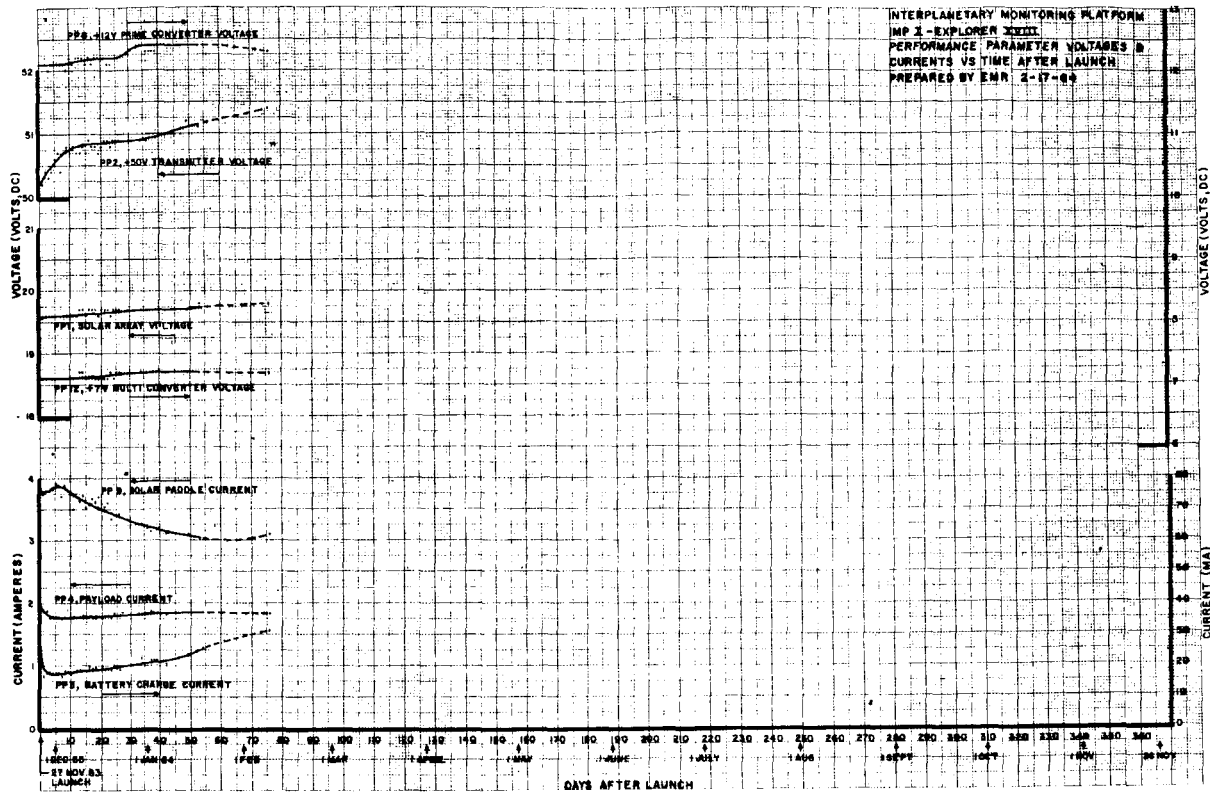


FIGURE 5.—Voltages and currents vs time after launch.

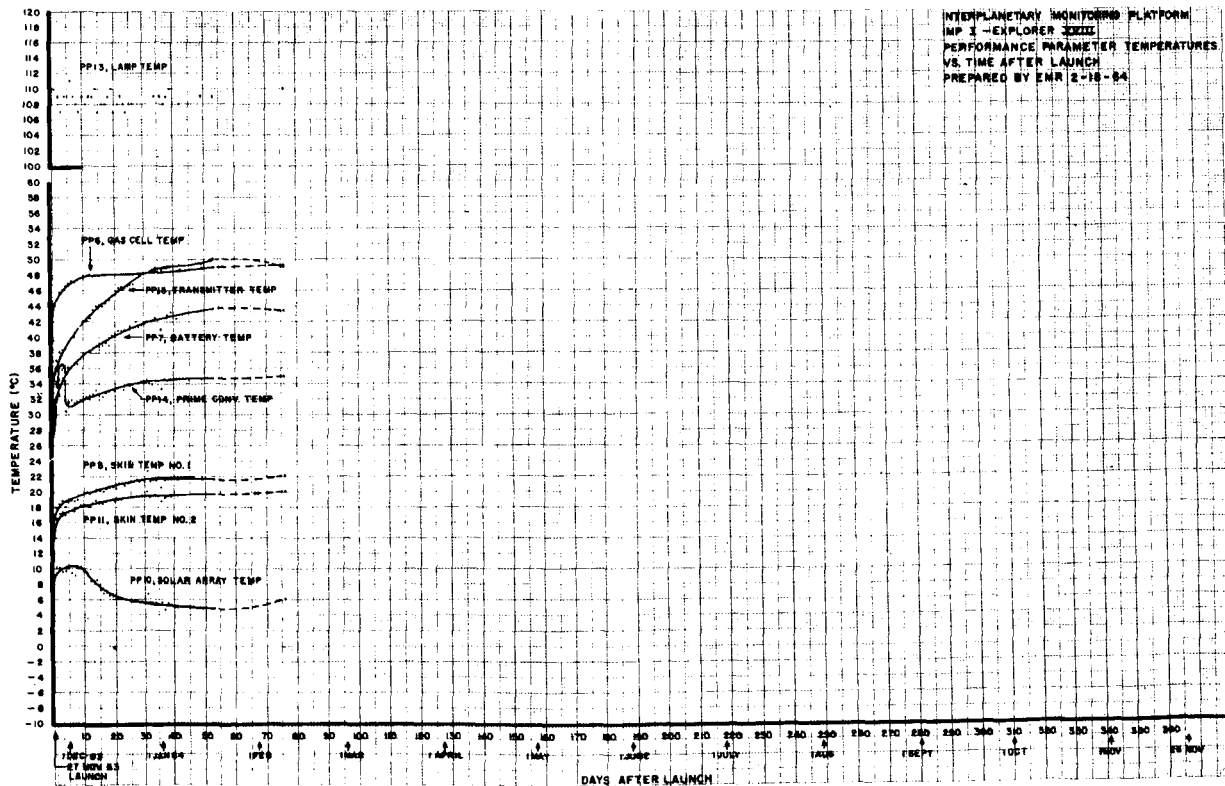


FIGURE 6.—Temperatures vs time after launch.

ending just after perigee. The temperature variations are plotted in Figure 11.

Expected temperatures for the Battery (PP7), Prime Converter (PP14), and Transmitter (PP15) at the end of 8 hours shade and after 1½ hours of subsequent sunlight are as follows:

	Temperature after 8 hours shade (°C)	Temperature after 1½ hours sunlight (°C)
Battery.....	-58	-52
Prime Converter....	-82	-65
Transmitter.....	-68	-59

The temperature of the solar paddles during the shade is expected to fall to roughly -150 to -200° C. However, based on previous testing*, it is believed that this will not seriously affect the operational capabilities of the paddles.

*A portion of a paddle, similar to that flown on IMP, satisfactorily survived a liquid nitrogen immersion.

Computations for a 3¼ hour shadow period result in the following predictions:

	Temperature after 3¼ hours shade (°C)	Temperature after 4½ hours sunlight (°C)
Battery.....	-23	-16
Prime Converter....	-40	-32
Transmitter.....	-35	-28

The effect of a long shadow on the over-all operation of the spacecraft should prove very interesting. It must be noted, however, that the spacecraft thermal design did not take into account extended shadow periods; hence, even partial survival will be considered to be somewhat of a "bonus".

Subsequent to this deep shadow condition, 100 percent sunlight is predicted from day 164 thru day 295. At this time, up to 25 minute shadow periods per orbit will be resumed and will continue

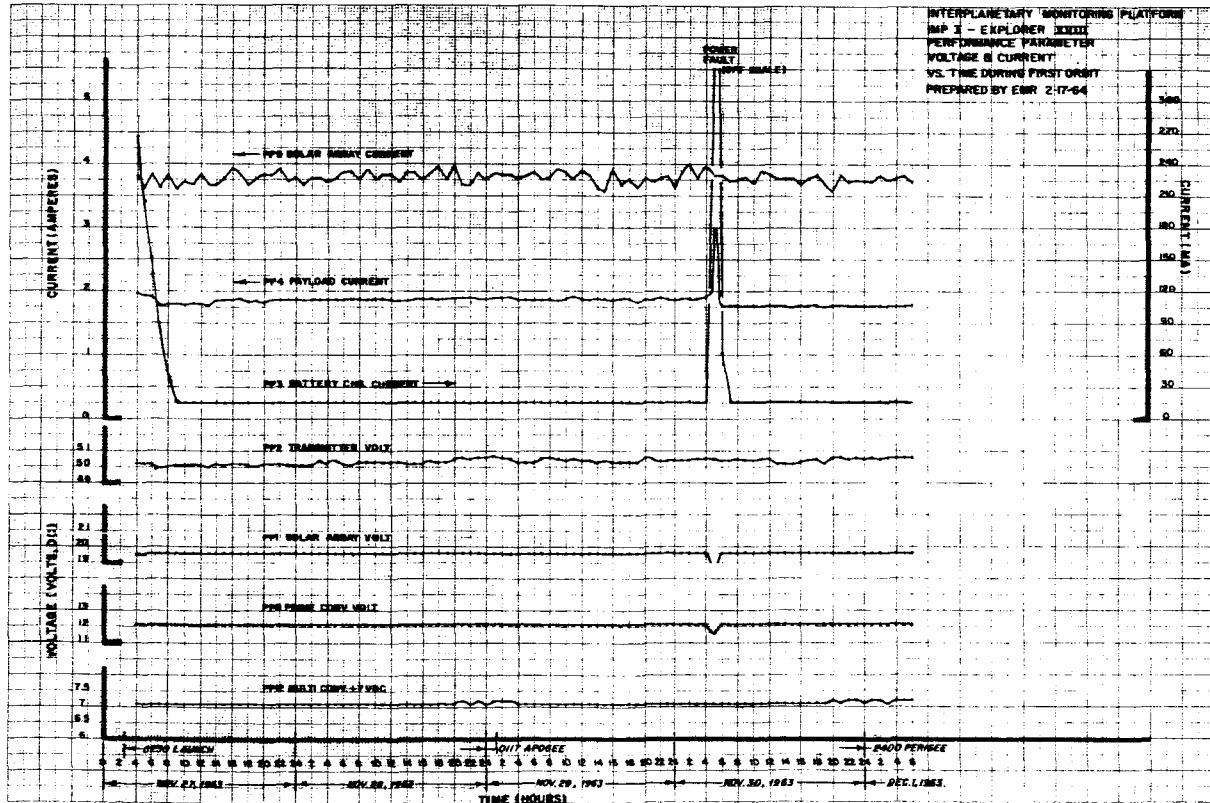


FIGURE 7.—Voltages and currents during orbit #1.

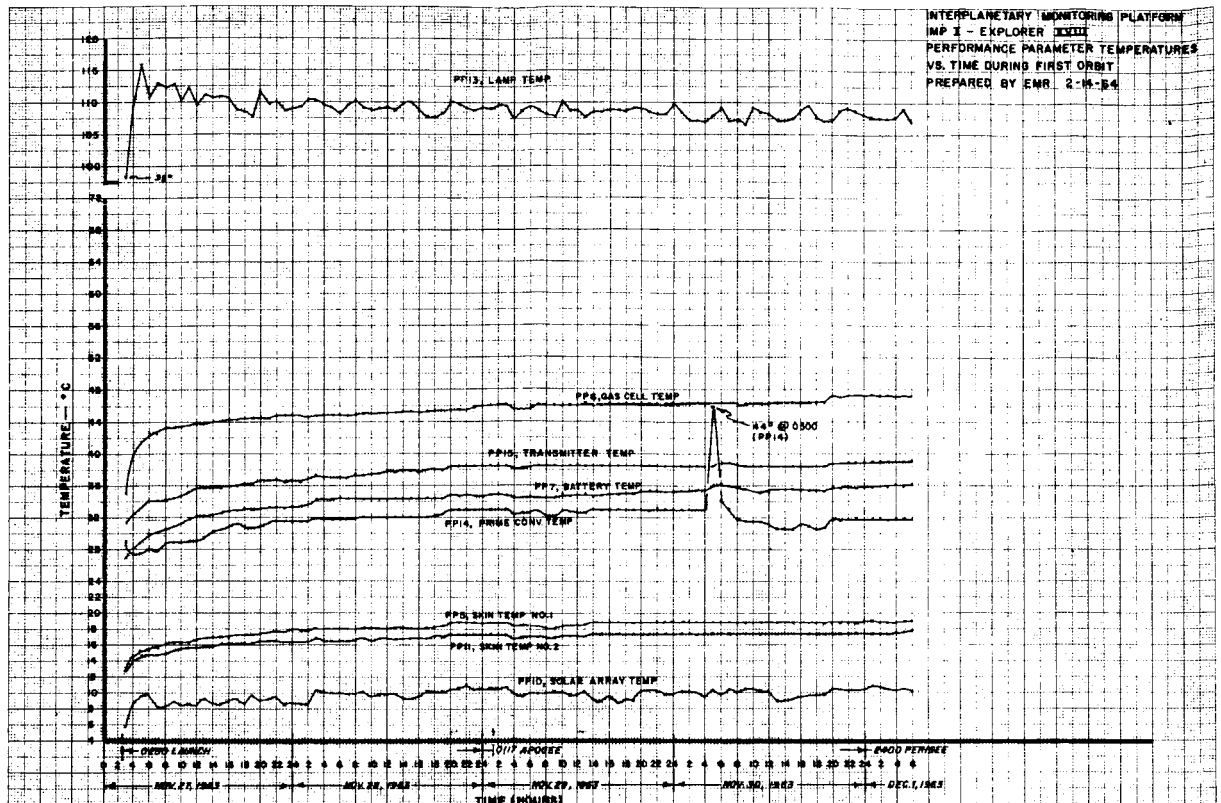


FIGURE 8.—Temperatures during orbit #1.

thru the end of the one-year period following launch.

Abnormalities in Spacecraft Operation

On November 30, 1963 at 0453 hours UT, the spacecraft experienced an apparent temporary "malfunction". At this time, the spacecraft load current increased by (at least) 50 percent, the system voltage decreased from 19.6 to 15.8 volts, the 12 volt regulated output decreased to 11.6 volts, and the prime converter temperature increased at least 11°C within 11 minutes after the first sign of "trouble" (Table 3). By 0504 hours, the spacecraft had resumed normal operation. Each experimenter has reviewed data before and subsequent to this period of time; all report *no change* in the operation of their experiment.

The only known difference in spacecraft operation is an apparent average decrease ($\sim 115\text{ ma}$ or $2\frac{1}{4}\text{ watts}$) in the spacecraft power requirements (Table 3).

It is believed that a partial short, in an unknown assembly, on the $+12\text{v}$ line caused the problem. Quite possibly, a component burning open, in a redundant or non-critical circuit may have removed the problem. The $2\frac{1}{4}\text{ watt}$ decrease in spacecraft power consumption is not understood at this time.

No recurrences of this type of problem have been observed.

A second, unresolved, problem area is the apparent upward "drift" of (at least) some of the data, notably the voltage parameters (Figure 5). It is not likely that each of the four telemetered voltages would drift simultaneously, in the same direction at roughly the same rate; hence, the accuracy of the data is being questioned. Since two voltages are telemetered through one analog oscillator, and the remaining two through a *second* oscillator, it seems reasonable to eliminate simultaneous drift of both oscillators. Another possibility—drift of the -2.7 volt supply to the analog

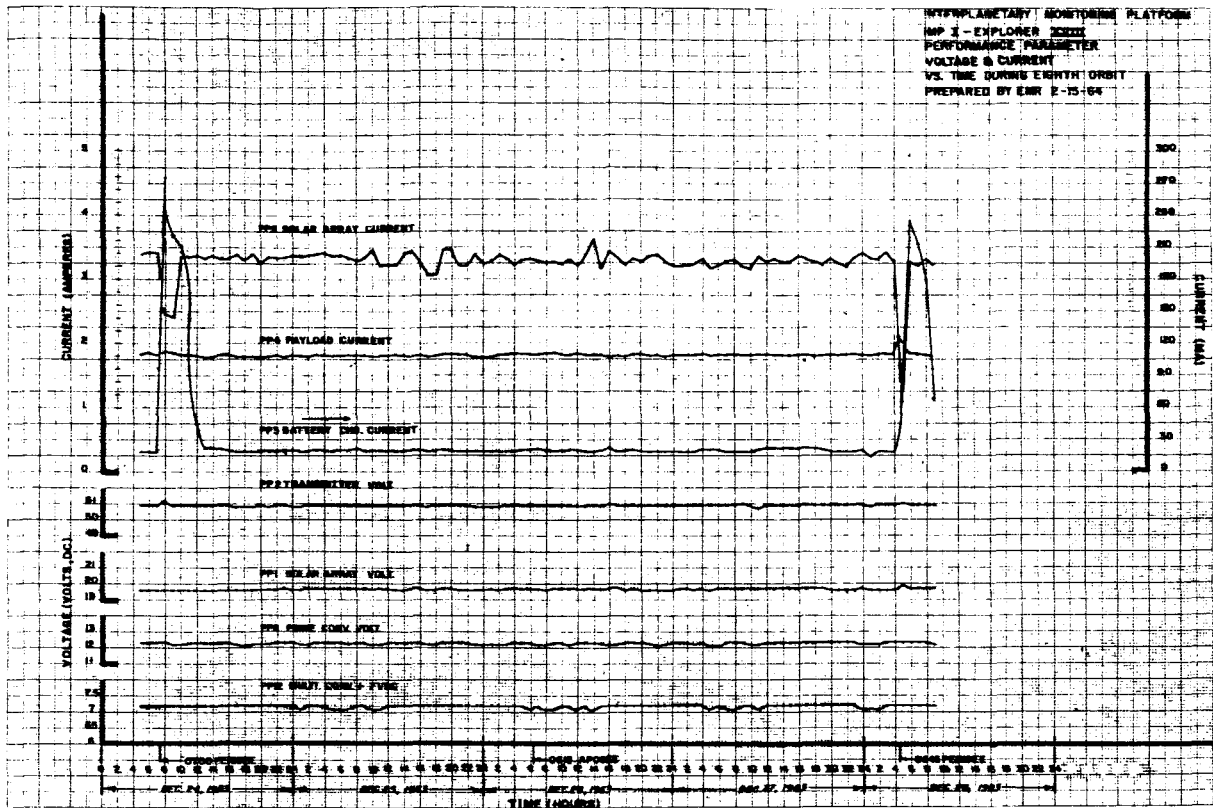


FIGURE 9.—Voltages and currents during orbit #8.

TABLE 3.—IMP I—Performance Parameter Data—(Spacecraft Problem of 30 November 1963)

Time* (11/30/63)	Sequence Count	PP1 (Volts)	PP2 (Volts)	PP3 (Ma)	PP4 (Amps)	PP8 (Volts)	PP9 (Amps)	PP14 (°C)
0450:53.664	3290	19.6	50.9	18.5	1.98	12.1	4.01	33.2
0453:37.492	3292	19.6	50.2	18.5	3.18	12.1	3.54	34.7
0454:59.406	3293	15.8	51.6			11.7	4.16	34.7
0456:21.320	3294	16.3	50.9			11.6	3.86	36.0
0459:05.148	3296							
0500:27.062	3297							
0501:48.966	3298							
0504:32.804	3300	19.6	50.2	502	1.85	12.2	4.01	44.0

*Times shown represent the time of Channel 1 (PP1), Frame 2. Data was lost for channels as shown.

PP4 Averages

	Seq 1	Seq 2	Seq 3
Average of 2 hrs data before problem, amps	1.997	1.885	1.893
Average of 2 hrs data after problem, amps	1.886	1.770	1.775
Net decrease, amps	-0.111	-0.115	-0.118

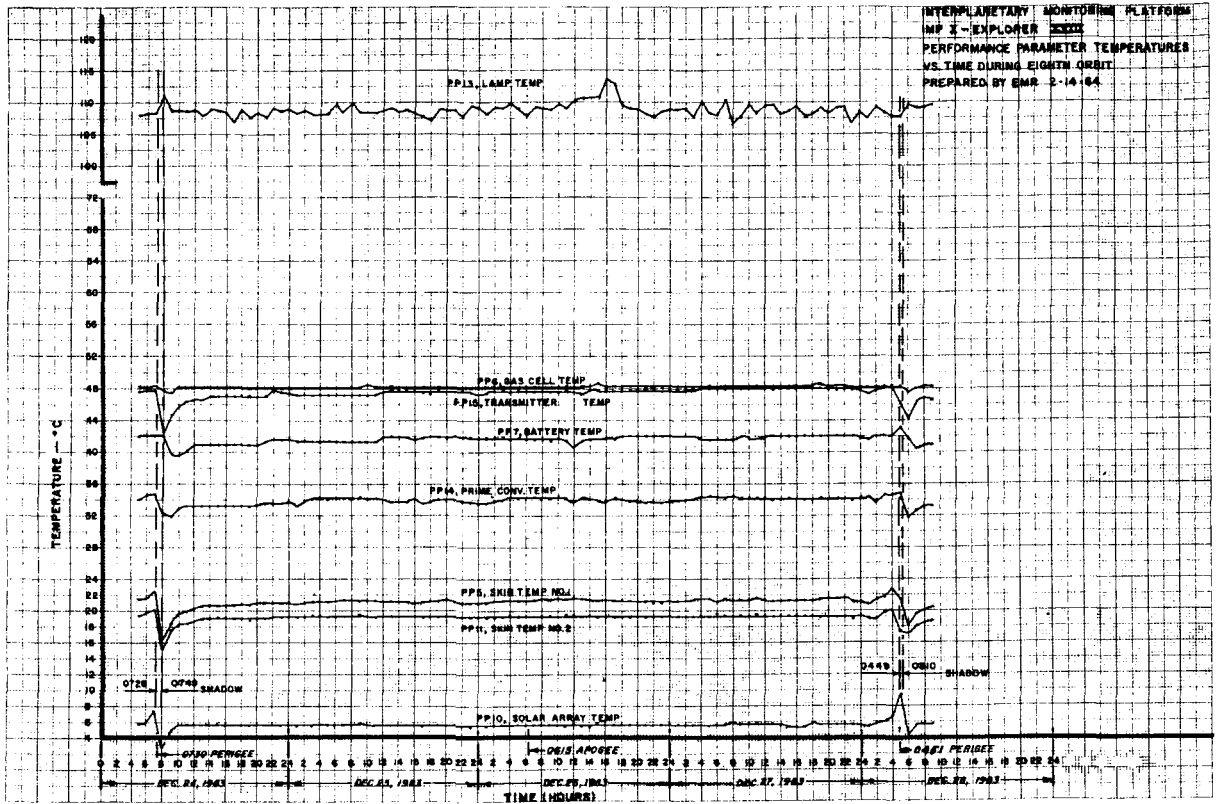


FIGURE 10.—Temperatures during orbit #8.

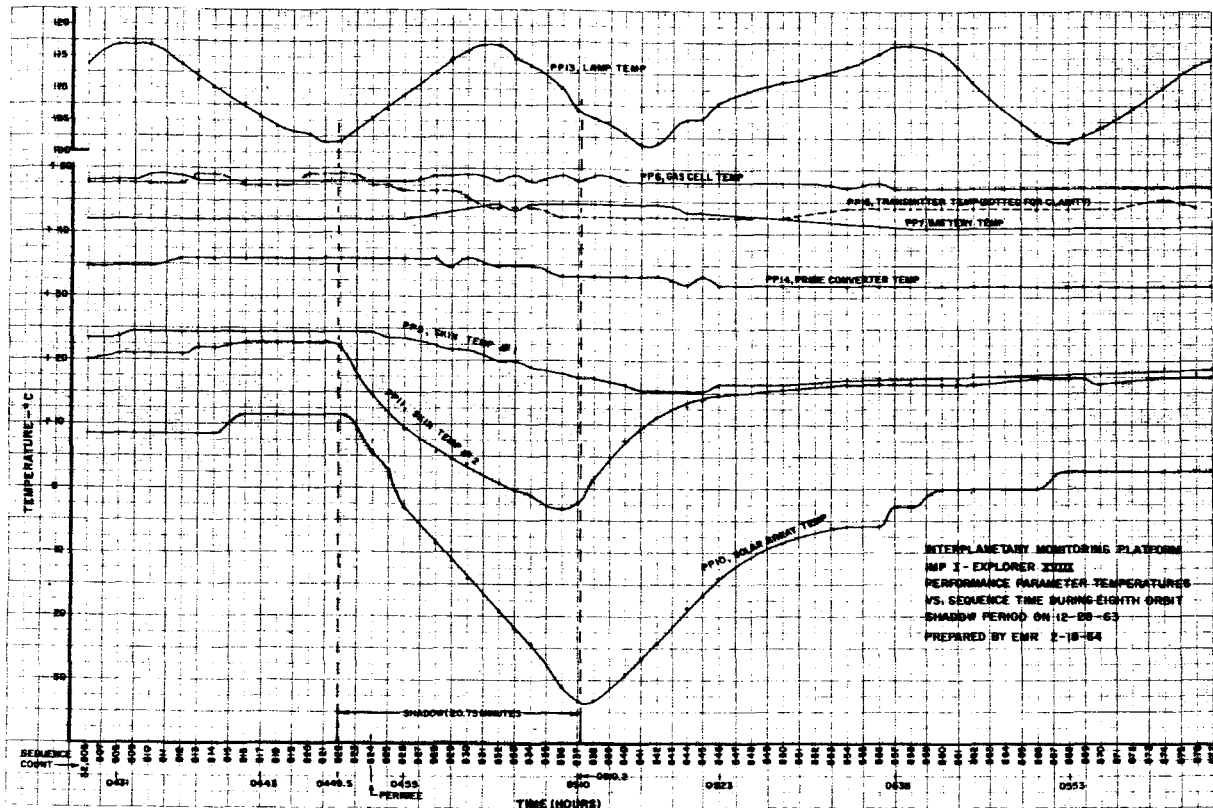


FIGURE 11.—Temperatures during shadow period of orbit #8.

oscillators was ruled out when a check of the 4.5 kc sync frequency (generated by a digital oscillator also dependent on the -2.7 volt line) showed no detectable change in the frequency. Hopefully, a satisfactory explanation of this ambiguity will be forthcoming in the not-too-distant future.

CONCLUDING REMARKS

The IMP I spacecraft has demonstrated extremely successful operation since its launching, nearly three months ago. The operation of the basic spacecraft subsystems (power, telemetry, etc.) as well as the scientific experiments has been quite satisfactory.

As the often quoted, critical, "90-day lifetime" limit is approached, all indications point toward continued successful operation. The spacecraft

carries no "killer timers" or other specially designed means of terminating operation, and so it is hoped that IMP will continue to operate for some time. However, it appears that the extended shadow periods (due, early May 1964) will present an almost insurmountable obstacle to the spacecraft electronics and could very well write the final epitaph for one of Goddard's most successful spacecraft.

REFERENCES

1. "Delta 21 Flash Flight Report," GSFC Field Projects Branch document, 26 November 1963.
2. "Field Flight Report, Delta 21," GSFC Field Projects Branch document, yet to be published.
3. CARR, F. A., "Performance Parameter Measurements on IMP I," GSFC Report X-672-63-238, November 1963.

ADVANCED RESEARCH FOR COMMUNICATIONS SATELLITES*

ROBERT J. MACKEY, JR.

To advance communications-satellite technology, the Goddard research program investigates technology in three areas: the ground systems, propagation path, and the satellites themselves.

GROUND SYSTEMS

In the ground systems we are investigating distortion and linearity in high-powered transmitters and the development of broad-band phase-lock receivers. We are evaluating the relative performance of high-powered klystrons and traveling-wave tubes for operating with either a frequency-modulated or a single-sideband carrier. The information obtained will provide data on high-power microwave tubes of various designs for multichannel wide-band communications systems.

The wide-band phase-lock receiver work has advanced the state-of-the-art substantially. This receiver reduces threshold and distortion, thereby adding a considerable margin of performance to a multichannel telephony communications system.

In the future we will investigate elements of ground terminals to improve over-all operating efficiency. These include low-loss transmission components and improved transmitting-antenna illumination, and lower antenna-systems noise. These studies will be done throughout the microwave- and millimeter-wave regions.

PROPAGATION PATHS

All of the communications satellite radio frequencies, to date, have been below 10,000 mc. This is because most common carrier and military equipment and techniques employ these frequencies. Because of overcrowding in the present frequency bands, it may be necessary to move to

higher frequencies. Atmospheric water-vapor absorption at 20,000 mc. and oxygen absorption at 70,000 mc. have been investigated. However, most of this work has been done within the atmosphere at spot frequencies of interest.

We would like to gain information on the use of millimeter waves between 15,000 mc. and 35,000 mc. for communications with satellites.

ACTIVE SATELLITES

For the active type of communication satellite, work is going on in several areas. Work was begun last year to develop a low-noise microwave front-end for general application to satellite repeaters. Receiver noise figures of 5-6 db have been achieved with tunnel-diode amplifiers and mixers at 6000 mc. At present the effects of space environment on these components are being investigated thoroughly and initial results obtained so far are encouraging although the investigation is far from complete.

Additional effort is getting underway to explore the use of "cold" cathodes. These do not require heat for the release of electrons but rely on "tunneling" effects in solid-state materials. Such cathodes have already been employed in traveling-wave tubes but under pulsed conditions only. The object here is to first investigate cathode materials suitable for continuous-wave use; then to develop higher efficiency traveling-wave tubes (TWT) at communications frequencies.

MICROWAVE CONVERSION

For about the past six months, methods of directly converting one microwave frequency to another without resorting to an intermediate frequency have been under investigation. Two methods are being explored; one using a "serro-

*Published in *Electronics World* 71(6):64-69, June 1964.

dyne" technique and the other a re-entrant TWT (traveling-wave tube) technique.

In the serrodyne method, the TWT slow-wave structure is modulated by a linear saw-tooth at the difference frequency. Hence, the incoming signal is linearly phase-shifted along the slow-wave structure at the rate of the saw-tooth modulating frequency. The TWT output frequency is the output frequency shifted by the saw-tooth frequency with the amplification available in the TWT. If the saw-tooth is not linear with negligible flyback, undesired harmonics are generated.

The second approach involves passing the input frequency through a TWT which amplifies it. Down conversion of this frequency then takes place by conventional local-oscillator/mixer techniques to the desired frequency to be transmitted. This frequency is then amplified by passing it back through the TWT. This involves some tricky microwave hybrid networks, filter design, and broad-band TWT's. A system of this nature has been developed and is about to be tested.

Other phases of development include microwave frequency generation using varactors and tunnel diodes, as well as microwave filter design for harmonic isolation and frequency selection. Other

solid-state devices are being explored for possible use as power generators at microwave frequencies. New types of electronic components are tested when they become available.

SPACECRAFT ANTENNAS

In the area of spacecraft antennas, a study of all electronic beam shaping and steering techniques was initiated last fall. This applies to the microwave- and millimeter-wave region. The study will include prototype spacecraft antennas, having high gain with self-steering and remote-steering features. The most promising techniques will be selected and developed to demonstrate their feasibility. The purpose here is to form an antenna beam in a satellite, point it at a desired ground station or stations, and tracking the station(s) whether the satellite or station is moving, regardless of spacecraft attitude. In addition, such antennas provide spacecraft-to-spacecraft communication links.

Finally, work is also being done in the advanced passive satellite area. Improved lightweight structural materials, components, plastic removal, and erection techniques are under study.

II. SPACE DYNAMICS AND CONTROL SYSTEMS

A LIGHTWEIGHT FOLDING BOOM FOR ACCURATELY PLACING SCIENTIFIC EXPERIMENTS UP TO 25 FEET FROM LARGE SPACECRAFT*

JOHN P. BAUERNSCHUB, JR.

A feasibility study was conducted on large folding booms for spacecraft applications. The resulting design featured precise repeatability, permitted the passing of a large electrical cable directly through each knuckle, and used a constant-speed spring-driven motor to operate the knuckles by means of a cable. The entire assembly, including the spring motor, is completely nonmagnetic and is ideally suited for placing a magnetometer experiment a considerable distance from the spacecraft structure. It can also place many other instruments, such as antennas and solar-celled paddle systems, away from the structure.

INTRODUCTION

Certain experiments flown on spacecraft are susceptible to contamination or interference by the spacecraft operational system, by other experiments, or by both of these. One way to overcome this problem without imposing undue restrictions on other spacecraft components is to place the sensitive experiment some distance from the main body. If the required separation is much longer than the body of the spacecraft, the use of a multisection boom appears logical. This report presents the results of a study to develop a folding boom explicitly for such applications (Figure 1).

DESIGN CRITERIA

This boom system was designed to meet the non-magnetic, high strength, and alignment requirements imposed by several experiments and it incorporates the following features:

1. Power to unfold the boom in subzero temperatures, with a large electrical cable passing through the knuckle;
2. Slow speed, approximately $\frac{1}{2}$ rpm for each joint, to prevent whiplash of the experiment when the last section of the boom opens;
3. Positive locking of the knuckles to prevent any relative motion after closing, and repeatability of the closing position to give the highest degree of accuracy possible;

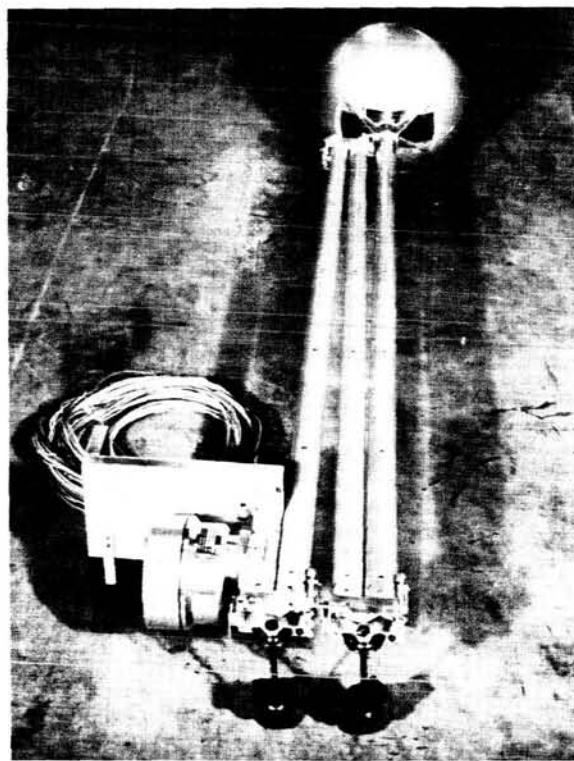


FIGURE 1.—A 12 ft experimental folding boom in folded condition with a hermetically sealed, nonmagnetic spring motor.

4. Reliability through simple design and redundancy where possible;
5. Exclusive use of nonmagnetic materials;

*Published as NASA Technical Memorandum X-941, January 1964.

6. Flexibility in conforming to the spacecraft shape and allowable volume under the shroud;
7. Easy adaptability to any desired unfolding sequence.

BOOM OPERATION

The sequence of events for erecting the boom are initiated by an electric impulse from the spacecraft which starts the motor. The motor unfolds the boom by reeling in, at a constant speed, a wire cable which applies a high torque to each joint. This cable is fastened to the last section of the boom, goes around a pulley in each knuckle, and is fastened to a reel on the motor. The pulley in each joint is free to rotate, so that the cable can move easily when other knuckles are closing.

With the motor started, the normal sequence is the successive opening of each joint beginning with the one closest to the spacecraft. As one knuckle closes and locks, the next section unfolds. A microswitch in each knuckle signals when the joint is closed and turns off an electric drive motor (if one is used) when the boom is fully opened. If a spring motor is used the motor stalls when the boom is fully deployed.



FIGURE 2.—Boom deploying. This illustrates the feature which prevents an experiment from hitting the boom during opening.

BOOM CONFIGURATION

The boom can be designed in almost any configuration or length to adapt to the spacecraft body and booster fairing. The sequencing of knuckle operation can be in any order, even to the point of having a joint close in steps with other knuckles closing between these steps. This would be done to prevent the experiment from hitting the boom or spacecraft while unfolding. An example of this is illustrated in Figure 2 where the first knuckle has closed and the second is about to lock. However, the third knuckle has already opened a sufficient distance to prevent the sphere from hitting the first section of the boom as the second knuckle closes. This opening of the third knuckle occurred when the motor was started, but further motion has been restricted by a cable to the sequencing device on the second knuckle. Figure 3 shows the boom fully deployed.

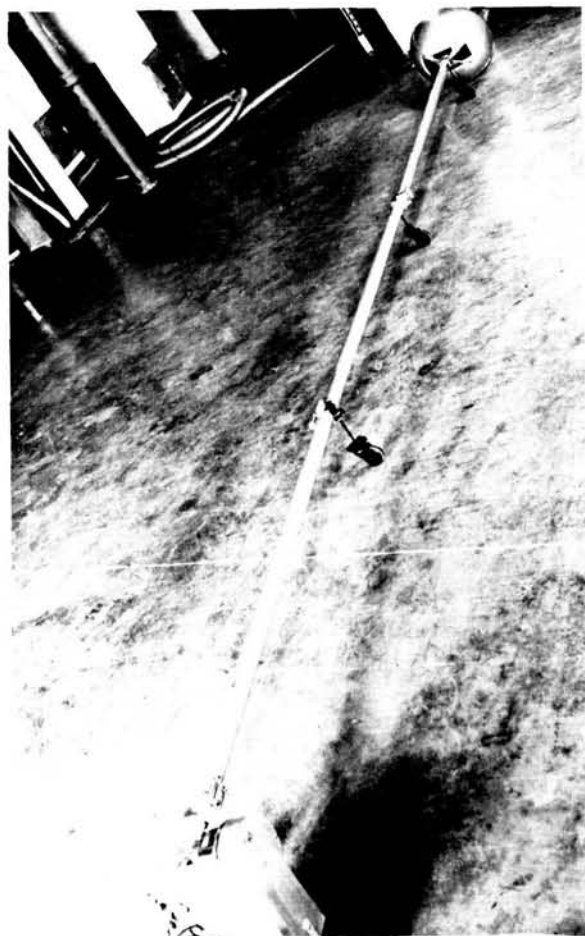


FIGURE 3.—Boom fully deployed.

Possible variations of the boom configuration include:

1. The motor can be mounted inside the spacecraft, on the boom before the first knuckle, or on any other section of the boom;
2. The sections need not be of the same length;
3. The sections do not have to be folded in the same plane.

TRIGGERED KNUCKLE LOCK

Figures 4, 5, and 6 show the details of another feature of the boom. The accuracy of this unique lock is not dependent on the clearance allowed for the knuckle pivot, or on the precise location of any of its parts; yet it will repeatedly lock the knuckle in the same position and prevent any relative movement between the knuckles of a joint.

The operation sequence of the triggered knuckle lock starts when the joint is almost closed and "B" knuckle pushes the trigger in "A" knuckle. When the trigger pin is withdrawn from between

the spring retaining arms these arms release compressed springs. However, just before the springs are released, the "T" catch on "B" knuckle drops over the wedge on the claws on "A" knuckle. When the springs are released, they pull the claws which wedge in slots of the "T" catch, closing the knuckle and firmly locking it. The knuckles are guided into place by a tapered pin in "A" knuckle which wedges into a hole in "B" knuckle. The "B" knuckle is also positioned and locked by the wedge of the claw seating in the "T" catch. The "T" catch itself is wedged into place in the "B" knuckle. An additional feature of this design is that only about 5 lb of tripping force is required to bring 100 pounds of locking force into play.

A sequencing device attached to the side of the knuckle works in a similar way. It begins to actuate as the knuckles close and pushes a pin. This pin moves a slotted block off a hook attached to the next section to open. This release occurs just as the knuckle is locked. A microswitch is

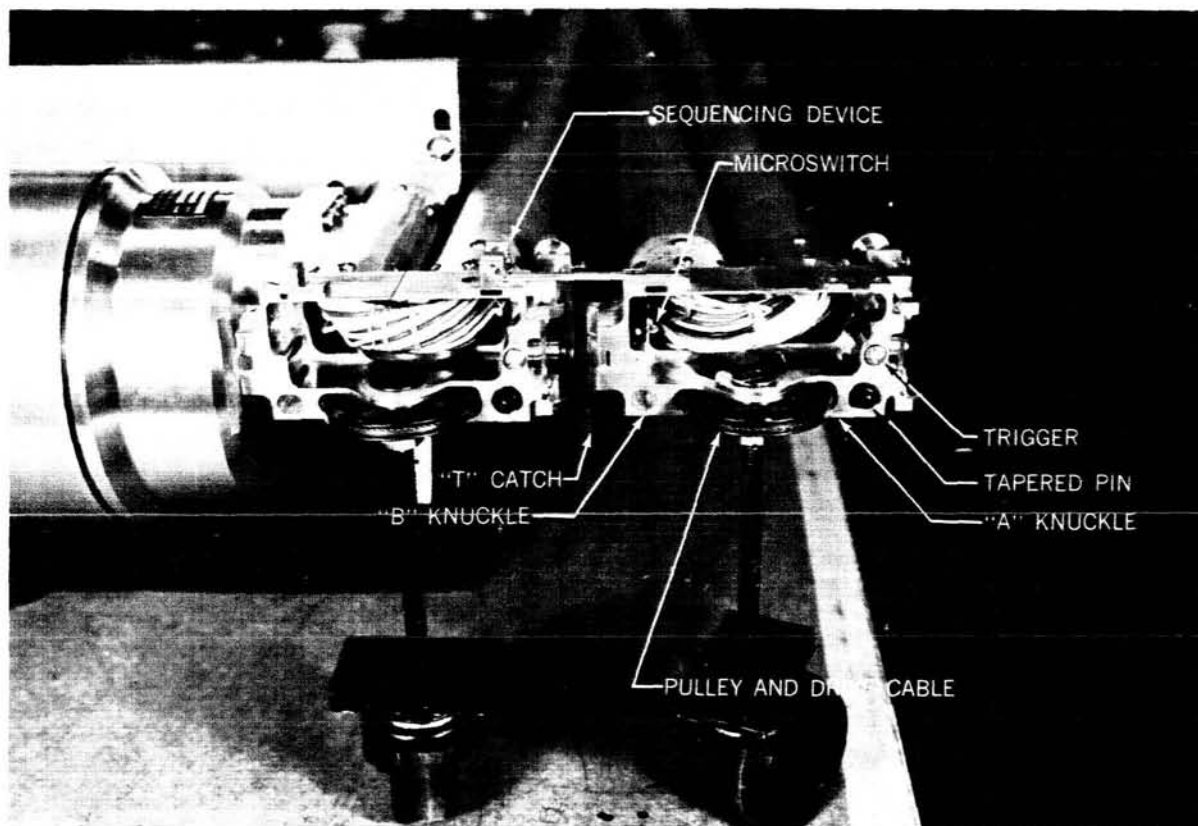


FIGURE 4.—Details of the triggered knuckle lock and internal passage of the electrical cable.

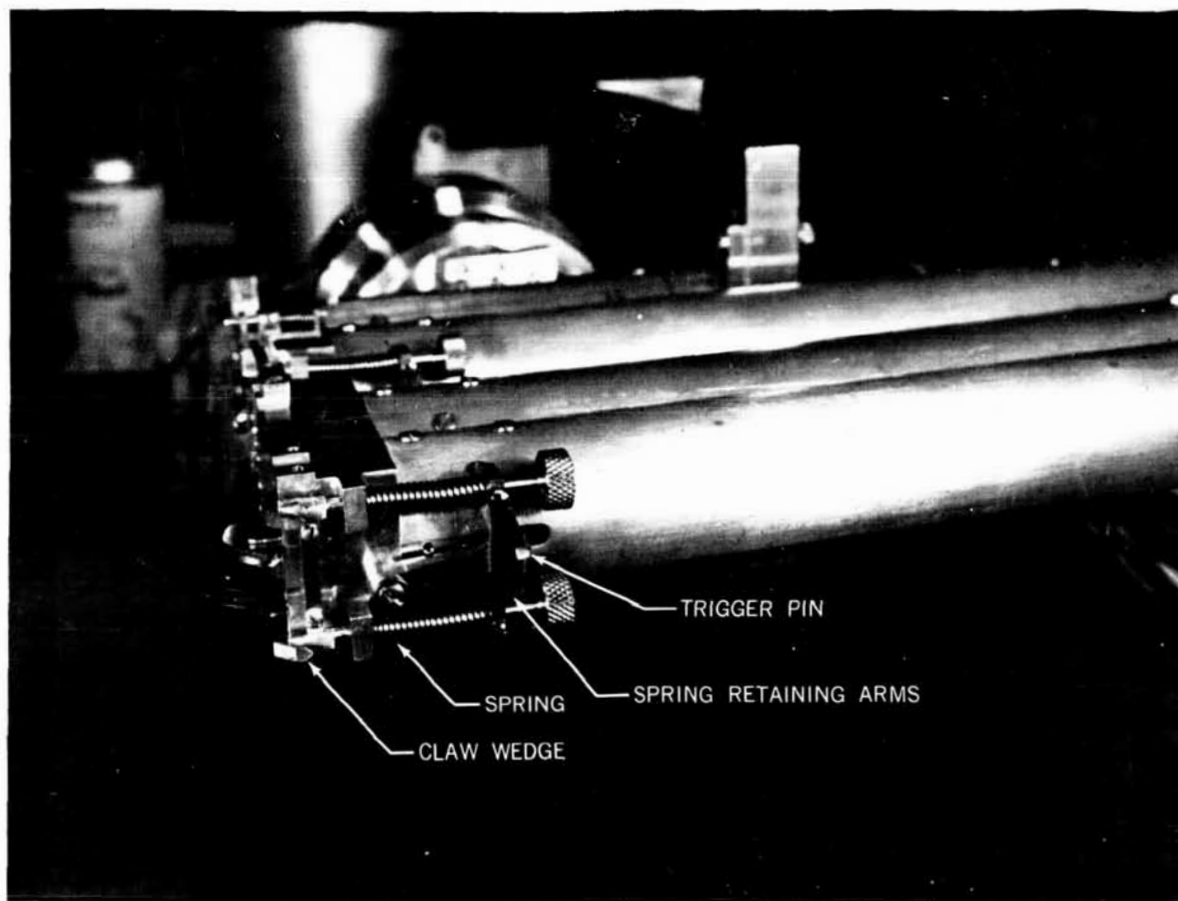


FIGURE 5.—Triggered knuckle lock cocked.

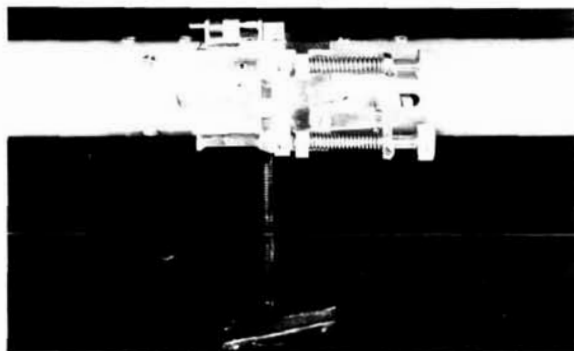


FIGURE 6.—Triggered knuckle lock in closed position.

built into each knuckle to signal when the joint is closed.

MOTORS

Either of two motors are used for operating this boom, depending on the particular case. Figure 7

shows a completely nonmagnetic spring motor developed specifically for this application. Its governed output supplies sufficient torque to open a four section boom with a bundle of 46 electrical wires at 0°F. Electric dimple motors are used for starting the mechanism. The total weight of the assembly is less than 2 lb.

The other motor which has been used is a 24 volt, wound field dc motor (Figure 8). It requires about 25 watts and weighs 15 ounces. The permanent magnetic field of this motor is 72 gammas at 1 foot. If magnetic contamination is not a problem, an even lighter, permanent magnet dc motor can be used.

WEIGHT

The data in Table 1 is given in order that an estimate can be made on the weight of a particular boom. A four section, 25 foot boom using 1 inch

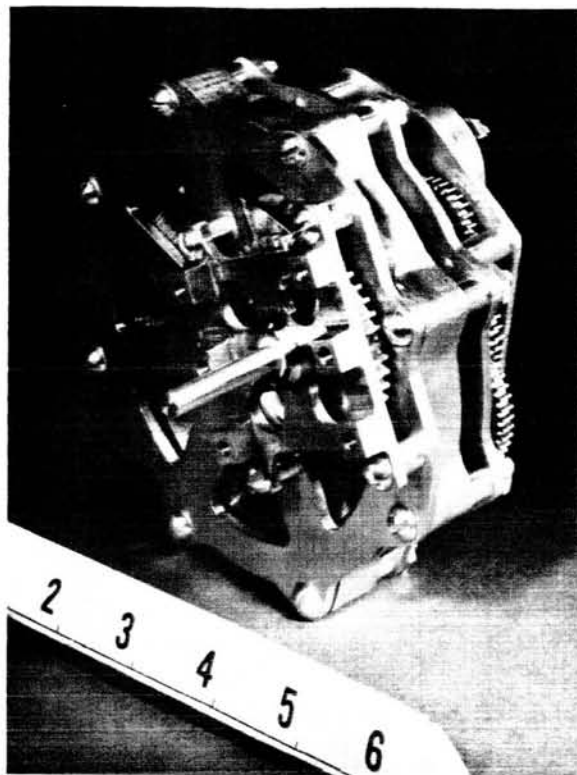


FIGURE 7.—Nonmagnetic, constant speed spring motor. (The measurement is in inches.)

diameter aluminum tubing would weigh 4.4 lb with an electric motor and 5.4 lb with a spring motor.

EXPERIMENTAL BOOM

The 12 foot, three section boom shown in the photographs was built for determining the torques required for such a boom, with the electrical cable passing through the joints, to unfold, and for determining the relative horizontal (axial) position accuracy of the boom. The sections are 1½ in. O.D. aluminum tubes with a 0.020 in. wall.

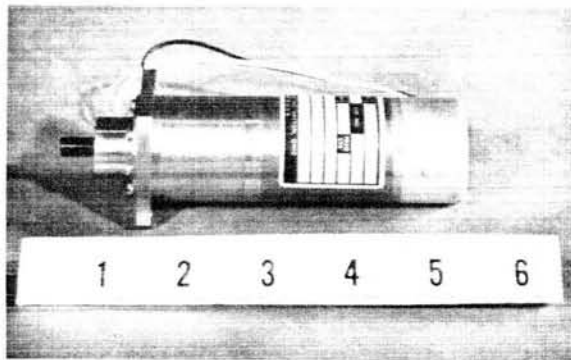


FIGURE 8.—Wound field dc motor. (The measurement is in inches.)

Table 1

Boom Instrumentation

Item	Material	Weight
Knuckles	Aluminum	0.35 lb each
Tubes	Aluminum, 1-1/2 in. O.D. x 0.020 in. wall	0.11 lb/ft
	Aluminum, 1 in. O.D. x 0.020 in. wall	0.074 lb/ft
Spring motor	Aluminum, magnesium, and elgiloy	2 lb
dc motor (wound field)		0.94 lb
Spacecraft interface and motor mounting pad	Aluminum	0.25 lb

The spring motor shown in these pictures is an earlier version that was hermetically sealed. Present knowledge of lubricants and the brief operational life of the motor eliminates the need for the seal.

As a result of having the knuckle powered by an external source there is sufficient clearance to allow passage of a large electrical cable through the knuckle. This results in a more compact design and the best way for securing the cable loop against launch vibration. Thermal control of the electrical cable is also facilitated by this feature.

The electrical cable used in this boom consisted of 33 strands of No. 20 Teflon-covered wire and 12 No. 22 shielded strands. Figure 4 shows the cable passing through the knuckles when the boom is in the folded condition, and Figure 9 shows its position when the boom has unfolded. This cable weighed $\frac{1}{4}$ lb/ft and was lashed to an aluminum strip fastened to the tubing. The purpose of this strip was to secure the electrical cable during vibration.

To provide a completely nonmagnetic boom, all springs and the wire cable were made of elgiloy, and all other parts were aluminum. This was also true for the spring motor. The small residual magnetic field of the wound field dc motor would be negligible to a magnetometer placed 25 feet away.

The boom is supported by castors attached to the pivot bushing of each joint and high enough

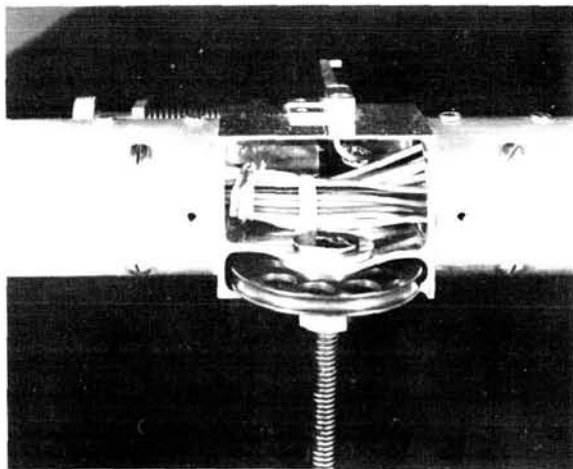


FIGURE 9.—Electrical cable passage through the closed knuckle.

to allow the sphere to clear the ground. The sphere on the end of the boom is a bias sphere for a rubidium vapor magnetometer.

Tests using the spring motor shown in Figure 7 were run for determining the torque required to open the boom at different temperatures. However, the data obtained were inaccurate because of the opposing torque introduced by the friction of the castors and their 4 foot moment arm. Below 0°F the castors sometimes failed to align in the direction of motion. Even with this handicap, the spring motor was capable of extending the boom. The load presented by the electrical cable pack was also extreme, since an actual experiment would use No. 26 or 28 wire for the purpose of saving weight.

Because of the problems encountered with the castors, it was impossible to make the desired accuracy checks. A system to perform such a test has been developed in which a boom is supported by air bearings on an exceptionally smooth surface; but this is not yet available at the Goddard Space Flight Center.

Nevertheless, a study of the knuckle design indicates that such a deployed boom is structurally a solid member. Therefore, the tip position will be precisely repeatable when the boom is in equilibrium under similar external forces, provided the boom has not been stressed beyond the elastic limit.

This points out the main advantage of a cable-driven boom. Since it deploys at a slow, constant speed, the bending moment imposed on the boom by the sudden stop of the last knuckle is much lower than that caused by a spring-driven boom, which is continuously accelerating until stopped. For this reason a cable-driven boom can be lighter or can support a larger experiment than a spring-driven boom.

CONCLUSION

Because of its ability to precisely orient, the boom system described in this report is suitable for use on a highly stabilized space platform whose attitude, in relation to some known reference, can be accurately determined. The basic design can also be applied in the construction of a boom to place other instruments, such as antennas and solar-celled paddle systems, away from the main body of a spacecraft.

VIBRATION ISOLATION OF SATELLITE TAPE RECORDERS*

JOSEPH H. CONN

A standardized method of shock and vibration isolation for satellite tape recorders has been developed. Using readily available components, the isolators combine small size, light weight, and damping characteristics with good load-carrying and elastomeric properties. This system is now being used in the International Satellite UK-2 and will be used in the Orbiting Solar Observatory (S-57 model). It may, however, be applied to many other components, such as electron tubes, bearing assemblies, and sensitive electronic assemblies.

INTRODUCTION

The early tape recorders designed for NASA satellite applications had extremely low survival rates when subjected to satellite design vibration tests. The majority of the failures could be accounted for in the following ways: (1) The high dynamic acceleration levels in the medium and high frequency ranges caused surface fatigue of the bearings in the motors and capstan housing. (2) When the pinch roller assemblies resonated, the acceleration loads were so great that tension would be released from the tape causing it to loop and occasionally to tangle.

As the development of the satellite recorder progressed toward standardization of components, it was apparent that a vibration isolation system should also be designed, if possible, with standard, readily available components. A standardized recorder and standardized mounting hardware would allow greater design flexibility and permit interchange between programs.

The first approach to the vibration problem was to increase the design load requirements on the bearings and springs of the recorder. However, this was impractical because the technological and material requirements were too stringent. The bearing assemblies became impossible to standardize because of the constantly changing loads. Furthermore, increased preload requirements on the pinch roller springs increased their physical dimensions. The preload also increased the drag torque on the drive system,

requiring an increase in both motor sizes and power requirements. It became apparent that the tape recorder would have to have a filter for the environment which it experienced. A shock and vibration isolator was carefully selected which had a low resonance frequency and was compatible in size with the tape recorders.

DESIGN CONSIDERATIONS

The particular design discussed here was used in the International Satellite UK-2. A similar design is being used in the Orbiting Solar Observatory S-57. The UK-2 tape recorder shown in Figure 1 was the first commercial modularized tape recorder designed by the Recording Techniques Section at GSFC. The complete tape recorder is enclosed in an envelope 7.00 inches in diameter and 3 inches in height with a total weight of 4 lb or less. The container is pressure-tight to eliminate bearing lubrication problems and motor arcing in a high vacuum environment. The UK-2 satellite was subjected to a vibration environment 1.5 times greater than the expected flight environment. However, since the satellite structure was a multidegree-of-freedom elastic body the input levels were amplified, the major amplification occurring at the structure's first mode. For 70 to 250 lb satellites the first mode usually occurs between 75 and 200 cps. Table 1 indicates the design or prototype vibration levels for the tape recorder assembly. In addition to the sinusoidal levels, the recorder was subjected to Gaussian random vibration in the three orthogonal axes. These levels were of 0.07 g²/cps

*Published as *NASA Technical Memorandum X-942*, February 1964.

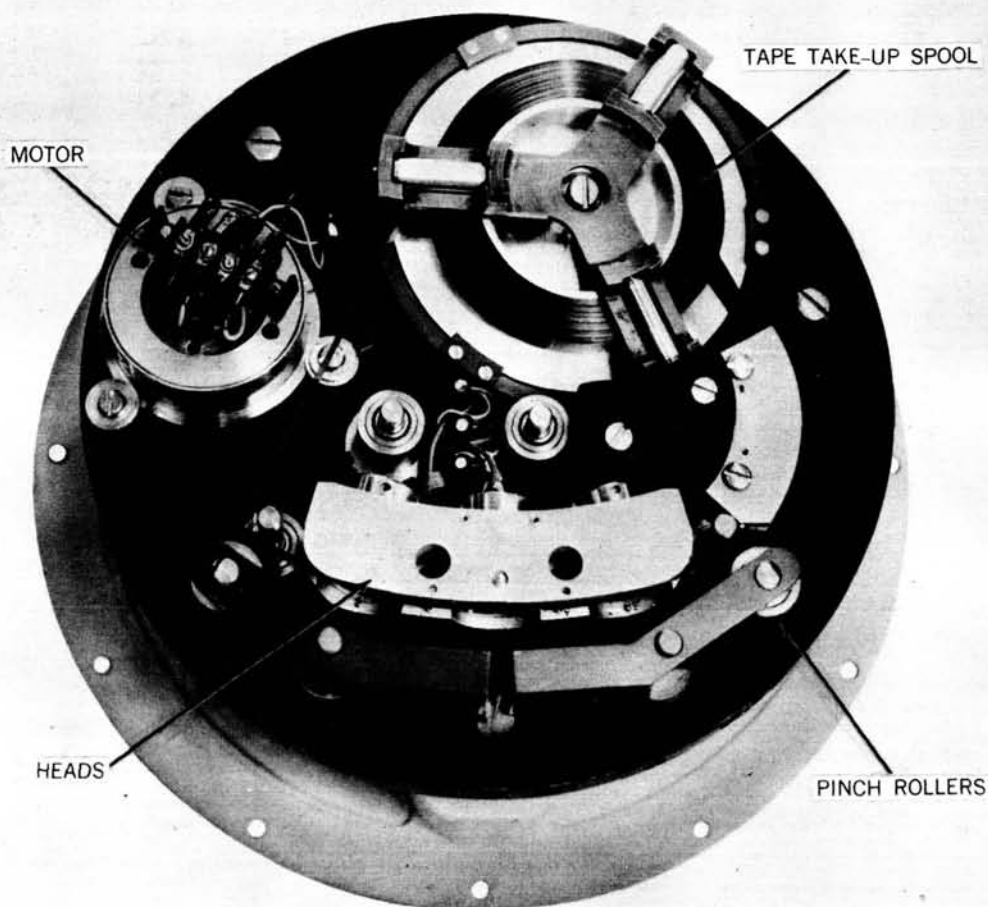


FIGURE 1.—UK-2 tape recorder.

Table 1
Design Vibration Levels for UK-2

Vibration Axis	Frequency Range (cps)*	Level
Thrust Z-Z	5-50	±2.3 g
	50-75	0.072 D. A.
	75-150	±21.0 g
	150-500	±10.0 g
	500-2000	±21.0 g
Lateral X-X and Lateral Y-Y	5-50	±0.9
	50-500	±2.1
	500-2000	±4.2

*The time rate of change of frequency is proportional to the frequency at the rate of 2 octaves per minute.

power spectral density in the frequency band range of 20 to 2000 cps for a duration of 4 minutes. The flight vibration levels are approximately 30 percent less than the design vibration levels.

A further design specification considered was steady state accelerations. The tape recorder was subjected to an acceleration of 27 g's for prototype testing and 18 g's for flight levels. The 18 g acceleration level was the maximum expected in flight and occurred in the last few seconds of the X-248 rocket motor's burning period.

These specifications were used to test particular mounts for the resonance frequency, the amplification decay rate, and the maximum deflection. The flight level test results were used for evaluating the expected flight performance with respect

to the combined effect of steady state accelerations and vibration. Since the isolator system would be exposed to only the individual prototype test levels when tested with the integrated spacecraft, the total maximum deflection for a combined test at these levels would not be significant.

ISOLATOR SELECTION

The primary consideration in selecting an isolator is the frequency or frequency band to be isolated. In the particular design for spacecraft components the frequency of interest is the spacecraft's design parameter of 40 to 50 cps was selected for the isolator resonance. In addition, it was necessary to have a sufficient amount of damping in the mounts to control the amplification at isolator resonance and to insure a rapid decay rate afterward. Furthermore a nonlinear isolator was desirable. This would result in a shift in the apparent resonance frequency of the isolator as the load levels increase. By selecting an isolator system that would have a resonance just about 50 cps for the ± 1.5 g's of the flight vibration levels and a resonance just below 50 cps for inputs of ± 7 g's or more, the maximum acceleration would be somewhat lower than that of a simple resonance (Figure 2).

To accomplish this particular result an isolator with a decreasing spring constant K for an increasing stress level is required. In Figure 3 are shown four typical isolator deflection curves, one with an increasing K , one with a linear K , and two with a decreasing K . An elastic isolator can be designed to have any one of these character-

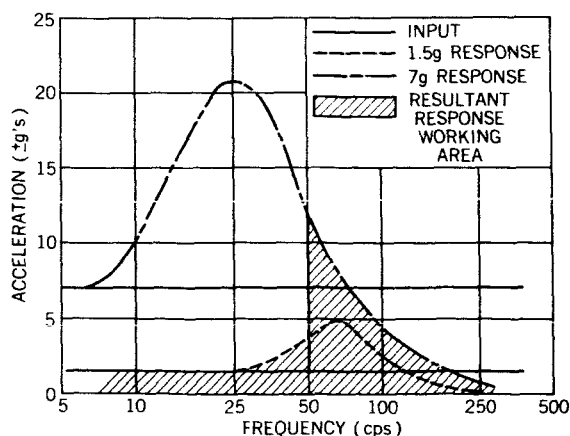


FIGURE 2.—Ideal case, resonance shift at increased loads.

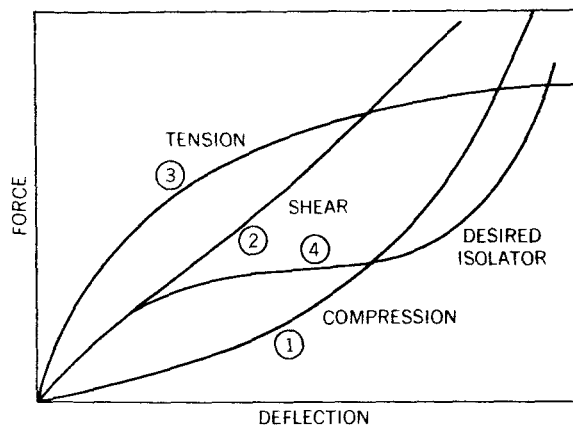


FIGURE 3.—Elastomer deflection curves, $K = \Delta \text{ force} / \Delta \text{ deflection}$.

istics. The first case (curve 1) results from compressive loading of the elastomer, the second (curve 2) from pure shear loading, and the third (curve 3) from pure tension loading. For a mount that has a characteristic of shear for light loads, tension for moderate loads, and compression for severe loads, an isolator with characteristics similar to those for curve 4 should be used. It should be noted, however, that the isolators function best when operating in the shear or linear area of the curve.

Before the final selection of an isolator can be made, a few additional design parameters must be considered. During flight the spacecraft is essentially a free body with a number of external and internal forces, causing side as well as thrust loads to be imparted to the isolator system. Therefore, the amounts must have good isolation properties in the transverse axis as well as in the thrust axis. Furthermore, the physical and chemical properties of the isolator must be unaffected by the temperature requirements and the local environment. Finally, the physical arrangement and size of the isolator must be as simple as possible for the particular system.

TAPE RECORDER ISOLATOR DESIGN

From these design considerations a particular system was proposed for testing. The choice of vibration isolators was limited because of the following requirements:

1. Relatively small size.
2. Good load-carrying properties.

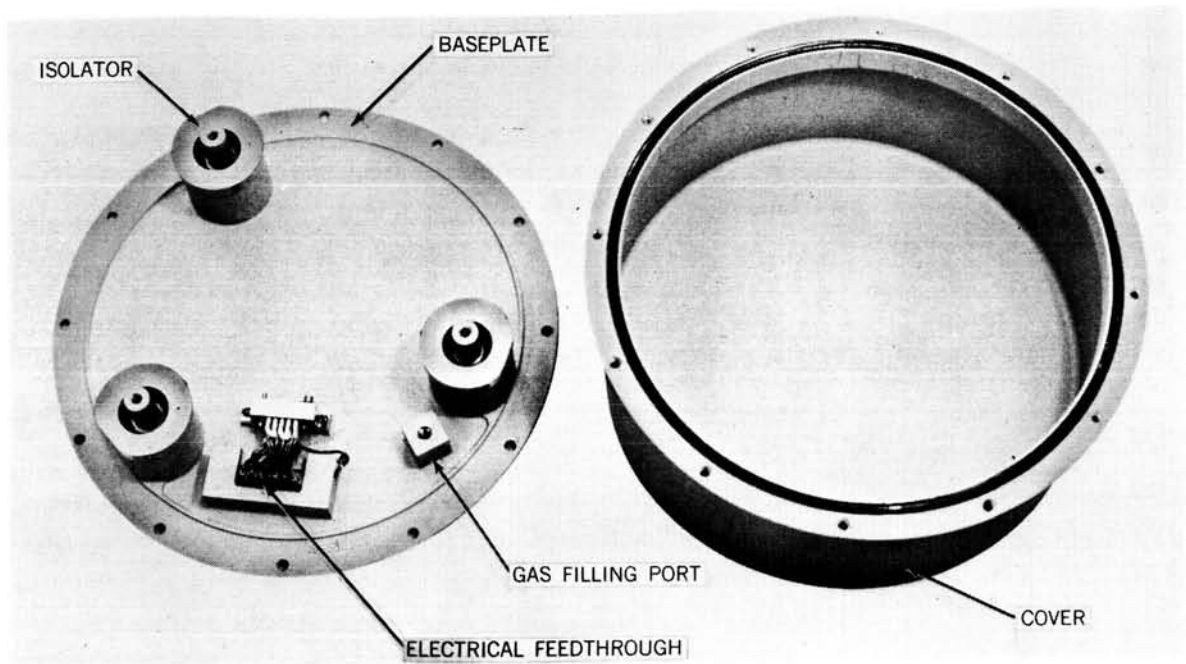


FIGURE 4.—Tape recorder container and isolator system.

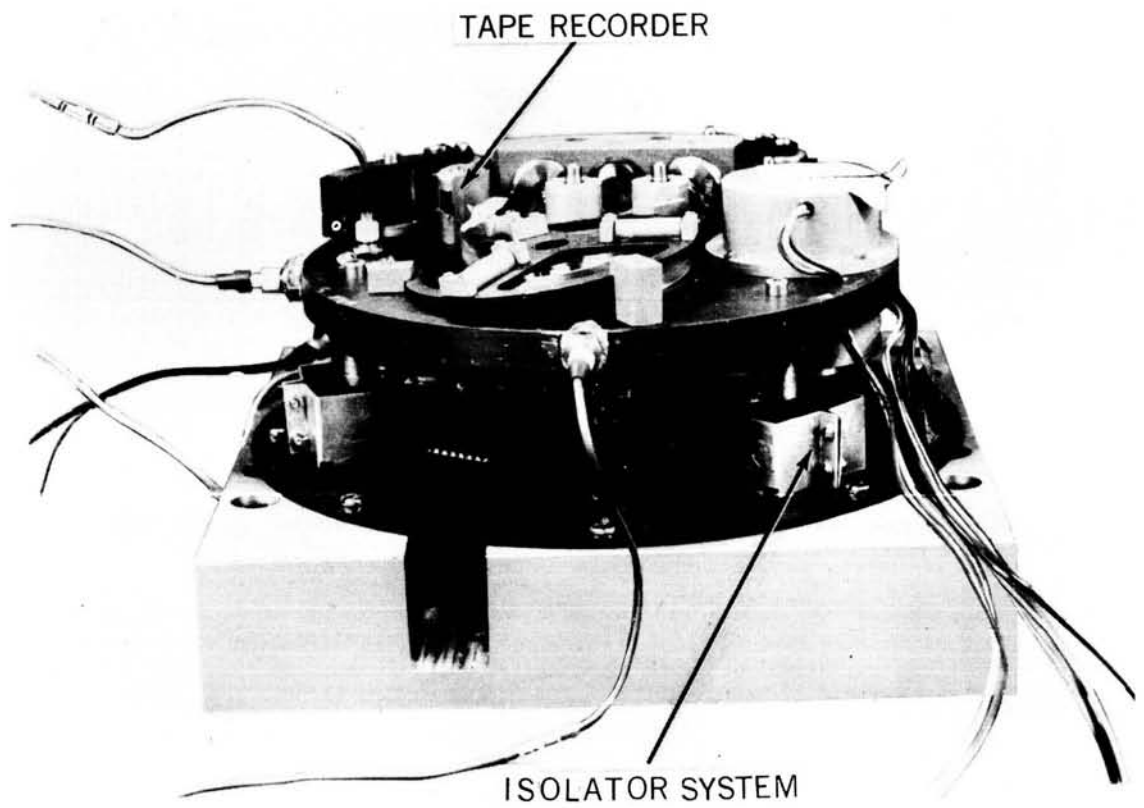


FIGURE 5.—Vibration test setup.

3. Decreasing resonance frequency with increasing stress.

The particular mount chosen was a standard Lord shock and vibration mount of the HT-0 series. The isolator system (Figure 4) consisted of three mounts arranged to be as nearly symmetrical as allowable by the tape recorder transport configuration.

An engineering test system (Figure 5), using the desired arrangement and a number of standard mounts in the 2 and 3 lb load range, was assembled. This unit was used for selecting the proper load range of the mount and determining the deflection characteristics. A pre-prototype model of the container, incorporating the isolator housings into the cover, was then designed and fabricated. A decrease of the inside diameter of the mount housing from that of the standard housing was used to increase the friction damping and lower the amplification at 50 cps.

RESULTS

During the preliminary vibration tests the 2 lb mount was selected as the most acceptable. The acceleration responses are described in Figures 6 and 7. The individual accelerometer outputs

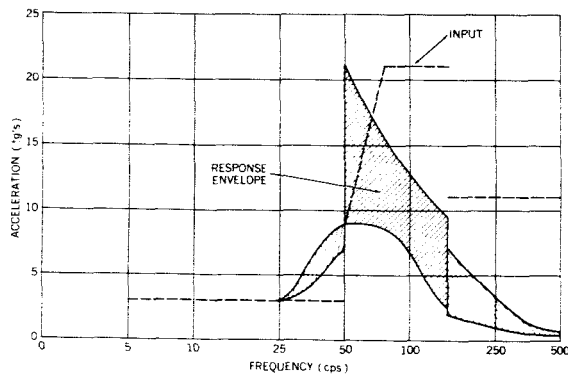


FIGURE 6.—Tape recorder prototype response, thrust axis.

were enveloped to produce an overall picture of the maximum accelerations. Furthermore, the range of accelerations at any particular frequency shows the relative amount of rocking.

In Figure 8 the maximum relative deflection between the tape transport and the baseplate at flight levels is plotted. The minimum clearance envelope for the tape transport was determined

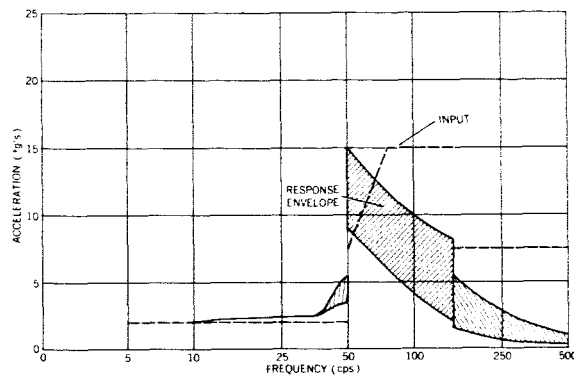


FIGURE 7.—Tape recorder flight response, thrust axis.

from this plot plus the maximum deflection due to steady state acceleration. The final design clearance for all combined loads, with 100 percent safety factors, was only 0.25 inch. The final tests on the preprototype model of the tape recorder and container were performed at the design levels and at 3 times the flight levels. No failures of the tape recorder system resulted.

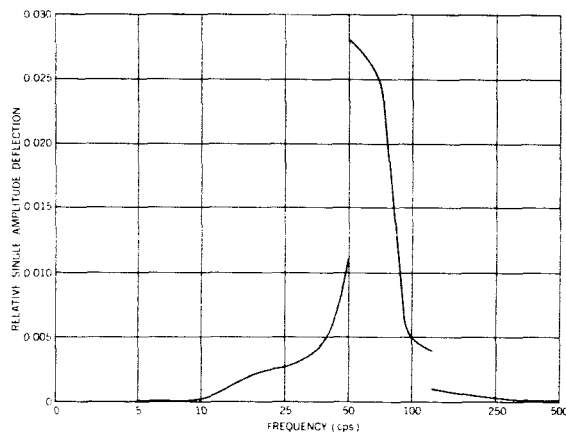


FIGURE 8.—Tape recorder deflection, flight levels.

CONCLUSION

As a result of the tape recorder's performance during all of its tests, future tape recorders will incorporate an isolation system. However, tape recorders are only one area of space satellites in which isolation can be used effectively. Other areas include electronics, electron tubes, mechanical mechanisms, and possibly the complete spacecraft.

ROLLING ELEMENT SLIP RINGS FOR VACUUM APPLICATION*

EDWARD J. DEVINE

Electrical slip rings employing rolling contact elements (balls and raceways) were tested in a hard vacuum environment. Low noise operation was achieved at speeds up to 5000 rpm. Slip rings lubricated with MoS₂ operated with lower noise and had a longer life than slip rings employing solid metal film lubrication. Noise generated by rolling slip rings is more severe in air than in a hard vacuum. A technique was developed which operated satisfactorily in air for a time sufficient for preflight testing. A method of replenishing MoS₂ lubrication on the slip rings was devised. The final configuration operated in a vacuum of 2×10^{-9} torr at 2000 rpm for over 100 million revolutions at a noise level of 0.002 ohm rms.

INTRODUCTION

This study investigated the feasibility of rolling contact electrical slip rings in a hard vacuum environment. Conventional sliding slip rings have been reasonably satisfactory in air, but are subject to problems in a hard vacuum environment. Conventional graphite brushes are unsatisfactory and metal wipers are suspect because of increased friction, greater wear, and the possibility of cold welding. Considerable research is underway on lubricants and materials to alleviate friction and wear. Although progress has and will continue to be made, it is evident that sliding contact elements will have definite limitations on speed and operating life.

It was surmised that friction and wear might be drastically reduced by replacing the conventional sliding slip rings with rolling contact elements. Accordingly, a limited study was conducted to investigate this approach.

APPROACH

Rolling elements have been used in the past for the transfer of electrical energy. Early dc motors employed rolling metal brushes. On occasion, the ball bearings of machines have been used to provide electrical contact between the rotor and stator. These methods are seldom used in a normal atmosphere because there is no wiping action and as a result oxide films and contaminants form which deteriorate the contact. This

usually leads to arcing which quickly pits and destroys the rolling elements and races.

In a hard vacuum, two factors result in favorable environment for rolling contact:

1. Surfaces are cleaned by outgassing and film formation is reduced.
2. Below a critical pressure, arcing is extinguished and damage from this source is eliminated.

Rolling contact slip rings can be used in many configurations. For this investigation, a thrust type of ball bearing configuration was selected (Figure 1).

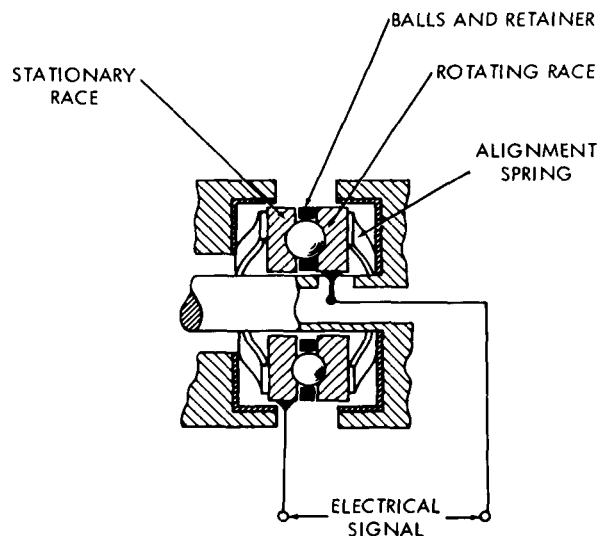


FIGURE 1.—Basic rolling element slip ring.

*Published as NASA Technical Note D-2261, April 1964.

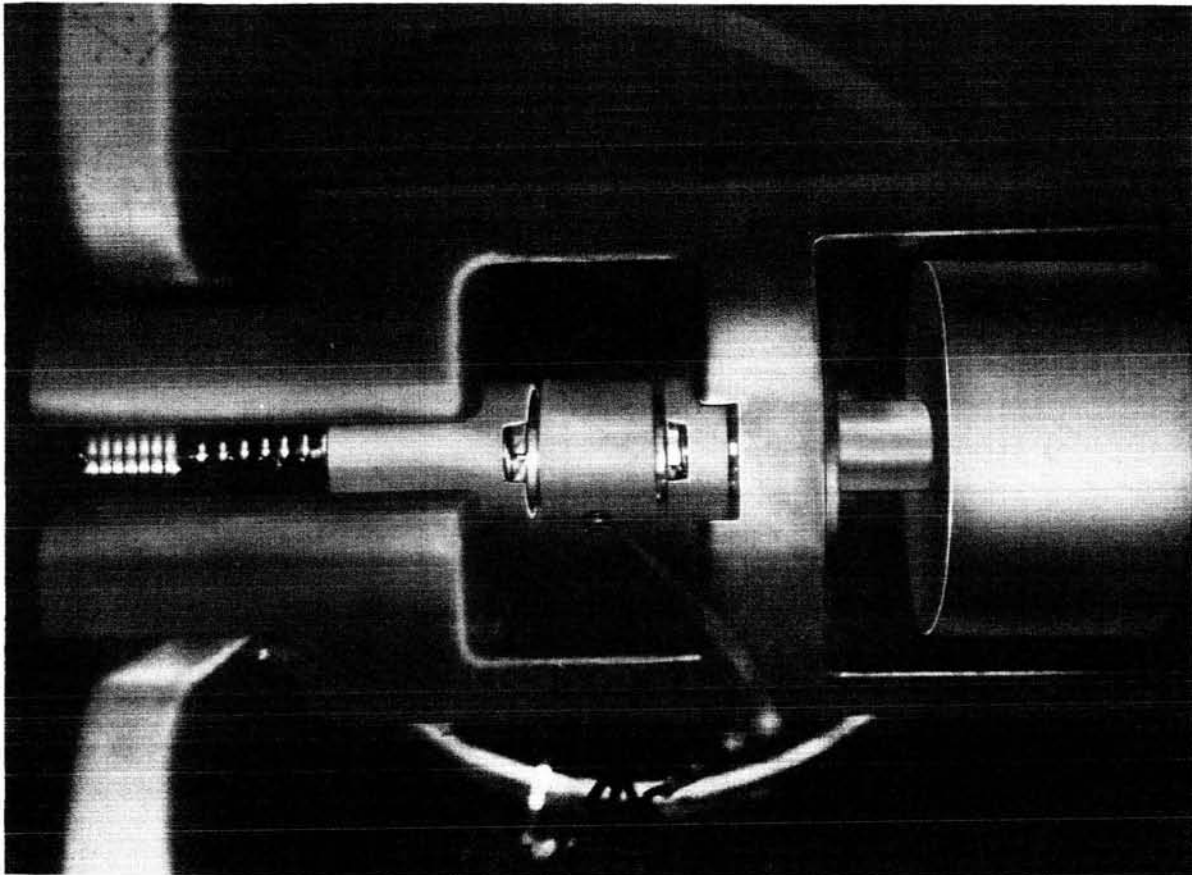


FIGURE 2.—Slip ring test fixture.

The following features make this arrangement suitable for a basic study of rolling contact:

1. The contact load is uniform on each ball.
2. The contact load may be simply and accurately controlled.
3. Since the balls contact each race at the same diameter, nearly pure rolling action is achieved.
4. High quality commercial balls and races can be readily obtained.

TEST FIXTURE

The test fixture is shown in Figures 2 and 3. Two pairs of slip rings are employed in series to permit transfer of a signal to and from the rotating frame. Load is applied to the slip rings by means of a calibrated spring which forces on the rear bearing support. The rotating races are driven by a rotor which is keyed to the shaft. The rear bearing support and rotor are lubricated

with MoS_2 to assure free motion and uniform loading on the front and rear slip rings. The

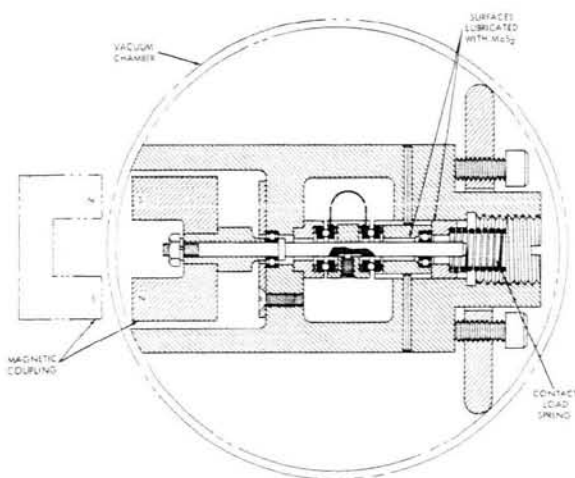


FIGURE 3.—Slip ring test fixture assembly.

radial support bearings are gold plated, and they are lubricated with MoS_2 applied in an alcohol slurry. A magnetic coupling capable of 0.87 in.-oz. slip torque drives the assembly through the wall of the vacuum chamber.

Extremely low outgassing of the entire fixture was achieved by complete elimination of organic materials. The slip rings are electrically insulated by a spray coating of aluminum oxide ceramic on the bearing seats. This material is an excellent insulator and can be honed to precision dimensions. Experience showed that a minimum spray coating of 0.010" was required for proper fabrication.

One requirement for low contact noise is uniform loading on the slip rings as they are rotated. The initial design attempted to achieve this by accurate (parallel) machining of the seating surfaces. However, tests established that a self-aligning feature was required to assure uniform contact pressure. A simple and effective method

was provided by annular wavy spring washers behind one or both races of the slip ring.

More uniform contact was achieved with a loose radial fit (0.004") on the slip ring races rather than a close fit with extremely close tolerances on concentricity.

To facilitate assembly of the test slip rings, electrical connection was originally made by pressure contact with a gold-plated washer under the alignment spring washer. These additional contacts proved to be a source of noise. The problem was eliminated by resistance welding of the electrical connections directly to the edges of the slip ring races.

It was found that a contact load of 1 lb. was sufficient to maintain continuous contact in the final test fixture. This load was maintained constant throughout all tests. This load probably could be reduced without a sacrifice in performance, but the lower limit was not experimentally determined.

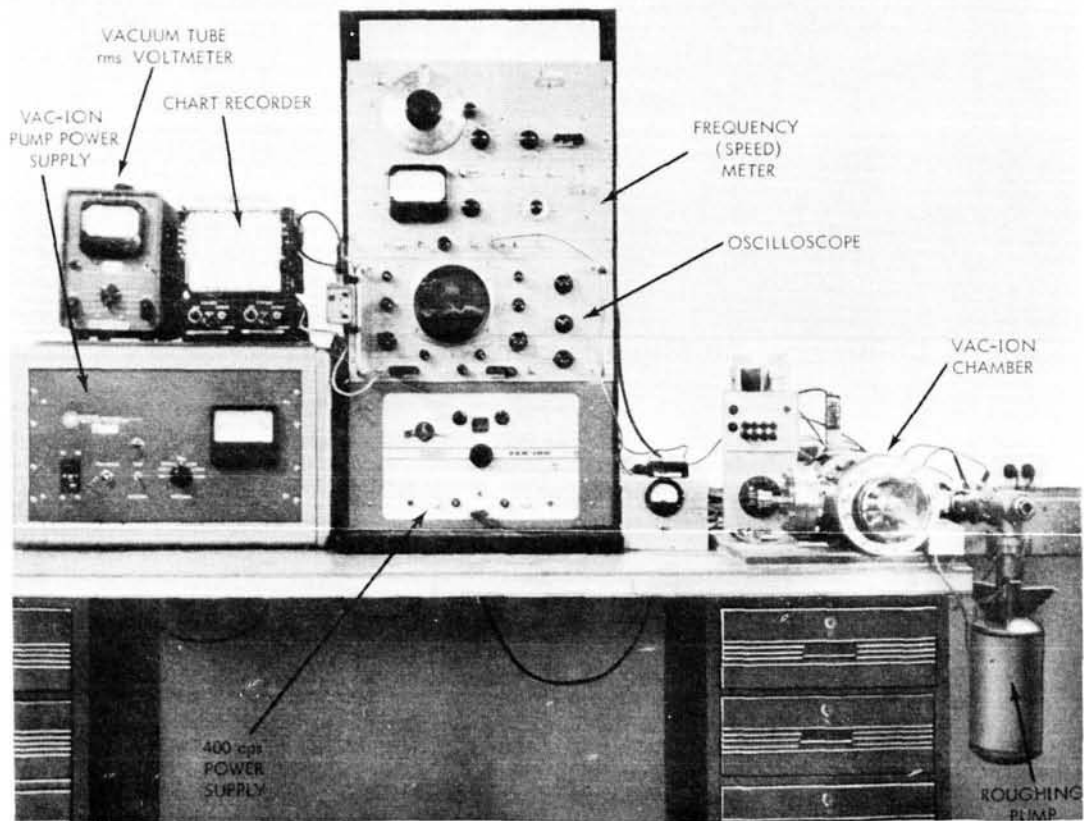


FIGURE 4.—Test equipment.

EQUIPMENT

The vacuum facility employed was a stainless steel chamber with a 75 liter per second Vac-Ion pump (Figure 4). No mechanical forepumping or other source of oil contamination was present.

Rotary drive for the experiment was provided by a 400 cps induction motor which was capable of 600 to 5000 rpm directly coupled (60–500 rpm with a 10 to 1 gear reduction). The motor was located outside the vacuum chamber and connected to the experiment by means of a magnetic coupling.

INSTRUMENTATION

The principal measurement made was contact noise. The circuit employed is shown in Figure 5.

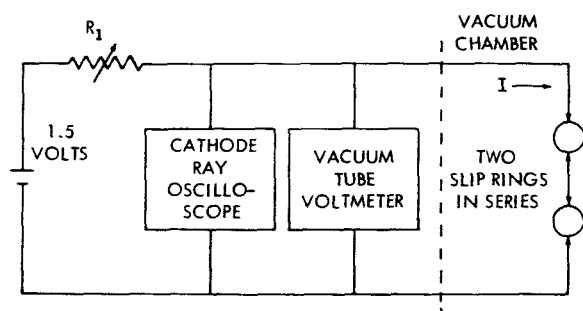


FIGURE 5.—Circuit diagram.

In all cases where noise is of an acceptable level, the contact resistance R_c is small compared with R_1 , resulting in essentially constant current conditions. Contact noise is found by measuring the variations in voltage drop across R_c . This will be reported as peak-to-peak and rms voltage (E_n) at a given current and also as contact resistance variation calculated from:

$$R_n = \frac{E_n}{I}$$

All noise readings will be reported as measured and apply to 2 slip rings in series. Peak-to-peak noise is thought to be independent of the number of contacts, whereas the rms values should be reduced by approximately a factor of 2 for a single slip ring.

The instruments employed were:

For Peak-to-Peak Noise

Cathode Ray Oscilloscope
Hewlett Packard Model—130 BR
Sensitivity—1 mv/cm
Bandwidth—dc to 300 kc

For rms Noise

ac Vacuum Tube Voltmeter
Hewlett Packard Model—400 D
Sensitivity—1 mv rms full scale
Bandwidth—10 cps to 4 Mc

For Contact Resistance

Recording Voltmeter
Varian Model—B22A
Full span balancing time—1 sec.

Background noise from the line voltage and nearby laboratory equipment was reduced to less than 0.0001 volts peak-to-peak by shielding.

TEST RESULTS

Four test runs were required before the final configuration of the test fixture was decided upon. Because of the changes in the mechanical configuration resulting from these tests, the data obtained are not of interest. The test results of interest are summarized in Table 1 and discussed below.

Test 1

This test employed only gold-plated 440C stainless steel slip ring elements without lubrication. Initial operation in air gave low noise. The noise level in vacuum was moderate, but became severe after a short time (2.4×10^6 revolutions). Upon exposure to air the contact resistance became very erratic, averaging about 0.6 ohms. It rose to 20 ohms after a day in the atmosphere. Examination revealed that the gold plating was badly worn. It is surmised that the stainless steel base metal was exposed and that a nonconducting oxide film formed.

Test 2

To explore the nature of the nonconducting film, the same balls and races were again subjected to a vacuum environment. With the slip rings rotating at 200 rpm, contact resistance and noise dropped dramatically with decreasing pres-

Table 1
Slip Ring Test Results

Test	Configuration			Air Operation		Vacuum Operation						Remarks		
	Balls	Races	Retainer	Lubricant	R _c (ohms)	p-p Noise (ohms)	R _c (ohms)	p-p Noise (ohms)	rms Noise (ohms)	p-p Noise (mv at 10 ma)	Vacuum		Speed (rpm)	Life (rev × 10 ⁶)
1	Gold-plated 440C	Gold-plated 440C	Gold-plated crown	None	0.095	0.010 (0 rpm)	0.56	0.6	0.02	6	1 × 10 ⁻⁷	200	2.4	Noise increased to 140 mv p-p
2	See remarks				20.0	0.360 (0 rpm)	0.23	0.4	0.08	4	2 × 10 ⁻⁷	5000	44.2	Noise increased to 1.5 volts p-p; configuration consisted of worn parts from test 1
3	Gold-plated 440C	Gold-plated 440C	Gold-plated crown	2 dips MoS ₂	1.0	Very high	0.5	0.2	0.004	2	1 × 10 ⁻⁷	5000	50.0	Failed mechanically
4	Gold-plated 440C	Gold-plated 440C	Gold-plated crown	2 dips MoS ₂	2.0	2.0	0.3	0.03	0.002	0.3	5 × 10 ⁻⁸	2000	16.0	Failed mechanically
5	Gold-plated 440C	Gold-plated 440C	Gold-plated crown	4 dips MoS ₂	150.0	20.0	0.4	0.05	0.004	0.5	1 × 10 ⁻⁷	2000	34.5	Failed mechanically; noise at 300 rpm, 20 mv p-p; relubricated assembly from test 4 was used
6	Tungsten- carbide	Rhodium- plated 440C	Gold-plated crown	None	—	30.0	—	60.0	—	600.0	2 × 10 ⁻⁷	600	0.003	Failed mechanically
7	Tungsten- carbide	Gold-plated 440C	Gold-plated crown	None	100.0	40.0	0.3	20.0	—	200.0	4 × 10 ⁻⁶	2000	0.12	Failed mechanically
8	Tungsten- carbide	Rhodium- plated 440C	Gold-plated crown	None	4.5	0.6	—	0.04	0.002	0.4	3 × 10 ⁻⁷	2000	0.24	Failed mechanically
9	Tungsten- carbide	Rhodium- plated 440C	Machined gold	None	—	—	—	0.03	—	0.3	1 × 10 ⁻⁸	2000	31.6	Failed mechanically
10	Gold-plated 440C	Gold-plated 440C	Machined duroid	None	8	60.0	25	15.0	2.5	150.0	4 × 10 ⁻⁸	2000	13.4	Still running; noise very high
11	440C	<i>in situ</i> MoS ₂	Crown <i>in situ</i> MoS ₂	<i>in situ</i> MoS ₂	600	High	0.3	0.04	<0.002	0.4	1 × 10 ⁻⁷	2000	57.5 ^a	Failed mechanically; may have been due to support bearings
12	Gold-plated 440C	Gold-plated 440C	Machined, 85% gold 15% MoS	Very light MoS ₂	0.25	0.1 (2000 rpm)	0.3	0.04	<0.002	0.4	5 × 10 ⁻⁹	2000	115	Failed mechanically

sure. At a pressure of 2×10^{-7} torr, contact resistance had dropped to 0.5 ohm and noise was less than 0.05 ohm peak-to-peak. The assembly was then run continuously at 5000 rpm with the results shown in Table 1. The noise level was fairly low until a sudden increase after 44×10^6 revolutions (Figure 6).

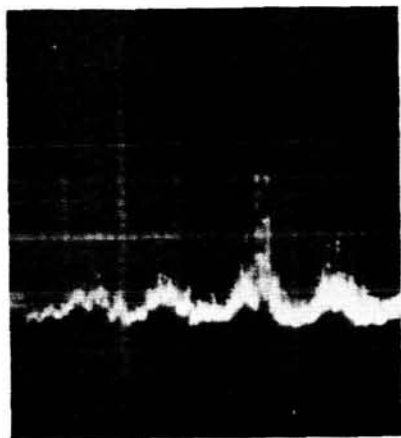


FIGURE 6.—Noise trace for test 2, unlubricated gold-plated elements, 50 ma current, 5000 rpm; vertical scale—20 mv per grid division, horizontal scale—5 msec per grid division.

Test 3

It was clear that some additional lubrication was required both to reduce noise generated by “sticky” operation and to increase operating life. Ideally, the lubricant should be conductive and have extremely low vapor pressure. Liquid metals were considered, but they were not tested because of problems of corrosion and containment. Molybdenum disulphide was selected for testing despite the fact that in bulk it is an insulator. The MoS_2 was applied to gold-plated bearing elements by dipping them in an alcohol suspension. The initial resistance was 8000 ohms, but after a brief period of operation at 1000 rpm it dropped to 1 ohm. Electrical noise remained high when the fixture operated in air.

This assembly was tested in vacuum and operated with very low noise at 5000 rpm until mechanical binding occurred at 50×10^6 revolutions. Post-test examination revealed pitting of one set of balls (Figure 7), indicating a failure of the lubricating film and the resulting adhesive wear.

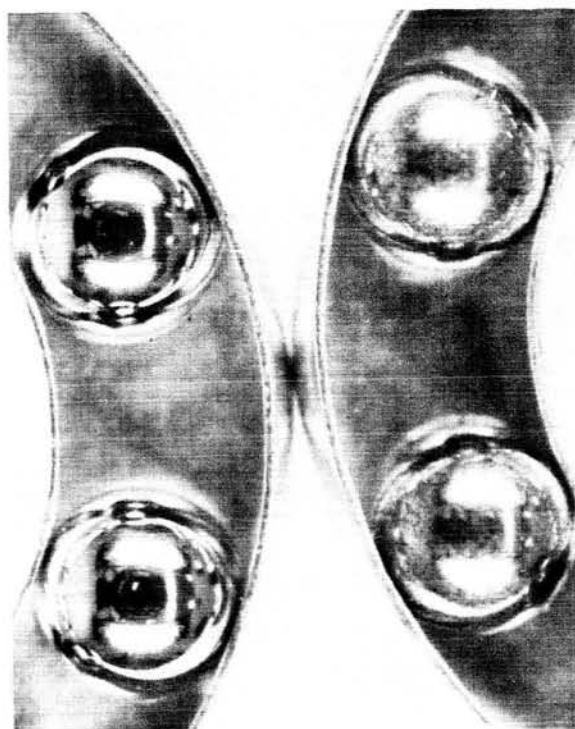


FIGURE 7.—Balls and retainers from test 3. Note pitting on one set of balls.

Test 4

This test was a repeat of Test 3. Very low noise operation was noted before 16×10^6 revolutions when a sudden increase occurred.

Test 5

To determine if the failure in test 4 was due solely to a loss of lubricant, the assembly for test 4 was relubricated with MoS_2 applied in an alcohol suspension to the slip rings without disassembly or disturbance of the fixture. After relubrication, low noise operation resumed for an additional 34×10^6 revolutions.

Test 6

Other materials were investigated in an effort to find a combination which would have the desired characteristics of (1) low contact resistance; (2) low friction in vacuum; (3) good wear properties. Test 6 employed tungsten-carbide balls and rhodium-plated races. Contact resistance and noise in air were disappointingly high. Attempts to operate the combination in vacuum were not successful.

Test 7

This was a test of tungsten-carbide balls and gold-plated races and was also unsatisfactory.

Test 8

This test was conducted to determine whether the crown retainer (which is difficult to plate satisfactorily) was responsible for the failures in tests 6 and 7. The test was conducted with a machined gold retainer and results were not significantly better.

Test 9

This test employed tungsten-carbide balls, rhodium-plated races, and a machined retainer with MoS_2 lubrication. Very good vacuum results were noted until lubrication failure at 31×10^6 revolutions.

Test 10

This test employed gold-plated stainless steel balls and races and a machined duroid retainer. This combination operated very well mechanically, but with very high electrical noise.

The observations of contact resistance and noise reveal much about the nature of the lubricating film formed with the duroid material. An initial contact resistance of 1 ohm increased to practically open circuit after 1 hour of operation in air. Under vacuum conditions with 10 ma current, contact resistance dropped to 15–25 ohms. The much higher resistance of the duroid film in both air and vacuum when compared with MoS_2 films, clearly shows the presence of teflon in the lubricating film. This test was stopped after 31×10^6 revolutions because of continued high noise.

Test 11

The lubricant employed in this test is a coating of MoS_2 applied *in situ*. Stainless steel races and crown retainers were first electroplated with molybdenum which was then converted to MoS_2 by subjecting the specimen to H_2S gas at an elevated temperature and pressure. This process is described in detail in Reference 1. Balls were plain 440C stainless steel. This combination was noisy in air, but gave very good results in vacuum. The test was terminated after 54×10^6 revolutions because of mechanical difficulties. The slip ring components appeared to be in good condition

(Figure 8) and it is suspected that failure was due to the support bearings.

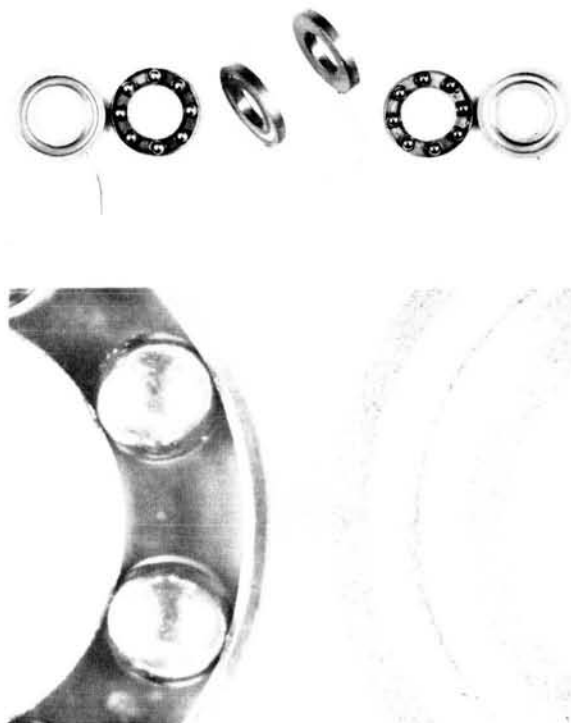


FIGURE 8.—Balls, retainer, and race from test 11.

Test 12

Previous tests showed that MoS_2 films and coatings are effective lubricants, but have limited life. In addition, MoS_2 films are electrically noisy in air, an undesirable limitation. It is believed that the best way to employ MoS_2 is in a continuous manner.

For this purpose a ball retainer was machined from a sintered compact 85 percent by weight gold and 15 percent MoS_2 . Since the retainer was in continuous light sliding contact with the balls, a continuous transfer of MoS_2 was expected.

Assembly 12 employed gold-plated balls and races and the machined gold- MoS_2 retainer. When operation was attempted in vacuum, mechanical difficulties and high noise were experienced. Contrary to previous experience, it was evident that operation of unlubricated gold-plated elements is not satisfactory under a 1 lb. contact load.

The balls and races were lubricated with a very light suspension of MoS_2 which was to serve until a film could be transferred from the retainer. It was found that this light film resulted in low noise operation in air. The test was then resumed in a vacuum of 5×10^{-9} torr. Contact resistance was 0.28 ohm and noise was less than 0.04 ohm peak-to-peak (Figure 9). A life of 115×10^6 revolutions was achieved before mechanical failure.

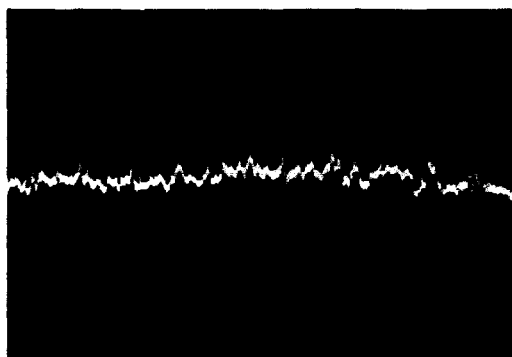


FIGURE 9.—Noise trace for test 12, gold- MoS_2 retainers, 50 ma current, 2000 rpm; vertical scale—1 mv per grid division, horizontal scale—2 msec per grid division.

Miscellaneous Tests

Most of the noise measurements were made by observing the variation of low level dc signals. Long term tests were made with a constant 10 ma dc current through the slip rings. Spot checks were made at currents up to 100 ma and in all cases the voltage noise increased in an approximately linear fashion, indicating that R_n is constant over this range of current.

A brief test was made (test 9) at higher dc current (up to 1 ampere) in a vacuum of 2×10^{-7} torr. No visible arcing or damage to the slip rings resulted.

During test 3 a thermocouple output was connected through the slip rings which were run at 2500 and 5000 rpm in a vacuum of 3×10^{-8} torr. The accuracy and sensitivity of the potentiometer readings were not affected.

During test 12 an ac signal was transferred through the slip rings operating at 2000 rpm and 5×10^{-9} torr. The signal was varied from 1 to 20 kc at an amplitude of 100 mv peak-to-peak. No visible noise or distortion was observed.

CONCLUSIONS

It has been demonstrated that rolling contact can be employed for electrical signal transfer at low noise in an ultra-high vacuum environment. Contact noise can be less than 2 milliohm rms even at speeds up to 5000 rpm. There are indications that the higher the vacuum, the less will be the contact noise.

Lubrication is required for maximum performance from rolling contact slip rings. Precious metal platings without additional lubrication appear unsatisfactory because of high torque and "sticky" motion. The best results obtained to date have been with stainless steel balls and races, gold-plated and with a thin initial film of MoS_2 . Continuous lubrication is provided by ball retainers (separators) machined from compacts with a high MoS_2 content.

A direct comparison of the performance of rolling and sliding contact is difficult because of the paucity of reported data and wide variation in test conditions and instrumentation methods.

Assuming a reasonable specification for high quality miniature sliding slip ring operating in air, we could expect:

- Noise—50 milliohms peak-to-peak,
5 milliohms rms,
- Maximum speed—200 rpm,
- Maximum life— 20×10^6 revolutions.

Expected performance in vacuum is more difficult to predict. Reference 2 reports that satisfactory operation is not possible and recommends oil vapor lubrication. Reference 3 reports favorable results for brushes impregnated with MoS_2 at 17.8 rpm and a vacuum of 2×10^{-6} torr. Noise figures from Reference 3 are not directly comparable with the present data because of the low band-width (8 cps) reported.

No data are available on the performance of sliding contacts at the speeds and vacuum levels explored in this study. Based upon the data obtained in this study, it appears that rolling contact elements are capable of the following performance:

- Noise—50 milliohms peak-to-peak,
5 milliohm rms,
- Maximum speed—greater than 5000 rpm,

Maximum life—greater than 100×10^6 revolutions,
Vacuum environment— 10^{-6} torr and lower.

REFERENCES

1. VEST, C. E., "Adaptation of a MoS_2 *in situ* Process for Lubricating Spacecraft Mechanical Components," NASA Technical Note, to be published.
2. SALMON, W. A., and SCAMMELL, H. (ed.) "Proceedings of a Symposium on Lubrication in Space," published by Baird-Atomic, Inc. and Arthur D. Little, Inc., 1963.
3. GOETZEL, C. G., and SINGLETARY, J. B., (ed.), "Space Materials Handbook," published by Lockheed Missiles and Space Co., Sunnyvale, Calif., under Air Force Contract AF 04(647)-673, 1962.

ENERGY ABSORBER FOR THE ARIEL I INSTRUMENT BOOMS*

T. L. ENG

The instrument boom energy absorber aboard the Ariel I was designed on the principle of runaway escapement. Its function was to reduce the erection rate of the instrument booms, thereby minimizing the kinetic energy transferred to these booms during their deployment and thus further safeguarding the experiments carried by the booms from shock damages. This device consisted of a pulley coupled to a gear train, which in turn was regulated by a pallet on a scape wheel. Since the frequency of the pallet is a function of the actuating torque, its amplitude, and moment of inertia, the rate of movement of the escapement mechanism can be adjusted by making the pallet moment of inertia variable (the others constant).

Ariel I was successfully launched on April 26, 1962. Telemetry data from the satellite indicated that all experiments carried by the booms had functioned normally.

INTRODUCTION

The first successfully launched satellite of the International Space Exploration program, Ariel I (1962) had two booms—each carrying a scientific experiment (the electron density and the electron temperature sensors) on its outer end. These booms were placed along the side of the last stage during flight and were released at a predetermined time and allowed to deploy by centrifugal force. Calculations indicated that, for the spacecraft to achieve the desired final spin rate, these booms must be released at the time when the satellite is spinning at 74 rpm. Further calculations showed that approximately 0.3 second would be required for the unrestrained booms to be fully deployed. At the instant when these booms cease traveling and are locked in their final position, they would be required to absorb 33 ft-lb of energy. Therefore to safeguard the above-mentioned experiments and ultimately protect the useful life of the satellite, the energy transferred to the booms during their deployment should be reduced.

DESIGN CONSIDERATIONS AND ANALYSIS

Requirements

The selected design, which would reduce the energy transferred to the booms, must also be:

1. A self-actuating device containing no out-gassing materials which might contaminate the experiments;

2. Capable of controlling the booms simultaneously so that they may erect together or within a fraction of a second of each other;

3. Capable of varying the erection time to not less than 1 second at the anticipated boom release spin rate and capable of withstanding a load equivalent to a 20 percent overspin or 90 rpm;

4. Light and capable of being housed in existing available space; and

5. Able to meet the environmental test specifications at the prototype level.

The second condition indicated that the mechanism should be centrally located. The only available room in the central area was the space (7 in. diam. \times 2 in.) below the tape recorder, where the middle portion was occupied by an electron temperature gauge (Figure 1).

To satisfy all conditions, a constant velocity device—an escapement mechanism (similar to the clock mechanism)—was selected (Figure 2). Since the kinetic energy of a moving mass is proportional to its velocity square, $KE = \frac{1}{2}MV^2$, theoretically, by controlling the erection rate or V of the booms, the energy transferred to them is also controlled.

Force Analysis

To simplify the force analysis on the booms, frictional forces at the hinge were neglected. It was also assumed that at the time of boom deployment the satellite was in an environment of

*Published as NASA Technical Memorandum X-936, January 1964.

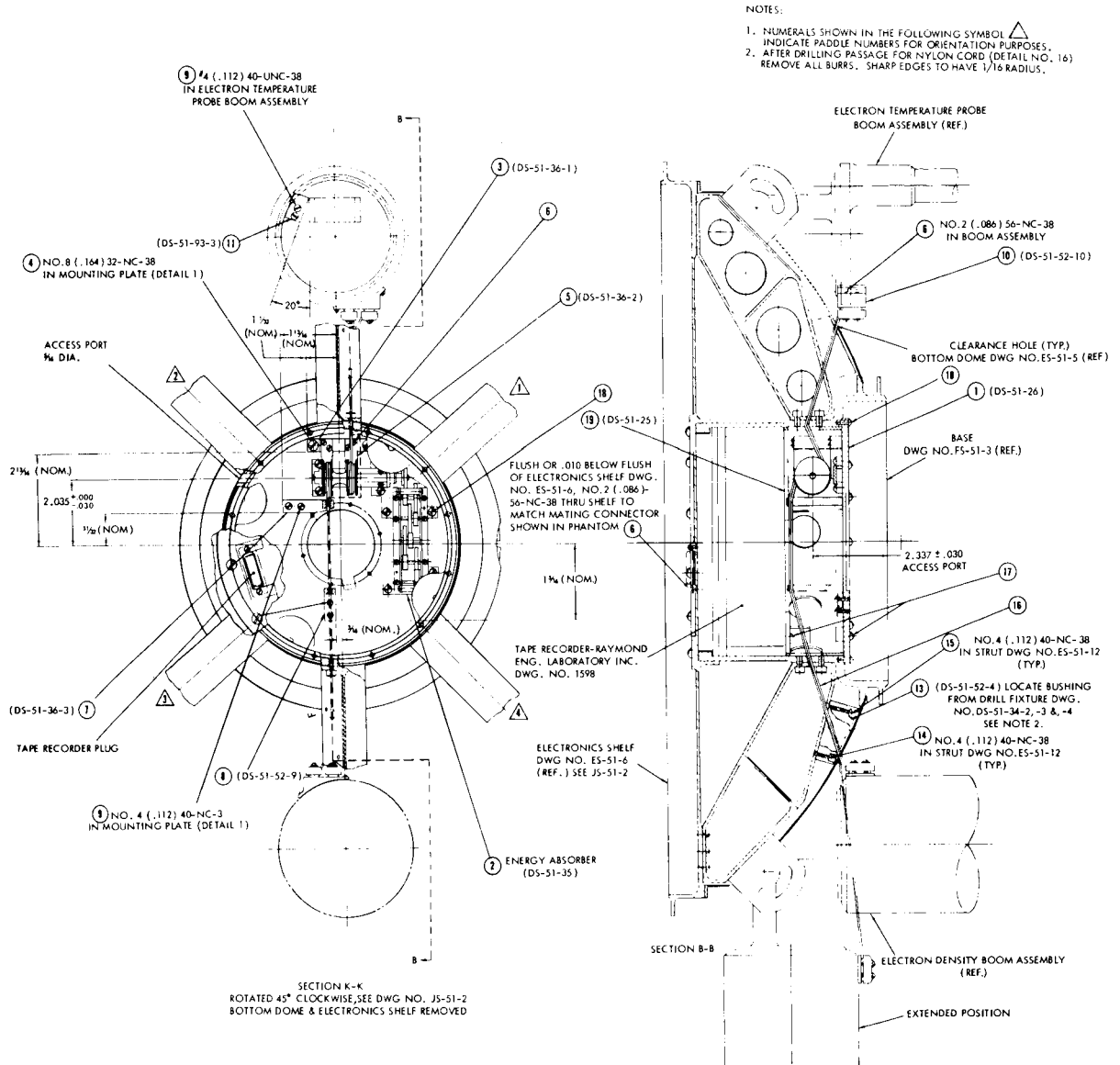


FIGURE 1.—Location of energy absorber in relation to the instrument boom.

weightlessness; consequently, the weight of the boom assembly was not considered. The force with which the energy absorber is directly concerned is the centrifugal force acting on the booms, and it was further assumed that this force acted through the center of mass of the boom assembly (see Reference 1 and Figure 3):

$$F_{m+M} = M' \bar{r} \phi^2, \quad (1)$$

where $M' = m + M$, m is the mass of concentrated

weight, M the mass of the boom, \bar{r} the distance to the boom's cg from the spin axis, F_{m+M} the centrifugal force acting on the center of mass, and ϕ the spin rate of the satellite at any given θ , the boom position at any given time. If

$$\bar{r} = a + \bar{L} \sin \theta, \quad (2)$$

we have

$$F_{m+M} = M' \phi^2 (a + \bar{L} \sin \theta), \quad (3)$$

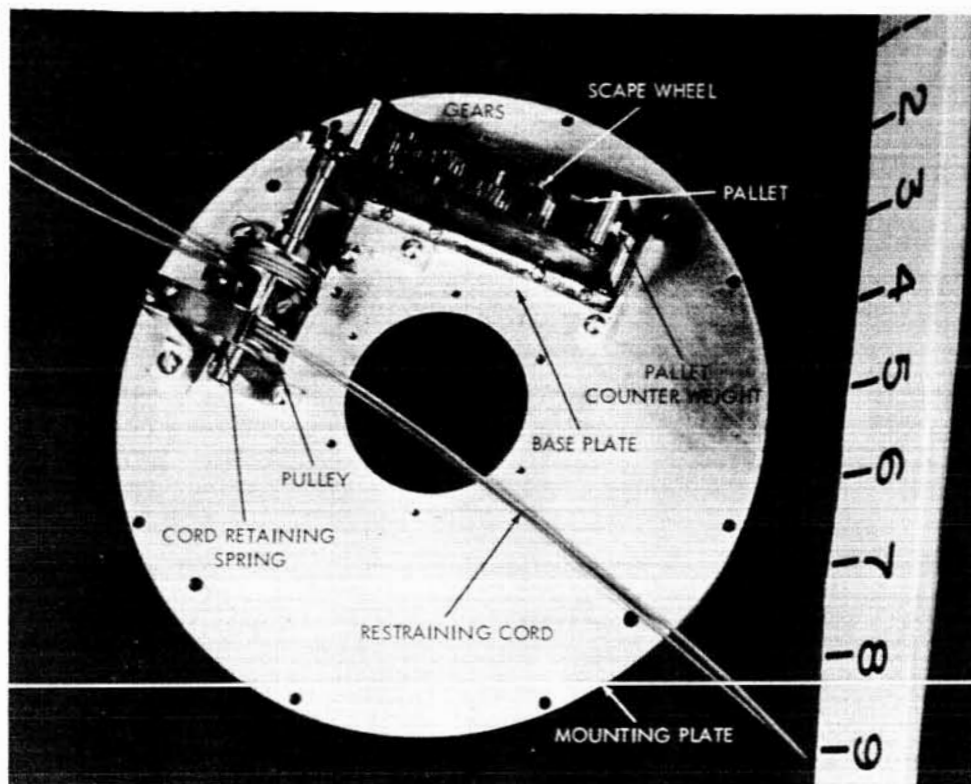


FIGURE 2.—The escapement mechanism.

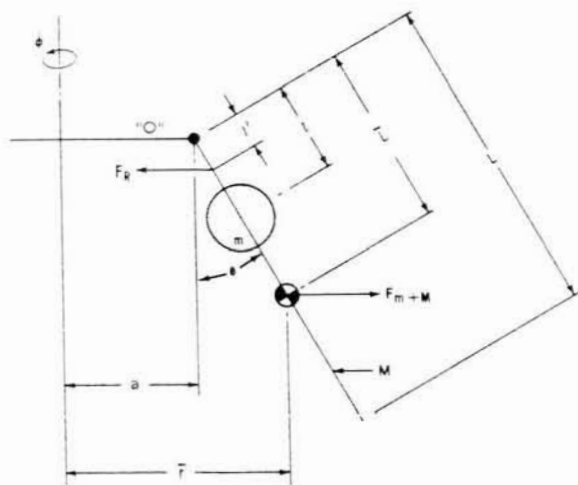


FIGURE 3.—“Free-body” schematic of satellite.

where \bar{L} is the distance from the pivot end of the boom to the center of mass M' , and a the distance from the spin axis to the boom hinge. The sum of the moments about the pivot point is zero:

$$\Sigma M_0 = 0;$$

therefore the reaction force

$$F_R = \frac{F_{m+M} \bar{L} \cos \theta}{l' \cos \theta} \quad (4)$$

where l' is the location of the restraining point of the energy absorber. Thus, for $\theta < 90^\circ$, we have

$$F_R = \frac{F_{m+M} \bar{L}}{l'} \quad (5)$$

and by substituting Equation 3 into 5, we obtain

$$F_R = \frac{M' \dot{\phi}^2 (a + \bar{L} \sin \theta) \bar{L}}{l'} \quad (6)$$

where F_R is the restraining force from the escapement mechanism required to set one boom at equilibrium at any given position, θ . Since there were two booms, the total force acting on the energy absorber would be $2 F_R$. To evaluate ϕ in terms of θ at any given $\dot{\phi}_0$, the satellite's spin rate prior to deployment, the following

equation was used:*

$$\frac{\dot{\phi}}{\phi_0} = \frac{(m+M)a^2 + I}{m(l \sin \theta + a)^2 + M\left(\frac{L}{3} \sin^2 \theta + aL \sin \theta + a^2\right) + I}, \quad (7)$$

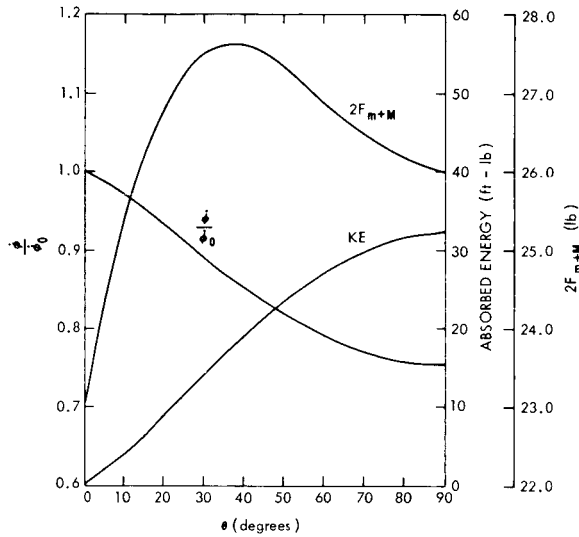


FIGURE 4.—Calculated results based on $\phi_0 = \text{rpm}$

where Figure 4 shows that F_{m+M} maximized at approximately 35° ; therefore the design load for the energy absorber would be $F_{m+M} \bar{L}/l'$ at $\theta = 35^\circ$. I is the moment of inertia of the satellite prior to appendage deployment.

The actuating force, F_A , on the escapement mechanism (per boom) is as shown in Figure 5.

$$F_A = \frac{l'}{\bar{L}} F_{m+M} \quad (8)$$

To insure the actuation of the energy absorber under most conditions, a minimum F_A was chosen. In this case, where F_A does practically no work when $\theta \geq 60^\circ$, it was assumed that F_A is minimum when $\theta = 60^\circ$.

Escapement Design

The expression for a pallet half-cycle period for a runaway escapement (References 2 and 3) is

$$t_1 = \sqrt{\frac{2I_{p_1} \theta_1}{T_{p_1}}}, \quad (9)$$

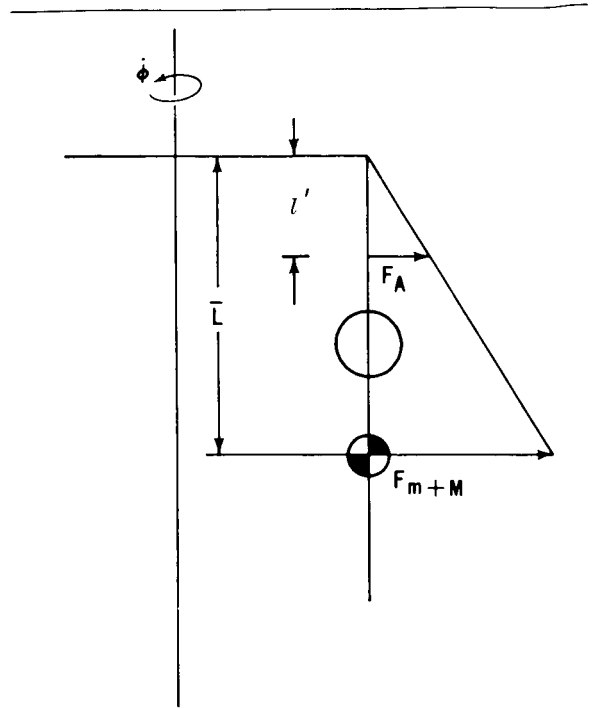


FIGURE 5.—Forces acting on the energy absorber.

where θ_1 denotes the half-cycle amplitude, I_{p_1} the moment of inertia of one half the pallet, and T_{p_1} the torque acting on the pallet. Total pallet amplitude θ is $\theta_1 + \theta_2 = 2\pi/n$, where n is the number of teeth on the scape wheel. For cases where the first and second half-cycle motions are identical, the pallet frequency f is equal to $1/t$ or

$$f = \frac{1}{2} \sqrt{\frac{T_{p_1}}{2I_{p_1} \theta_1}}. \quad (10)$$

The first half-cycle period for the scape wheel is

$$t_1 = \sqrt{\frac{2I_{e_1} \theta_1}{T_w}}, \quad (11)$$

where $T_w/T_p = \omega_p/\omega_n$ and the equivalent moment of inertia $I_e = I_w + I_p(\omega_p/\omega_w)^2$ and T_w is the torque acting on the scape wheel, I and ω the mass

*J. V. Fedor, Private communication.

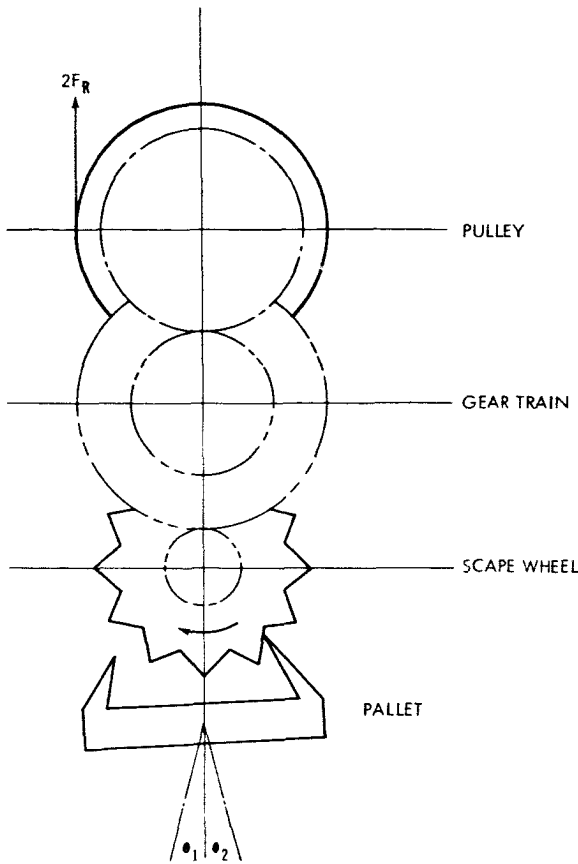


FIGURE 6.—Schematic of escapement design.

moment of inertia and angular velocity respectively. Subscripts p and w refer to the pallet and scape wheel. A similar expression is obtained for the second half of the pallet cycle:

$$t_1 + t_2 = \frac{1}{f} = \sqrt{\frac{2I_{e_1}\theta_1}{T_w}} + \sqrt{\frac{2I_{e_2}\theta_2}{T_w}} \quad (12)$$

In this instance, accuracy in timing was not an essential factor; thus we can simplify calculations by assuming that the motion of the first and second half pallet cycle were the same: $\theta_1 = \theta_2$, $t_1 = t_2$, $T_{p1} = T_{p2}$, $I_{p1} = I_{p2}$, and $I_{e1} = I_{e2}$.

Equations 9, 10, 11 and 12 now become:

$$t_1 = t_2 = \sqrt{\frac{I_p\theta}{T_p}} \quad (13)$$

$$f = \frac{1}{2} \sqrt{\frac{T_p}{I_p\theta}} \quad (14)$$

$$t_1 = t_2 = \sqrt{\frac{I_e\theta}{2T_w}} \quad (15)$$

$$t_1 + t_2 = \frac{1}{f} = \sqrt{\frac{I_e\theta}{T_w}} \quad (16)$$

respectively. To compensate for the inaccuracy of the mechanism owing to the simplified assumptions made in the calculations and to provide for a timing adjustment, I_p was made variable by attaching a counter weight to the pallet (Figure 2). By assuming that the actuating torque is constant and by varying the counter weight which changes I_p , t can be increased or decreased with the I_p changes indicated in the previous equations. The evaluation of T_w and T_p was based on F_a (Equation 8, where $\theta = 60^\circ$).

Gear Train Design

In calculating the stresses for the gears and shafts, the following assumptions were made:

1. The shaft between gears was in torsion;
2. The bearing portions of the shaft were in bending;
3. The pallet shaft was in bending.

The size of the shafts were evaluated by the following equations based on the maximum load determined by Equation 6 (Reference 4):

$$\text{For shear under bending, } D^3 = \frac{16}{S_s\pi} M_b; \quad (17)$$

$$\text{For shear under torsion, } D^3 = \frac{16}{S_s\pi} M_t; \quad (18)$$

$$\text{For tensile under bending, } D^3 = \frac{32}{S_t\pi} M_b; \quad (19)$$

$$\text{For tensile under torsion, } D^3 = \frac{32}{S_t\pi} M_t; \quad (20)$$

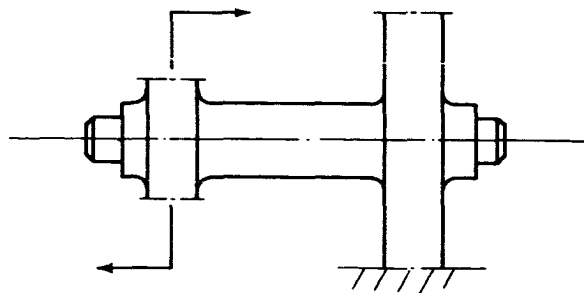


FIGURE 7.—Schematic of the shafts.

where S_s is the shearing stress, S_t the tensile stress, M_b the bending moment, M_t the torsion moment, and D the shaft diameter. From these calculations, the larger value of D was accepted as the final dimension. The design load for the gear train was the same as the shaft design load. Since there were two booms, the total load on the first gear and shaft would be $2F_R$. The load on each succeeding gear and shaft decreased by a factor of $1/R$, where R is the gear ratio. Gear size was governed by the Lewis formula:

$$S = \frac{2M_t P_d}{K\pi^2 yN} \quad (21)$$

where M_t is the torsion moment, P_d the diametrical pitch, y the form factor, N the number of teeth, and $K = P_d b / \pi$. In Equation 21, the allowable stress $S = S_0 (1200 / (1200 + V))$, where S_0 is the ultimate strength and V the pitch-line velocity. The selected gear size must also satisfy the dynamic tooth load equation:

$$F_d = \frac{.05 (bC + F)}{.05 V + (bC + F)} + F, \quad (22)$$

where C is a constant dependent on the material and tooth error and F is the total applied or transmitted load. F_d should not exceed the static load strength of the tooth as based on the static stress of the material. Normally, Buckingham's equation for tooth wear, $F_w = D_b K Q$, must also be considered; however, since this mechanism was expected to operate for only a limited number of cycles, this requirement was neglected.

Restraining Cord

Several cord materials were tested in an effort to find the most suitable. Solid stainless-steel wire was discovered to be most unsatisfactory since the high resiliency of this wire tended to unwind the mechanism; braided copper wires were much more satisfactory, but the size required did not possess sufficient strength. Dacron and nylon cords were found to be best suited to this equipment. Test performed with same size cords showed that the dacron has a higher strength than nylon; but because nylon has the ability to stretch more readily, it was found to possess better shock-absorbing characteristics. A 96-lb. test line of braided nylon was selected.

TESTING

To determine the adequacy of the design, both structurally and functionally, two component tests besides the systems environmental tests were utilized (Reference 5). The design passed the environmental tests without any major difficulty. One minor incident was encountered during the 600 cps resonance vibration test in the axial direction as the cords unwound from the pulley. The problem did not occur after the cord retaining springs were added (Figure 2). To determine the load capacity of the mechanism, a fixture with simulated booms was mounted (Figure 8). Dimensions A, B, and C were approximately the same as those on the actual spacecraft. The weights, equal to the maximum calculated centrifugal force acting on the booms, were attached to the cg of each boom. Calculations indicated that centrifugal force on the booms is maximum when $\theta = 35^\circ$; therefore distance c was selected so that the cord from the weight would be normal to the boom when the booms are 35° from the vertical axis. The booms were held in place by wrapping a cord around them and were released by cutting the cord, thus simulating the loading condition at the time of boom release. This test revealed several design weaknesses, which have been corrected accordingly.

The second component test was to determine whether the design functioned as intended. Figure 9 illustrates the test setup: The spacecraft structure (less the top dome) was mounted on the separation unit, which in turn was attached to the "dutchman." This configuration was then fastened to an empty X-248 bottle, and the system was inverted and mounted on a spin table with the roll moment of inertia properly simulated. The booms were held in position with scissor clamps (similar in design to proposed flight components); pressure was applied to the clamps through a wire wrapped around the legs of these clamps. The ends of this wire were locked in place by a pyrotechnic-actuated release mechanism. To measure the boom erection time, micro switches were placed at the folded and extended positions.

To simulate zero g, a constant force spring with a pull equal to approximately the weight of the boom was first tested by attaching one end of the spring to the cg and was later fastened to a point where equilibrium was achieved—between the

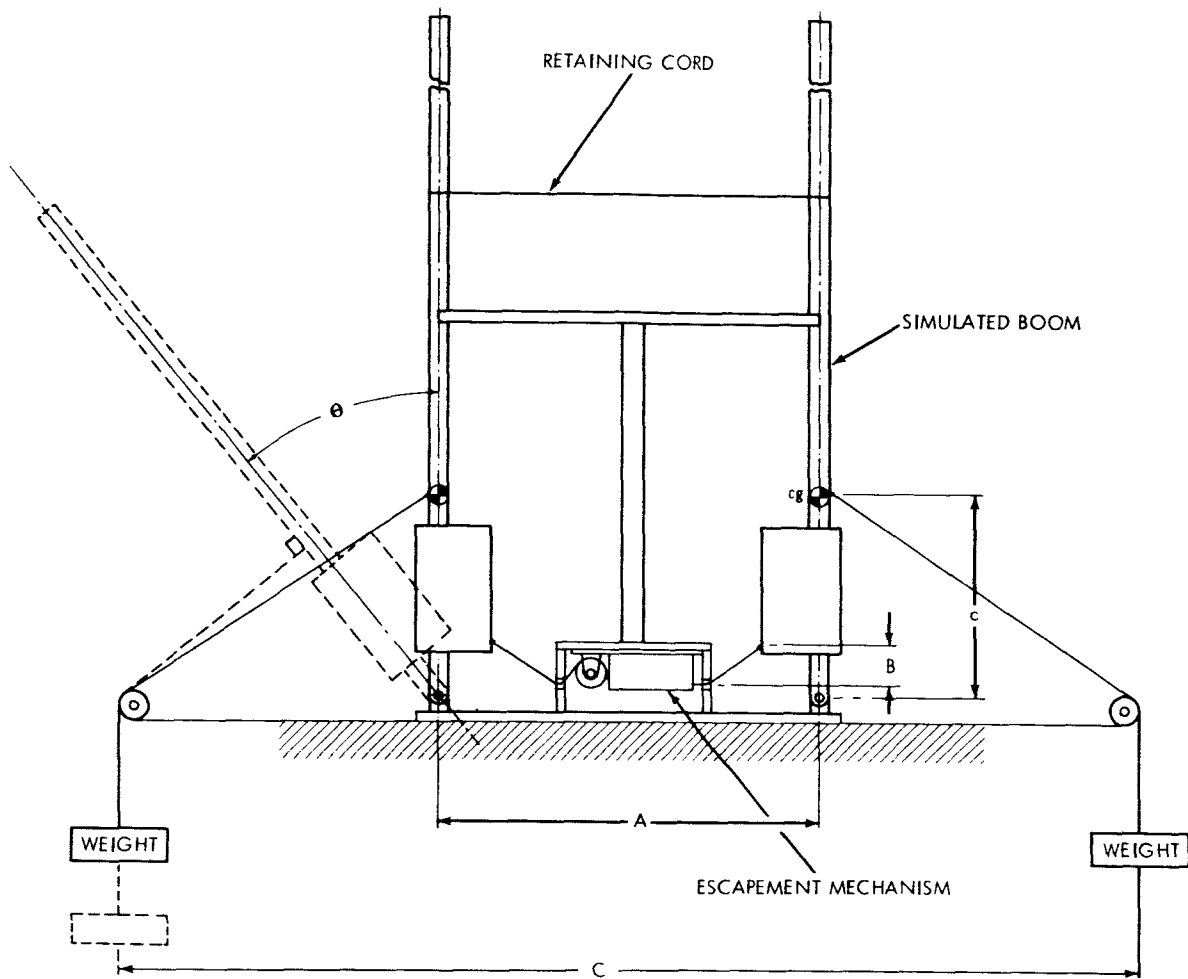


FIGURE 8.—Component test setup simulating initial loading condition on escapement mechanism.

weight of the boom and the pull of the spring—when the booms were very near the fully extended position. The booms were tested at 63, 74, and 90 rpm. Significant results were attained when the springs were attached to the boom's cg. Deployment time was within the expected range, and the booms locked in position each time. With the spring fastened to the second position described above, erection time remained approximately the same, but the booms were about 1/16 in. from lock-in position. Upon investigation to ascertain the reason for the difference in performance between the two systems, it was discovered that, with the spring attached to the cg, equilibrium condition could be obtained when the booms were between 15 and 75 degrees from the vertical axis. When the booms were at positions outside this

range, however, the weights of the booms predominated and overcame the spring force. When the spring was fastened to the second position described, equilibrium could be attained only when the booms were between 60 and 90 degrees from the vertical axis. If the booms were lying in any position outside this range, the spring force prevailed and pulled the booms toward their folded positions (Figure 9). Neither system offered a satisfactory test inasmuch as the first was inclined to under-test and the second to over-test.

The constant spring force was found to vary as much as 3 pounds depending on the amount of unwinding. Also the inability to duplicate the spin's decay rate affected the test results to some extent. To increase the possibility for complete

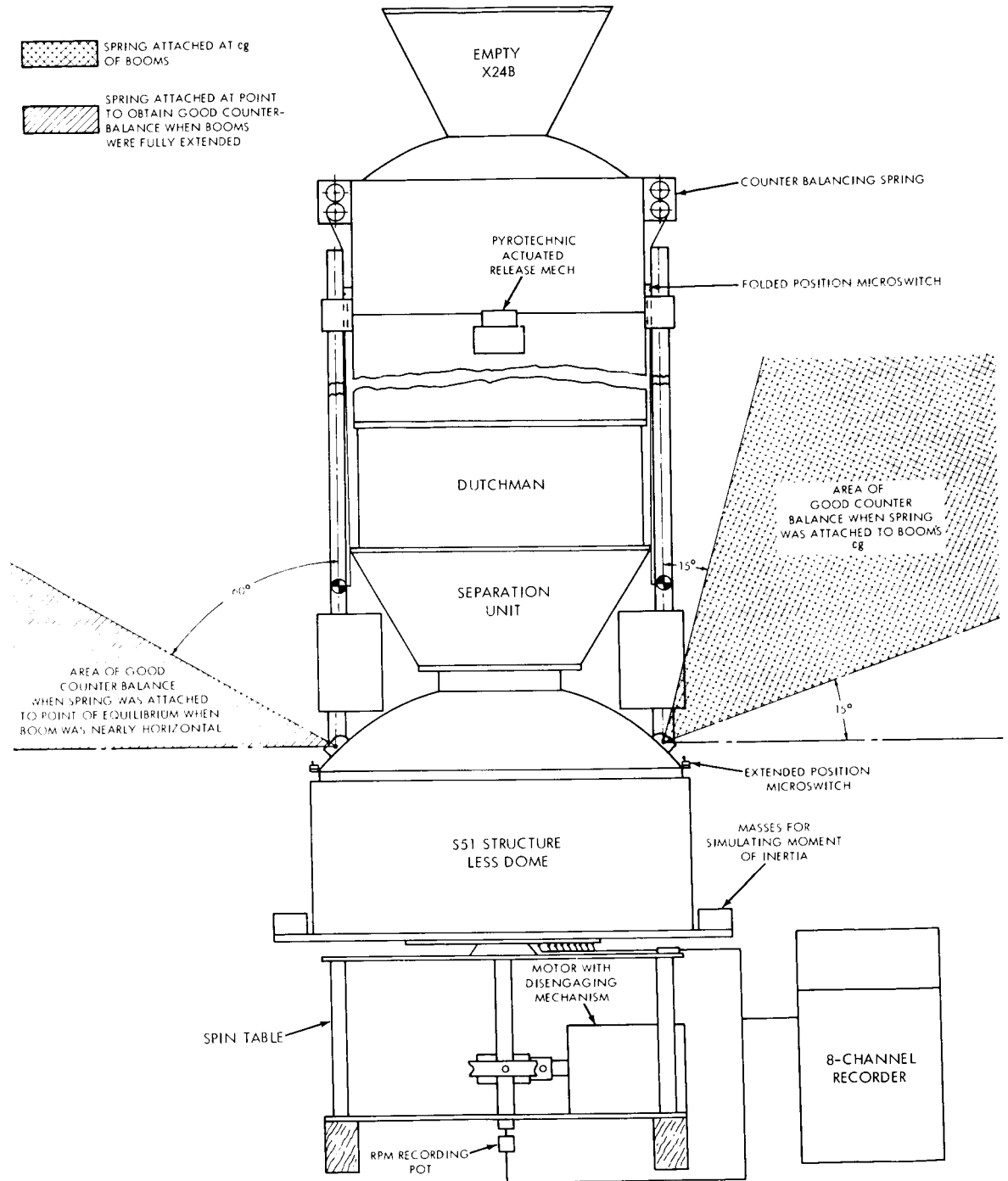


FIGURE 9.—Test setup to determine whether escapement mechanism met design requirement.

boom deployment, the boom hinge spring was made stronger so that the spring preload was approximately twice the torque needed to actuate the escapement mechanism. In the event that

the centrifugal force near the end of boom travel may be too weak to overcome the inertia of the energy absorber, the spring would be able to assist in completing this task.

CONCLUDING REMARKS AND RECOMMENDATIONS

Ariel I was successfully launched in the spring of 1962. All boom-mounted instruments were reported working satisfactorily. However, according to Dr. Willmore,* a senior U.K. scientist on the project, "The erection of booms and paddles occurred prematurely about 1 to 2 minutes after injection." The appendages were deployed in four separate events as indicated by the satellite's decreasing spin rate in "four well-defined steps." At the end of each of the first three de-spins "short duration shocks, such as might be caused by the booms snapping into place" were recorded from the accelerometers mounted orthogonally inside the dutchman. "For the fourth de-spin, where there is no shock, the duration of de-spin is slightly greater. The moment of inertia change corresponds with the erection of an experiment boom, so that this boom might have been successfully restrained by the escapement fitted for the purpose. Moreover, some 136-Mc telemetry was also received at Antigua. This shows a sudden and very marked change in the null of the electron density experiment at about this time, confirming that the boom erected here."

It may be also pointed out that the booms were so retained that any abnormal deployment could cause one boom to be released ahead of the other. Dr. Willmore's interpretation of the data indicated that the first boom—electron temperature—was released at a spin rate at least 10 rpm higher than the maximum design rate of 90. This situation probably caused the restraining cord to

break. The electron density boom was released at approximately 90 rpm and was successfully restrained.

With due awareness of the circumstances and observations as quoted above, we are of the belief that this principle can be utilized advantageously in other similar applications. For future designs, it is suggested that a refined rewind mechanism be added to alleviate the rewinding problem. Addition of a locking mechanism that would cage the energy absorber in place during power flight and that could be released by the pull of the restraining cords might be beneficial. The energy absorber should be designed to control only the initial movement of the moving body and allow the body to complete its movement unrestrained by momentum. Such control would facilitate testing and increase the reliability of the design.

ACKNOWLEDGMENTS

The author is grateful to Messrs. J. Kauffman, A. Pierro, and J. Sween for their recommendations and untiring efforts in testing the final product.

REFERENCES

1. FAIRMAN, S., and CUTSHALL, C. S., "Engineering Mechanics," 2nd Ed., New York: Wiley, 1946.
2. STEELE, T. K., "Clock-Escapement Mechanisms," *Product Eng.* 28(1):179-185, January 1957.
3. RAWLINGS, A. L., "The Science of Clocks and Watches," 2nd Ed., New York: Pitman, 1948.
4. NORMAN, C. A., AULT, E. S., and ZAROSKY, I. F., "Fundamentals of Machine Design," New York: Macmillan, 1938.
5. "Environmental Test Specifications for Design Qualification and Flight Acceptance Tests of The International Ionosphere Satellite—S-51," Goddard Space Flight Center, December 6, 1961.

*A. P. Willmore, Private communication.

EXPLORER XVII (1963 9A) REAL TIME PCM TELEMETRY DATA PROCESSING AND DISPLAY TEST STAND*

M. M. GRANT, C. C. STEPHANIDES, AND W. N. STEWART

The Atmospheric Structure Satellite (Goddard Space Flight Center projects S-6), successfully launched on April 2, 1963, was designated Explorer XVII. This spacecraft operated on command, consumed high power, and had the highest PCM telemetry data rate of any scientific satellite at the date of launch. A major problem related to the development, test, and orbital monitoring of this spacecraft was the quick and accurate evaluation of the telemetry data. The problem was solved by incorporating an automatic data processing system as an integral part of the spacecraft telemetry test stand. Two such systems were utilized in the integration, test, and prelaunch phases and for orbital monitoring throughout the scheduled 100-day satellite lifetime.

INTRODUCTION

Constant efforts are being made to improve the design and reliability of contemporary spacecraft electronic systems. The success of these efforts depends upon the ability of the designers to evaluate the performance of their equipment on the ground and in orbit. The employment of the most advanced performance evaluation techniques is necessary in the development of spacecraft systems of optimum design and reliability.

A wide selection of performance evaluation techniques is available, only one of which permits identical evaluations to be made on the ground and in orbit. This technique is the performance evaluation of spacecraft systems via telemetry.

Although in many cases it is desirable to measure a large number of quantities hard-wired from the spacecraft under test, the benefit to be derived by the evaluation of performance via RF telemetered data is believed to be much greater. Performance monitoring via RF telemetered data on the ground forces the designer to evaluate spacecraft operation in a manner similar to that used after the spacecraft is launched. By this method the spacecraft system is tested in its entirety; this includes payload and ground support equipment. The number of quantities measured via hard wires should be minimized since this

information may not be available once the spacecraft is launched. An effort should be made to avoid dependence on any of these measurements, if used, for significant performance evaluation. Included within the concept of spacecraft performance evaluation via telemetered data are the assumptions that the important experimental and housekeeping data are chosen for transmission, and that a telemetry system capable of transmitting these data is employed.

If spacecraft system performance is to be evaluated solely through telemetered data, the importance of an adequate test stand becomes magnified. Early spacecraft projects used combinations of hard-wired and telemetered data checkout systems and required relatively low data rate test stands for the decommutation and display of spacecraft data for evaluation. Their test stands employed slow speed display devices such as lights, meters, analog recordings, and low speed printers.

The advent of the Explorer XVII project (S-6†) imposed new and stringent requirements upon the spacecraft test stand because of the decision that all spacecraft checkouts be made via RF. A sophisticated system for the processing and display of the high data rate telemetered data had to be developed. The telemetry system produced 41 main and 42 subcommutator (subcom) channels

*Published as *NASA Technical Note D-2318*, June 1964.

†Prior to launch the Explorer XVII project was known as the S-6 project, a Goddard Space Flight Center designation only.

of information, exclusive of voltage reference, synchronization, and subcom counter channels. Its bit rate was 8640 bits per second of split-phase PCM, with a basic 48-word, 9 bits per word format.

Early in the project, the most meaningful data were selected to be telemetered from the spacecraft. This done, the project then established the requirements for a telemetry test stand, acknowledging recommendations that a computer be incorporated into the test stand. When the test stand was in operation for over 1 year a high speed computer complex was incorporated into the test stand system. This permitted a vastly more comprehensive spacecraft performance evaluation program; it required minimal manpower to conduct the evaluations; and it improved the efficiency and speed with which the evaluations were made. The digital nature of the telemetered data lent itself to the easy decommutation, formatting, and computer processing required in the test stand and computer system. An on-line computer processed, in real time, the telemetered digital data for presentation in meaningful form on a high speed line printer. Analog recordings were also made, but primarily for waveform checks on data. The accurate measurements needed for proper evaluation of the spacecraft system were obtained from computer-processed printouts. The

test stand processed data accurately and in easy-to-read form, in real time. The real-time feature of the test stand and computer system was one of its most important and useful assets. Real-time data processing had two significant values: It greatly reduced the amount of time required for testing; it permitted the test conductors to remain closely attuned to spacecraft operation.

The test stand system evolved from a simple one, which employed digital lights and analog recordings for display, into that described herein, a highly versatile test stand and computer system. The decommutation and computer interface equipment was designed and fabricated at Goddard Space Flight Center. The computer, its peripheral typewriter, tape units, and medium speed printer were leased from the Control Data Corporation.

Two test stand and computer systems were used for the checkout of the prototype and flight unit spacecraft during all phases of development. The first system, located in the Technical Control Center (TCC) at Goddard Space Flight Center, was used for the checkout of the spacecraft during integration, environmental testing, and real-time monitoring in orbit (Figures 1 and 2). A second test stand and computer system was assembled and used for a short period of time, exclusively,

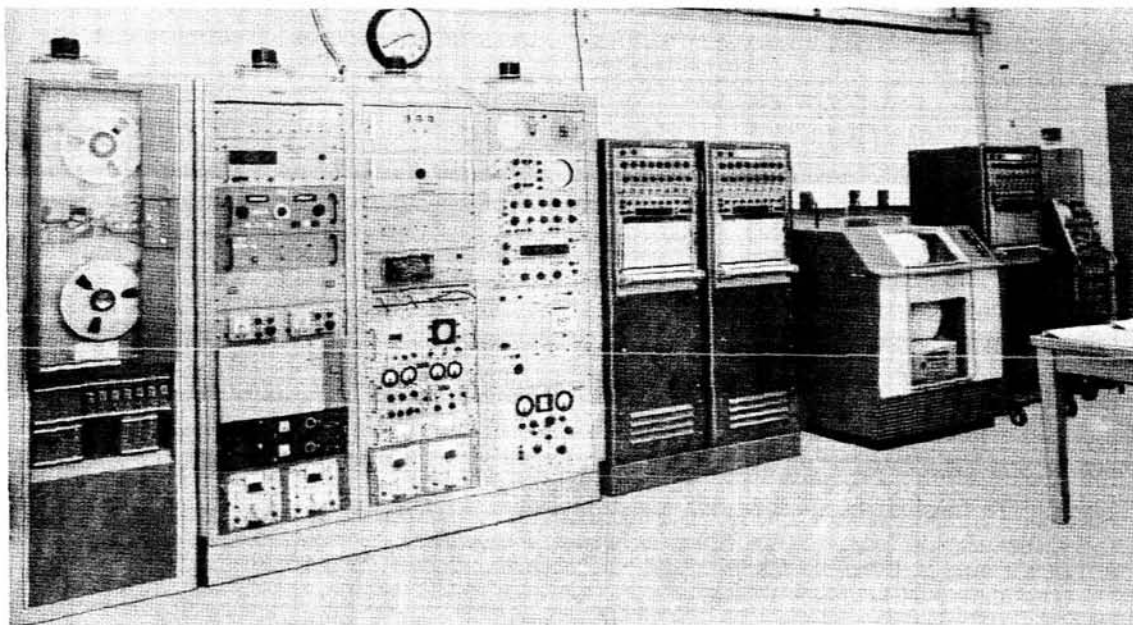


FIGURE 1.—Technical control center test stand.

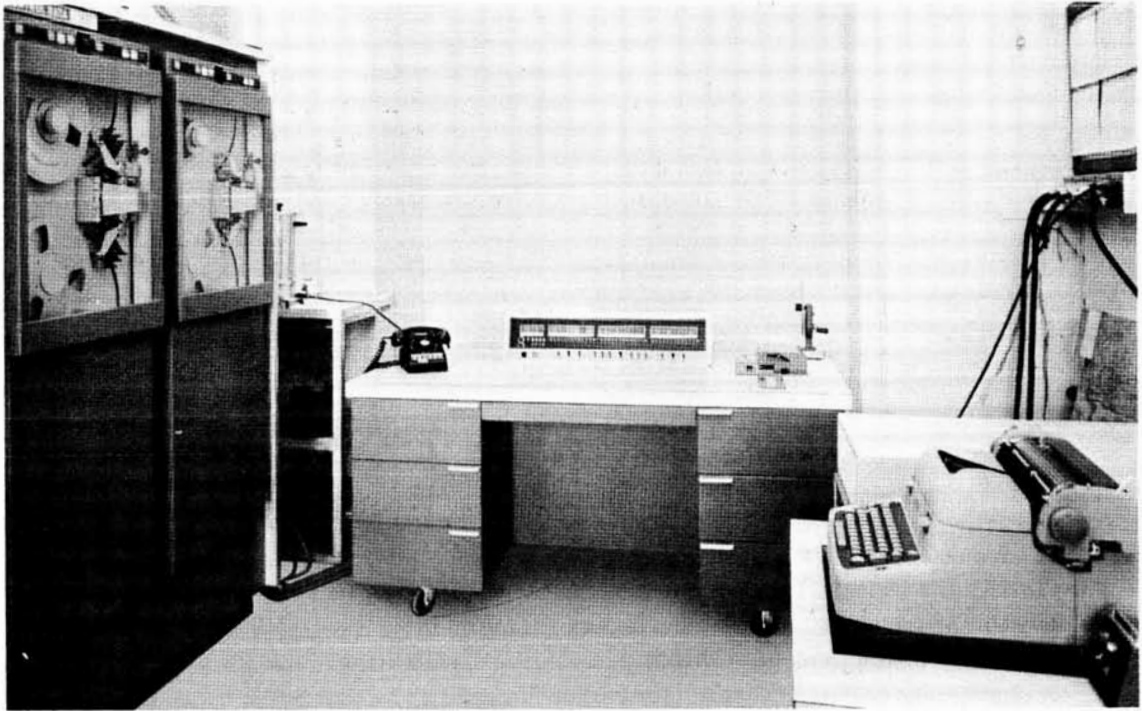


FIGURE 2.—Technical control center computer complex.

in the prelaunch checkout of the spacecraft at the Atlantic Missile Range. This system was dismantled shortly after launch.

It is the purpose of this report to describe in detail the development of the test stand and computer system and to discuss the spacecraft computer programs. Appendixes A through E cover telemetry assignments, computer interface electronic operation, computer program flow charts, computer printout formats, and computer program.

TELEMETRY FORMAT

This section is included in order to facilitate understanding of discussions to follow regarding the test stand and computer operations. As has been previously mentioned, the PCM telemetry operated at a bit rate of 8640 bits per sec in a serial data train of 9-bit data words (Figure 3).

Telemetry Main Frame

Telemetry main frame (Figure 3) was made up of 48 nine-bit words (channels). Word 1, as depicted in Figure 3, carried in digital form the subcom counter information, the experiment selector switch position indication, and an experiment calibration flag. Words 8 and 32 were optical

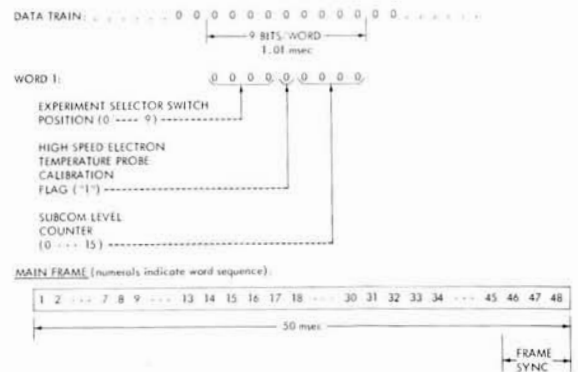


FIGURE 3.—Telemetry words and main frame.

aspect digital information, and words 46, 47, and 48 comprised a fixed 27-bit frame synchronization code. All other words contained analog information encoded into binary digital form, words 14 and 15 assigned to the main channel 0 and 5 volt references, respectively (see Appendix A).

Telemetry Subcom Frame

Words 17, 31, and 33 of the main frame were subcommutated channels. For example, if the main frame depicted in Figure 3 was frame 1

(level 1 indicated in a portion of word 1), words 17, 31, and 33 would carry sample sources 17-1, 31-1, and 33-1, respectively. For the next frame in sequence, word 1 would indicate level 2 and word 17 would sample source 17-2, word 31 would sample source 31-2, and word 33 would sample source 33-2. The level indications and subcommutation would change synchronously with each succeeding frame until 16 frames had been trans-

mitted; the sixteenth frame would contain level 16 in word 1 and words 17, 31, and 33 sampling sources 17-16, 31-16, and 33-16, respectively. In the next sequential frame, word 1 would again indicate level 1 and the subcom words would repeat their sample sequence, beginning with sources 17-1, 31-1, and 33-1. This is shown graphically in Table 1, which represents the telemetry matrix.

Table 1
Explorer XVII Telemetry Matrix

Level	Main Channel Number							
1	1	17-1	31-1	.. 33-1	48
2	1	17-2	31-2	.. 33-2	48
3	1	17-3	31-3	.. 33-3	48
4		17-4	31-4	.. 33-4	
5		17-5	...	31-5	.. 33-5	
6		17-6	.	31-6	. 33-6	
7		17-7		31-7	33-7	
8		17-8		31-8	33-8	
9		..	17-9		31-9	33-9	
10		.	17-10		31-10	33-10	...	
11		..	17-11		31-11	33-11	
12		17-12		31-12	33-12	
13		17-13	.	31-13	. 33-13	
14		17-14	..	31-14	.. 33-14	
15		17-15	31-15	.. 33-15	
16	1	17-16	31-16	.. 33-16	48

The subcom frame was defined as 16 sequential main frames beginning with frame 1. The subcom frame period was determined from the repetition rate of the individual subcom sources (one sample every 16 main frames). This period equalled 0.8 sec and resulted in a subcom rate of 1.25 samples/sec.

Telemetry Supercommutation

Information definition for some sources was improved by increasing the sampling rate of specific sources through supercommutation; that is, the assignment of more than one main channel (word) to a single source. For example, words 10, 26, and 42 were assigned to Redhead No. 1

ac output, a single source. The resultant sample rate for that source, at 3 samples/frame, was 60 samples/sec. Appendix A lists all the channel-to-source assignments for Explorer XVII and indicates their respective sample rates.

TEST STAND AND COMPUTER SYSTEM DEVELOPMENT

Early System

The first complete Explorer XVII PCM test stand is shown in Figure 4. Its main components were:

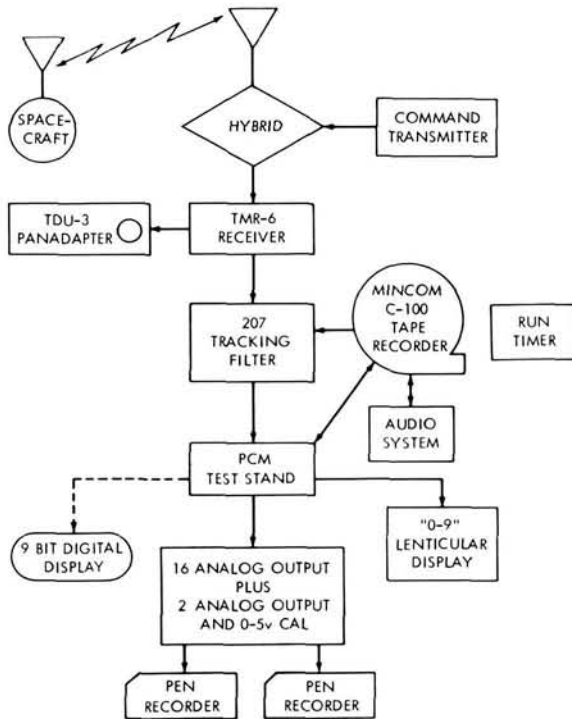


FIGURE 4.—Block diagram of early test stand system.

1. A three-tone command transmitter designed and constructed at Goddard Space Flight Center.
2. A Defense Electronics TMR-6 receiver and TDU-3 panadapter.
3. An Electrac 207 tracking filter.
4. A Mincom C-100 magnetic tape unit.
5. A PCM test stand decommutator (decom) designed and constructed at GSFC.

6. An audio recording and playback system for use with magnetic tape, designed and constructed at GSFC.
7. Two 8-channel pen recorders.

Test Stand Decom

The decom was originally designed as a prototype of models to be used at NASA's Woomera, Australia, tracking station and the central data processing facilities; it was delivered to the Explorer XVII integration group in June 1961. This decom was a basic telemetry decommutation system capable of accepting serial split-phase PCM data from a telemetry receiver. As a system it processed the incoming signal, recovering data and producing analog and digital outputs of selected data channels. A patchboard (Figure 5) was provided for these selections. In addition, the binary data relating to the position of the spacecraft experiment selector switch was gated to a 10 numeral lenticular display (Figure 5).

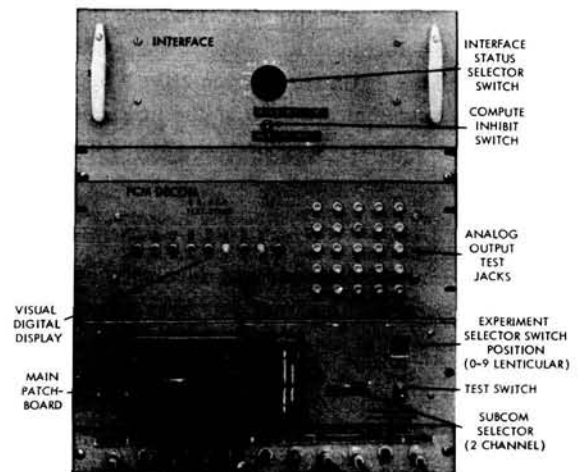


FIGURE 5.—Test stand decom and Mod II interface.

As the project progressed, additional refinements were incorporated into the PCM decom by the integration group, so that the unit was better adapted to the integration and testing of the spacecraft. Noteworthy among these refinements were: Automatic frame synchronization of the decom, an automatic strobe frequency generator (ASFG), a 0 and 5 volt calibration for analog outputs, two additional digital-to-analog convert-

ers, a visual digital display, digital parallel output, and signal conditioning equipment.

Automatic Frame Synchronization System

The synchronization system, designed and installed at Goddard Space Flight Center, substantially enhanced the operation of the PCM decom. Prior to the installation of this system, the services of one man were required to maintain the proper phase relation between the incoming and the timing data. The phase selection was made by repeated random manual switching between timing pulses, which occurred in either of two data phases or two inverted data phases (the manual switch is labeled "test switch" in Figure 5). This switching continued until a valid sync pattern was recognized and signal lock-on was achieved by the decom.

This problem was subsequently eliminated by the addition of an automatic frame synchronization circuit which consisted of two selectable, coded, 27-bit frame sync recognizers; two threshold detectors; one flip-flop; and associated gating circuits. One frame sync recognizer always was set to recognize the frame synchronization format and the other recognizer was set to recognize the inverse of that established code. The recognizers were manually set to their respective codes by the actuation of 27 switches that connected the main shift register flip-flops to the sync recognizers, one side of each flip-flop to "sync" and the other side to "not sync." These recognizer codes were "sync": 000110010011111000101011010; and "not sync": 111001101100000111010100101. The summing node of each recognizer was connected to its associated threshold detector. The outputs from both of the latter were used to reset word counters and other devices. The output of the "not sync" detector was also used to toggle a flip-flop that gated the proper timing pulses to return the decom to a synchronized condition. This autosync system required only 2 frames (100 msec) to regain frame sync after a dropout, whereas the same operation performed manually might have consumed a time on the order of seconds. Adoption of the automatic frame synchronization system sharply reduced the amount of data lost due to sync dropouts.

Automatic Strobe Frequency Generator (ASFG)

This circuit, designed and developed by the integration group, further reduced the number of sync dropouts and resultant data loss by generating a strobe frequency which automatically tracked the incoming PCM signal. The ASFG consisted of a controlled astable multivibrator and associated circuitry. It developed timing pulses at twice the bit frequency and maintained constant phase relation between that frequency and the frequency of the incoming data over a 100 cps bandwidth. Prior to adoption of the ASFG in the test stand system, a harmonic wave analyzer was used, with some difficulty, to extract the second harmonic of the bit frequency and develop a strobe frequency twice as great, and to provide tracking of second harmonic frequency variations over a 3.5 cps bandwidth. The ASFG increased this bandwidth and greatly reduced data deterioration resulting from bit-rate fluctuations introduced by variations in tape speed and system parameters.

Automatic Calibrator

The analog outputs of the PCM decom were all initially based on 0 and 5 volt references established in the spacecraft and transmitted in the telemetry data train; all analog outputs were scaled against these references. As the latter were "floating" and subject to variations (particularly from one spacecraft telemetry encoder to another), the accuracy of the analog pen recordings was questionable and visual data reduction of them was difficult.

The automatic calibrator (autocal) was devised by the integration group to simultaneously scale and calibrate all analog outputs by providing 0 and 5 volt reference outputs for use in setting up each pen recorder channel, with full scale representing 5 volts. Major components of the autocal device were astable multivibrators, relay drivers, and test switching circuits. Multivibrators were used to switch, via relays, all decom analog output lines between normal data outputs and the reference voltage outputs—which in turn were switched by a multivibrator-relay combination between the 0 and 5 volt references. Manual switches provided selection of several calibration modes. For setting up the recorder channels, the first calibration mode, a steady square wave of 0 and 5 volt

reference pulses (Figure 6) was delivered to all recorder channels, the data for these references in general being derived from magnetic tape recording of the spacecraft telemetry. For the most accurate analysis of the pen recordings, reruns were made with the recorders scaled from the references of the particular run desired. The second calibration mode was an occasional interruption of normal data flow with a small group (1 to 2) of scaling pulses. The repetition rate of these calibration bursts was preset at 1 calibration burst every 30 sec to provide minimum interference with data flow and convenient referencing for the analyst. The scaling pulses could be inserted at two rates, one of which is shown in the left (data) portion of Figure 6.

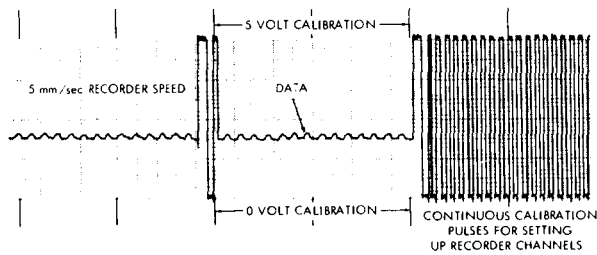


FIGURE 6.—Strip chart illustrating 0 and 5 volt calibration pulses.

Visual Digital Display

While the optical aspect and telemetry systems were being integrated and tested, the need arose for digital display of selected channels of the incoming data. The visual digital display, consisting of a 9-bit storage register, nine light indicator circuits, and the gating required to maintain compatibility with the PCM decom, was appended to the test stand (Figure 5). Channel selection for the display was made at the PCM decom main patchboard (Figure 5) and permitted display of one word or of several words superimposed on one another.

Other Changes

Also during integration, it became apparent that the original analog storage circuits were not able to store data for one complete subcom cycle (0.8 sec) without decay. This situation was rectified by adding two separate digital-to-analog converters, along with an auxiliary patch panel

for channel selection. (This patch panel was mounted in the PCM decom bucket, behind the output test jacks in Figure 5).

An RF link was used between the spacecraft and the test stand during all environmental tests. Adverse signal conditions, such as those expected in orbit (attenuation and noise), were simulated from the outset of this phase of the project. It became apparent that the tracking filter provided a marginally acceptable output under poor signal conditions. Therefore, in April 1962, a Dynatronics 5202 bit synchronizer and signal conditioner was added to the system (Figure 7). This new unit greatly improved data recovery in adverse signal conditions.

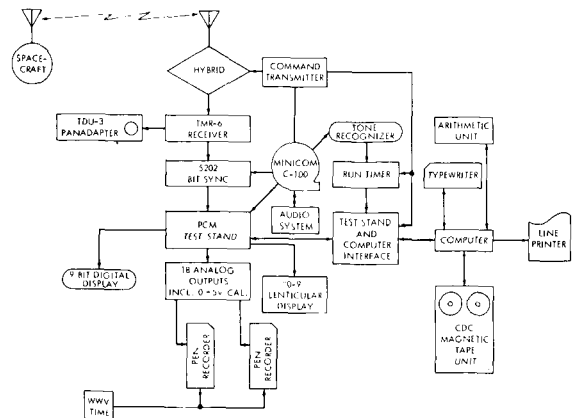


FIGURE 7.—Advanced test stand and computer system.

Initial Computer Installation

As integration progressed in the spring of 1962, it became increasingly apparent that the methods employed for voltage determination of the house-keeping telemetry channels were inadequate. The hand reduction of this information was a lengthy process, requiring more than an hour per commanded run to reduce readouts averaged from a visual digital display. Shortly after installation (May 1962), a CDC 160 computer was programmed to replace the notation-and-slide-rule conversion to volts. One averaged readout per run from each subcom channel was entered in Flexwriter code by punch tape and input to the computer. The conversions were output on punch tape and then typed out in column format by the Flexwriter (Table 2), "cutting" approximately 1 hour from the manual process.

Table 2

Subcom Data in Volts (June 29, 1962).*

Subcom Level	Subcom Data (volts)		
	Channel 17	Channel 31	Channel 33
01	3.97	3.98	0.00
02	4.15	4.32	5.00
03	4.43	2.72	3.56
04	1.61	2.91	3.34
05	1.38	3.64	1.47
06	2.27	3.41	1.54
07	2.30	0.71	3.11
08	2.33	0.69	2.75
09	4.07	1.02	0.00
10	1.85	0.00	0.00
11	4.57	0.00	0.00
12	3.75	0.55	0.00
13	2.26	2.23	0.00
14	2.26	2.91	0.00
15	0.00	5.00	2.52
16	5.00	0.00	1.39

*Note that each subcommutated channel has its own 0 and 5 volt references (levels 1 and 2 or 15 and 16).

Off-line reduction of spacecraft data for integration purposes reached a peak with these developments, and all further efforts were directed toward real-time data reduction, utilizing a computer on-line with the test stand system.

First Test Stand and Computer System (MOD I)

In conjunction with the PCM test stand, a real-time digital data conversion system was developed to improve digital data handling and analysis during spacecraft testing. This system required a *test stand and computer interface*, which was developed in two stages. The Mod I interface had limited capability and acted as an interim system until the Mod II interface could be brought to an operational state of development. Several limiting factors of the Mod I interface were:

1. It "slaved" the computer, preventing the use of any other peripheral equipment for input operations.
2. It required a limited computer program and a fixed computer output format.

3. It employed a gating arrangement that allowed only housekeeping data (subcom channels and spacecraft status) to be extracted from the test stand. The telemetry format chart (Table 1) shows that the spacecraft status word (word 1) was required by the computer to correctly establish the subcom channel levels (subchannels)

The first limitation requires further explanation. In a standard system, several pieces of equipment may be connected in parallel to a CDC 160 input-output channel, along which the computer transmits a "select code." Only the appropriate (desired) unit recognizes this code and performs the input or output functions to follow; all other units (in most cases) are inhibited from interfering with the selected communication link. These "selected-inhibited" functions are contained in what is broadly termed a "select package," and such a unit was not incorporated in the Mod I interface. Outputs from the computer were not affected because the interface used only the input portion of the computer input-output channel. The net result of the limitation was to require the physical disconnection of the interface-computer cable for the use of the input capabilities of the CDC 161 typewriter and 164 magnetic tape unit.

Development of this test stand and computer system required the solution of problems in the areas of logic level differences, timing, and synchronization. The digital logic incorporated in the test stand operated at different levels than that of the CDC 160 (or 160A) computers (see Appendix B, Figure B2). Therefore, appropriate logic interfacing circuits were designed to integrate the two systems. Line activity between the test stand and computer was high, with 960 data requests and 960 data responses—plus numerous status requests and responses—occurring every second. This requirement for a high data transfer rate placed utmost importance on inter-leaving the two independent system timing cycles. One of the important factors considered in the integration of these systems was line settling time for data and timing pulses from one device to the other. Data transfer coordination was obtained by properly delaying and gating the inter-system signals. The computer input word was of sufficient length (12 bits) to allow spacecraft data to

be gated on the low-order 9 bits and flagging information on the high order 3 bits.

The Mod I interface permitted the 160 computer to be used on-line with the test stand. After the installation of a CDC 166-2 line printer (120 characters/line, 150 lines/min), the computer was programmed to compute and print out house-keeping voltages in real time. The first successful real-time printout of Explorer XVII spacecraft data was made on July 26, 1962. Because of the time-share requirement for input-output operations and the limitation in printer drum speed, real-time processing was limited to every third subcom frame. Nonetheless, the amount of processed information presented decimally in volts was increased in excess of 100 times over prior methods and, most desirable of all, it was available during spacecraft operation.

Advanced Test Stand and Computer System (MOD II)

The development of a new (Mod II) test stand and computer interface (Figure 5 and Appendix A) and the replacement of the CDC 160 computer with the more sophisticated CDC 160A made possible further refinements and a data sampling rate increase for the automatic processing system. The Mod II interface retained all of the desirable features of the Mod I while overcoming the latter's limitations. Because the Mod II incorporated a "select package" and thus acted as normal peripheral equipment to the computer, it permitted other input devices to be used normally with the computer. The 160A provided extended memory (8192 memory locations) and an input-output buffer channel which, when paralleled with the normal input-output channel, made this advanced system (Figure 7) capable of inputting data from the test stand concurrent with voltage printout. The primary new feature of the improved interface was that all of the telemetry data were made available for real-time processing. Other new features of the Mod II interface and of the computer programs will now be discussed.

Interface Status Transponder

This was the digital decoder-encoder system described in Appendix B. It permitted the computer to interrogate the transponder requesting the status of the interface at the moment. The

latter would reply (via a program selection system) with a 12 bit code that could be used at the discretion of the computer to select a particular program, subprogram, or operation assigned to that code by the programmer. Computer recognition of one bit of the status word resulted in effectual remote start-stop control of computer data processing. Typical status request and response words are shown in Figure 8, along with a subprogram listing for status responses in Table 3.

Table 3

Interface Status Response Selections
(the program details are in the "Computer Programs" section).

Interface Selector Switch Position	Transmitted Program Code	Program Selected
1	0 0 1	Automatic program (Table 4)
2	0 1 0	Subcom program (continuous)
3	0 1 1	Bayard-Alpert program (continuous)
4	1 0 0	Program hold
5	1 0 1	Redhead program (continuous)
6	1 1 0	Electron temperature probe (ETP) program (calibrations)
7	1 1 1	Optical aspect program (continuous)

Interface Selection System

This system allowed the computer to select the test stand and use it as normal peripheral equipment. Described above as a "select package," it was in essence a decoder-encoder device that, upon receipt and recognition of a proper code word and appropriate signals from the computer, transmitted to that unit a 12 bit word in digital parallel form and in computer format, 9 bits comprising spacecraft data from the test stand and 3 bits expressing a preset flagging code for particular words (see Figure 8).

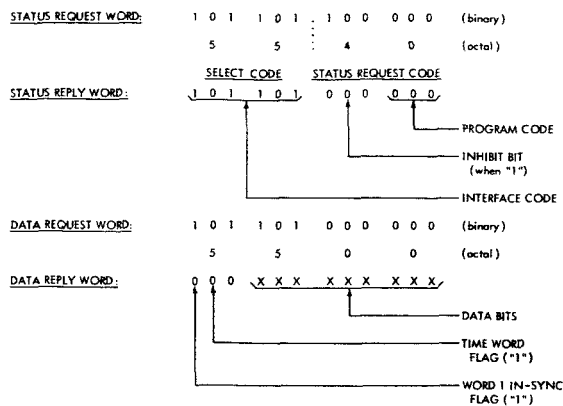


FIGURE 8.—Interface and computer word formats.

Program Running-Timer System

This was a clock system (Appendix B) that was started automatically at spacecraft turn-on time (or at run-start time as defined by a tone recorded on the tape) and counted running time in seconds throughout the spacecraft program cycle. The system consisted of a binary counter clock driven by 1 sec pulses from a modified Hewlett-Packard 523CR run timer, a storage register, and associated gating circuits. By an appropriate gating arrangement, one of the frame synchronization words (word 46) in the data format was replaced by a 9 bit time word taken from the running-time storage register. These data were handled as a part of the normal telemetered data and were transferred to the computer in the low-order 9 bits of the input word. In addition, one of the three high-order bits was preset to a "1" or a "0" to indicate to the computer that data in word 46 was a running-time word or the first frame sync word, respectively (Figure 8). This selection was controlled by a selector switch behind the interface panel and ordinarily was set to transfer time information. In the latter condition, the controlled bit provided a useful 1 bit flag for additional frame synchronization at the computer and thus released the computer program from the necessity of maintaining a rigid word count during lengthy computations.

Thus, the Mod II interface provided all-channel capability, a "test stand in sync" flag, program running-time, and position status of an interface selector switch (Figure 5).

Computer Program Refinements

At this time various new functions were introduced into the computer program. Spacecraft thermistor data were computed in °C, several main channels were periodically processed, and computer subprograms became remotely selectable at the test stand. Refinements continued on the programs, particularly when the line printer speed was increased to 300 lines/min by installation of a print drum with two fields of numerics and one split field of alphabets. With this new capability, it was possible to provide:

1. Continuous printout of optical aspect information in raw digital form.
2. Sampling of alternate subcom frames, an increase of 50 percent in data printed out.

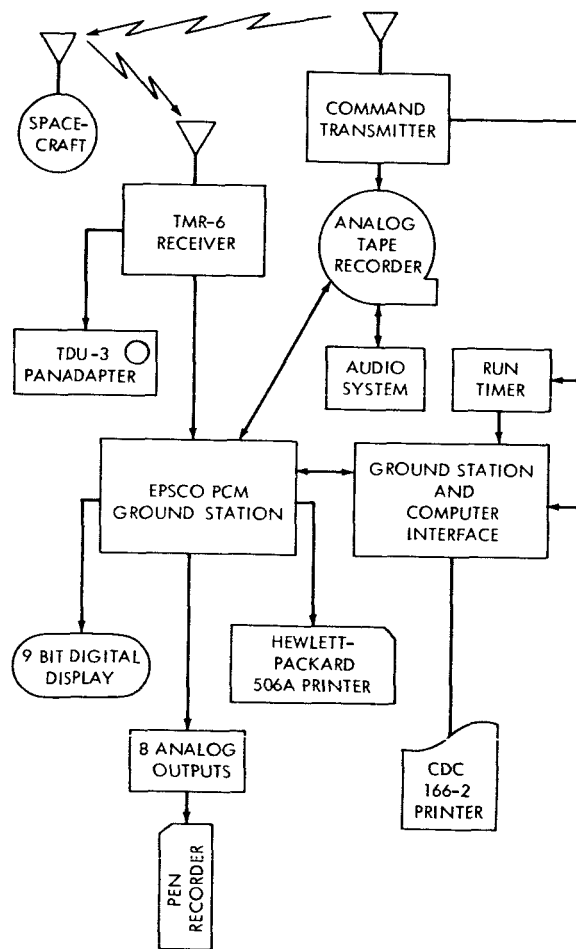


FIGURE 9.—Computer system used at the Atlantic Missile Range.

3. A preselected format (Table 4) to establish a quick-look cross-section of payload operation, control of the format depending on computer reduction of the spacecraft's experiment selector switch (ESS) data.

In addition, similar programs were written for a CDC 1612, 1000 line/min line printer, which was used for some time in the Technical Control Center (TCC). Printout formats were established through integration of equipment capability, programming practicability, and information priority as expressed by the recipients of the

data. Computer program flow charts and computer format facsimiles and descriptions are presented in Appendixes B, C, and D.

System Used at the Atlantic Missile Range (AMR)

In January of 1962 a universal PCM ground station was delivered to the integration group. The PCM test stand and the ground station were used in parallel during the final phases of spacecraft testing. The ground station was interfaced to the CDC 160A computer with the same type of interfacing system as that used on the PCM

Table 4
Automatic Program Format.*

ESS Positions 4, 5, 6, 7, 8	ESS Position 9	ESS Positions 1, 2, 3
T=24	T=24	T=24
Dateline	Dateline	Dateline
Subcoms	Subcoms	Subcoms
55	55	55
ETP II calibrations ETP I calibrations	ETP II calibrations ETP I calibrations	Subcoms
115	115	85
Bayard-Alpert and Redhead calibrations	Subcoms	Bayard-Alpert and Redhead calibrations
145	145	145
Optical aspect	Optical aspect	Optical aspect
175	175	175
Subcoms	Subcoms	Subcoms
205	205	205
Bayard-Alpert and Redhead levels	Bayard-Alpert and Redhead levels	Bayard-Alpert and Redhead levels
210	210	210
Subcoms	Subcoms	Subcoms
T=240	T=240	T=240

*Selection of these vertical sequence formats was determined from computer reduction of the spacecraft experiment selector switch (ESS) position information in telemetry word 1. Each block represents a printer page (see Appendix D), T = running time in seconds.

test stand but with minor changes. The ground station and computer system was shipped to Cape Canaveral (now Cape Kennedy) and used in pre-launch checkout of the spacecraft. This system is shown in Figure 9.

Flight Monitor System

With the exception of the RF data link, no significant changes in the test stand parameters were made for monitoring the spacecraft in orbit. As illustrated in Figure 10, spacecraft telemetry was acquired at the Blossom Point, Maryland, tracking facility and relayed through a Bell System microwave-and-landline link directly to the test stand bit synchronizer. Figure 11 shows the prototype as it was flown to simulate orbital passes over Blossom Point and to check out the described system.

The input-output operations of the computer complex during final checkouts were carried forward, with minor changes, into the flight monitoring phase. Once given a program start, the computer was controlled entirely by peripheral gear until shutdown. Operations for a typical run took the following sequence:

1. The pass number and command time were entered in the 160A computer via typewriter prior to a pass.

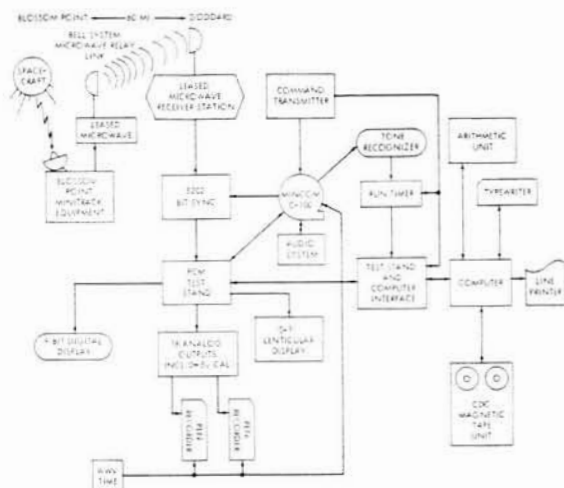


FIGURE 10.—TCC test stand and computer system for real-time monitoring of Explorer XVII in orbit.



FIGURE 11.—Prototype suspended from a helicopter in flight past the Blossom Point tracking and data acquisition facility.

2. Computer-interface operations commenced upon release of a "compute inhibit" switch located at the test stand and, as mentioned, controlled 1 status bit of the interface status word.
3. Initiation of data processing was subject to delays predetermined by run-time, the sync flag, or both.
4. Command counter data (at the subcom rate) were decommutated and converted to a decimal number, which was entered in the run identification line; the latter was then printed out.
5. Subprogram control was exercised at the interface; selections could be changed during the run as circumstances dictated.
6. Computer-interface operations were concluded, and the computer complex was set up for a new pass or rerun of the last pass by closure of the "compute inhibit" switch.

In addition to the data computation and readout functions described, the system was also utilized for the production of IBM-compatible data tapes and for the periodic computation and printout of Explorer XVII battery power drain summaries.

Computer Programs

For the most part, all Explorer XVII test stand computer programs were written in machine language. They began as relatively simple programs for the conversion of spacecraft data to volts and were subsequently expanded and modified during operations through launch. Though far more cumbersome in preparation than those of the assembly type, the machine language programs were comparatively easy to trouble-shoot and modify. The increased use of subroutines and the establishment of known memory locations for data storage allowed a considerable degree of capability and flexibility in assessing difficulties encountered with programs, the test stand interface, and the telemetry data train itself. In addition, a small library of short maintenance routines was established to implement the standard programs in this area. Another advantage derived from the machine language programming approach, as implemented, was the fine determination of

machine time for various loops and subroutines operating in periods where computation time intervals were critical. In this regard, the time required for several conversion operations was substantially reduced through unique logic techniques.

The computer complex was utilized for reduction of real-time or recorded telemetry data originating in the spacecraft. The basic computer operation was to *input* all telemetered data, retaining only those desired for the particular subprogram. Initially, computer frame sync was established through recognition of the frame sync flag on word 1 as sent from the test stand; reduction of the word 1 subcom level counter plus word counters in the computer program maintained orientation of that program in the telemetry matrix. However, problems arose when the conversion times occasionally exceeded the 1 msec word interval. Consequently, conversion operations frequently were delayed for a frame until the desired data and references for that frame were *input*. Once this was done, decision, conversion, and output operations commenced with computer frame sync being re-established through recognition of the time word flag on word 46. In the few cases where data word position in the frame allowed too small an interval prior to word 46, other methods were used to permit successful reduction of at least a large portion of an uninterrupted data chain. For most cases, requested data were rejected when the following conditions existed: Word 1 sync flag was absent; the subcom level counter was not in sequence; or the 0 and 5 volt reference range was illogical or beyond predetermined limits (defined as "illegal"). The rejection of any word was flagged by "***" in that word's normal printout location. Among other decisions, the program refused to convert temperature data to °C when that data indicated thermistor output outside the range of values for which voltage calibrations had been made. In such cases, the computed voltage data were printed out. Printout formats are shown in Appendix D and a typical computer program is listed in Appendix E.

Subcom Program

The subcom program was the most complex of all data subprograms because of the dispersal of the reference voltages throughout the subcom

frame. To establish a compute-start basis, a block of subcom data with legal references was stored in computer memory. The next block in sequence then utilized these references for computation, replacing them with the new references if the latter were legal. Similarly, if a data word was rejected as previously described, its last known value was retained in the printout format. Because of timing errors and telemetry sync dropouts, the retention of "last-known" values occasionally extended for several seconds and masked the event occurrence time on some channels. This difficulty was overcome through the use of the "***" indicator and more precise program timing; the last-known legal references were retained.

Following conversion to volts of all data and decommutation of the running-time data, conversion of the appropriate voltages to °C was accomplished and printout operations begun. Since the latter occupied a major part of the 0.8 sec subcom cycle, and to avoid timing difficulties, data were *input* on alternate subcom frames (for the more recent programs). A careful study of the programs and system capabilities indicates that, in the future, all subcom data of a similar nature can be computed, converted, and printed out in real time.

A typical subcom routine flow chart is depicted in Appendix C, Figure C1, and a flow chart for the digits-to-volts conversion subroutine is shown in Appendix C, Figure C2. The subcom printout format is shown in Appendix D, Table D1. To permit use of the 300 line/min print feature (i.e., minimum use of the split alpha field), temperatures were designated by no unit abbreviation and "-1" was indicated by "***". The programs written for the 1612 1000 line/min printer used "C" for temperature unit indication and "≤" for "-1" indication.

Bayard-Alpert and Redhead Pressure Gage Programs

Bayard-Alpert (BA) and Redhead (RH) experiments utilized main telemetry channels with sample rates of 3 samples/frame (60 samples/sec) for each channel, 4 for the BA's and 2 for the RH's. For checkout purposes, it was felt that reduction of a long continuous stream of data from either experiment was not necessary and that a reasonable real-time approach would be to print

out blocks of 3 or 4 successive frames of BA and RH data in volts. Time selection of these blocks could be exercised at the test stand interface, repeated there for spaced-block presentation, or internally selected by the program through the use of flags, such as the pressure gage calibration flags on sublevel 10 of channel 31.

The programs for these experiments were similar. In both cases, an experiment and source identification line was printed following initial selection of the particular subprogram. (This line was not reprinted if the same subprogram remained selected.) Then, after computer recognition of frame sync, the main channels comprising the reference voltages and the pertinent experiment outputs were selected from the data input stream. Data were computed and stored. The next two frames were similarly processed, with the aforementioned treatment of "no sync" or illegal reference conditions. Printout of the 3-frame block was then initiated, followed by an interface status check to determine whether the subprogram was to be repeated or another route was to be followed. Printout formats for the BA's and RH's are shown in Appendix D, Table D3.

Electron Temperature Probe Program

The subprogram for the electron temperature probes (ETP) was itself separated into two subprograms, one for the fast-sweeping ETP I and one for the slow-sweeping ETP II. Data rates for these experiments were 180 samples/sec and 40 samples/sec, respectively. Because calibration flags for either type may have appeared in the data stream after calibration initiation, and because these calibration cycles were determined to be most significant for proper checkout, a definite, continuously up-dated block of data had to be retained in memory so that the final printout format would encompass the entire calibration period. For ETP II this format also included three or four sweeps of live data following calibration and was repeated once in order to cover both high and low sensitivity calibrations. The format for ETP I did not guarantee inclusion of live data following calibration. To overcome this difficulty, the ETP I format was repeated 3 times after the first calibration cycle printed, with the following resultant sequence of printout blocks: calibration

(high or low sensitivity), live data, calibration (the sensitivity not covered by the first block), and live data for the latter sensitivity.

ETP II Subprogram

After printout of a dateline, the ETP program was first routed to the ETP II subprogram. The ETP II main channels and pertinent references were selected from the data stream and stored in such a manner that the most recent successive 80 data points and their references were retained in memory, with no consideration at that time for their legality. Upon recognition of the ETP II calibration flag, 240 more successive data points with references were stored in memory, sequentially following the original 80 points. Processing and printout *format* began on these 320 stored successive data points in the same manner as previously described. Printout was accomplished (Appendix D, Table D2), and the same routine was repeated through the next calibration. Control of the program was then routed to the ETP I subprogram.

ETP I Subprogram

This program was similar, with two significant exceptions, to that for ETP II. The first dissimilarity was that only one frame (9 data points) was kept current in memory until the ETP I calibration flag (1 bit on word 1) was recognized and only 7 successive frames were stored following that recognition. Another difference was that the ETP I process-printout cycle was repeated 4 times as described. These 4 blocks appear in Appendix D, Table D2, following the 2 blocks of ETP II data.

Optical Aspect Subprogram

For optical aspect (OA) a considerably different type of processing was utilized than that for the experiments. This spacecraft subsystem output digital data on two main channels (40 samples/sec), and proper analysis of these data required the presentation of all successive frames, including those in which sync was lost, for a minimum of 30 sec and a maximum limited by spacecraft operating time. Upon installation of the CDC 1612 1000 line/min printer, an acceptable OA program became a possibility. Prior to that time, limited printout capability had inhibited all-data

presentation of the OA information. For the 1612, two lines of 5 frames (individually processed into "1"s and "0"s) were transmitted in printer code to that device, and printout of those two lines was initiated, with an immediate program return to the data-search routine to acquire the next sequential frame. Sync loss was represented by all "0"s (an illegal OA code) for the pertinent frames.

After an analysis of the CDC 166-2 line printer with 300 line/min capability, it was determined that it could be utilized for an all-data OA program. Instead of a block of 10 successive frames *input* for storage and printout, an 8 frame block, in two printer lines, was incorporated. Thus, the data time for each line (200 msec for 4 frames) equalled the line-print time at 300 lines/min, or 1 line every 200 msec, and resulted in effectual synchronization between the optical aspect output and printout.

The OA printout format is shown in Appendix D, Table D4. With 12.4 sec of data represented on each printer sheet, an all-data printout of optical aspect would consume twenty sheets for a typical run. Despite this length, the format presented no obstacles for accurate and convenient analysis.

Automatic Program

This subprogram was of considerable value during final integration and checkout of Explorer XVII, particularly during closely spaced sequential operations. It permitted a quick-look cross-section of payload operation, enabling testing to continue without the delays imposed by long post-run analyses. Basically a routing scheme, the automatic program was also required to operate on the data stream for the determination of time, events, and position of the ESS. These determinations were then utilized in establishing sequences for entrances to and exits from the other selectable programs. In essence, the automatic program used the experiment, subcom, and OA programs as subroutines. Format sequences appear in Table 4.

Concluding Remarks

Test stand and computer systems have been used to provide data for the Explorer XVII spacecraft performance evaluation by processing

telemetered data in real time. This equipment was used to monitor Explorer XVII during the integration, environmental testing, and prelaunch checkout phases of spacecraft development. Later, it played a vital role in daily real-time monitoring in orbit.

Future plans call for the expansion of the spacecraft test stand and computer system to provide control signals for artificial experiment excitation of the spacecraft under test while the telemetered responses from the spacecraft are being automatically evaluated. This capability will "close the loop" on spacecraft testing. Such tests in the past have been carried out by the operation of experiment excitation equipment by technicians at the spacecraft and the remote evaluation of processed telemetered responses by other operators. Testing in this fashion resulted in difficulties both in the synchronization of efforts

between the exciter and the monitor operators, and in test reproducibility. Computer control of experiment excitation and monitoring will establish means by which experiments will be efficiently tested in a reproducible manner, will better simulate orbital conditions, and will eliminate human variability in test operations.

Finally, advanced test stand and computer systems are being programmed to satisfy such requirements as: Special quick-look data processing of the experimental data requested by experimenters before and after launch; power studies for power supply design and lifetime predictions early in spacecraft projects; generation and up-dating of spacecraft cabling and wiring tables; automatic and remote testing of spacecraft systems; generation of command schedules; and back-up digitizing of telemetry data for special data processing and information display.

APPENDIX A

TELEMETRY ASSIGNMENT FOR EXPLORER XVII

Main Commutator

Measurement	Channel Assignment	Sample Second
Subcommutator Position (digital) -	{ Last 4 bits First 4 bits Middle bit	20
Programmer Position (digital)-----		
Electron Temperature Probe High		
Frequency Calibration Flag-----		
Electron Temperature 1 Output----	2,7,12,18, 23,28,34,39,44	180
Electron Temperature 2 Output----	16,40	40
Bayard-Alpert Gage 1:		
dc-----	3,19,35	60
ac-----	4,20,36	60
Bayard-Alpert Gage 2:		
dc-----	5,21,37	60
ac-----	6,22,38	60
Aspect (digital)-----	8,32	40
Mass Spectrometer 1		
Log Amplifier Output-----	9,25,41	60
Electrometer Biased Output--	24	20
Redhead Gage 1-----	10,26,42	60
Redhead Gage 2-----	11,27,43	60
Mass Spectrometer 2		
Log Amplifier Output-----	13,29,45	60
Electrometer Biased Output--	30	20
Zero Volt Reference-----	14	20
Five Volt Reference-----	15	20
Subcommutator 1 Output-----	17	20
Subcommutator 2 Output-----	31	20
Subcommutator 3 Output-----	33	20
Frame Synchronization-----	46,47,48	60

**Subcommutator Deck 1 (appears on channel 17
of main commutator)**

Measurement	Channel Assign- ment	Word 1 Last 4 Bits	Subcom Counter Count
Mass Spectrometer 1:			
Ionization Current and Voltage.....	1	1111	15
High Voltage Monitor.....	2	0000	0
Repeller Monitor.....	3	0001	1
Amplifier B plus Monitor.....	4	0010	2
Mass Marker.....	5	0011	3
Log Amplifier Temperature-Calibration Flag.....	6	0100	4
Electrometer Temperature.....	7	0101	5
Mass Spectrometer 2:			
Ionization Current and Voltage.....	8	0110	6
High Voltage Monitor.....	9	0111	7
Repeller Monitor.....	10	1000	8
Amplifier B plus Monitor.....	11	1001	9
Mass Marker.....	12	1010	10
Log Amplifier Temperature-Calibration Flag.....	13	1011	11
Electrometer Temperature.....	14	1100	12
Zero Volt Reference.....	15	1110	13
Full Scale (5 volt) Reference.....	16	1110	14

**Subcommutator Deck 2 (appears on channel 31
of main commutator)**

Measurement	Channel Assign- ment	Subcom Counter Count
Bayard-Alpert Emission Current 1.....	1	15
Bayard-Alpert Emission Current 2.....	2	0
Bayard-Alpert Gage Temperatures 1.....	3	1
Bayard-Alpert Gage Temperatures 2.....	4	2
Bayard-Alpert Electrometer Temperatures 1.....	5	3
Bayard-Alpert Electrometer Temperatures 2.....	6	4
Bayard-Alpert Voltages, Bias 1.....	7	5
Bayard-Alpert Voltages, Bias 2.....	8	6
Sphere Pressure.....	9	7
Pressure Gage-Calibration Flags.....	10	8
Experiment Squib Monitor.....	11	9
Turn-on Counter.....	12	10
Telemetry Voltage (+20.15 v).....	13	11
9.3 Volt Battery Monitor.....	14	12
Full Scale (5 v) Reference.....	15	13
Zero Volt Reference.....	16	14

**Subcommutator Deck 3 (appears on channel 33
of main commutator)**

Measurement	Channel Assign- ment	Subcom Counter Count
Zero Volt Reference.....	1	15
Full Scale (5 v) Reference.....	2	0
Bayard-Alpert Card Temperatures 1.....	3	1
Bayard-Alpert Card Temperatures 2.....	4	2
Bayard-Alpert Ion Trap-Filament Monitor 1.....	5	3
Bayard-Alpert Ion Trap-Filament Monitor 2.....	6	4
6.2 Volt Battery Monitor.....	7	5
-27.9 Volt Battery Monitor.....	8	6
Thermistor 1 (low mass spectrometer boss).....	9	7
Thermistor 2 (skin, lower 45 degrees).....	10	8
Thermistor 3 (ambient, top of Redhead stack).....	11	9
Thermistor 4 (skin, equator).....	12	10
+13.95 Monitor.....	13	11
Thermistor 5 (skin, upper 45 degrees).....	14	12
Thermistor 6 (upper mass spectrometer boss).....	15	13
Electron Temperature Probe High Temperature Calibration Flag.....	16	14

APPENDIX B

COMPUTER-INTERFACE OPERATION

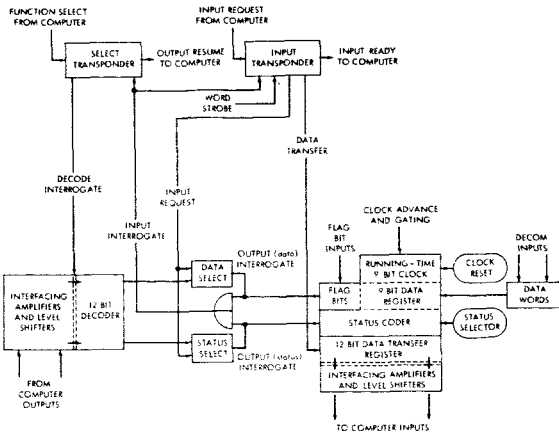


FIGURE B1.—Computer-interface operation.

Computer-interface operation can best be described by discussing the complete cycle of operation (Figure B1). The computer initiates an external function command that sends a function select pulse and the associated 12 bit select code to the computer interface. Upon receiving the function code and after a short period of line settling time, the decode interrogate pulse turns on the decoder. If the proper code is detected, an input interrogate pulse is generated which turns on the input transponder and completes the cycle of the select transponder by developing an output resume pulse. The output resume pulse is sent back to the computer, and upon completion of the computer's output* cycle the select transponder's cycle is terminated. Figure B2 shows the timing involved in operations for this and the following paragraph.

Unless the interface is deselected, it is now ready to send either data or status information to the computer, depending on which selector is interrogated by the code sent from the computer. The combination of the input request pulse and the input interrogate pulse (received during the select cycle) turns on the input transponder.

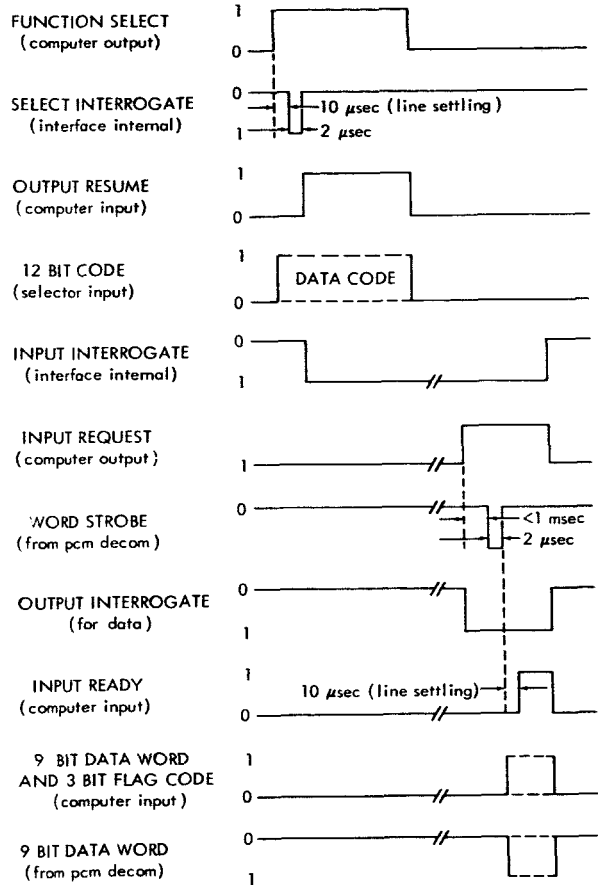
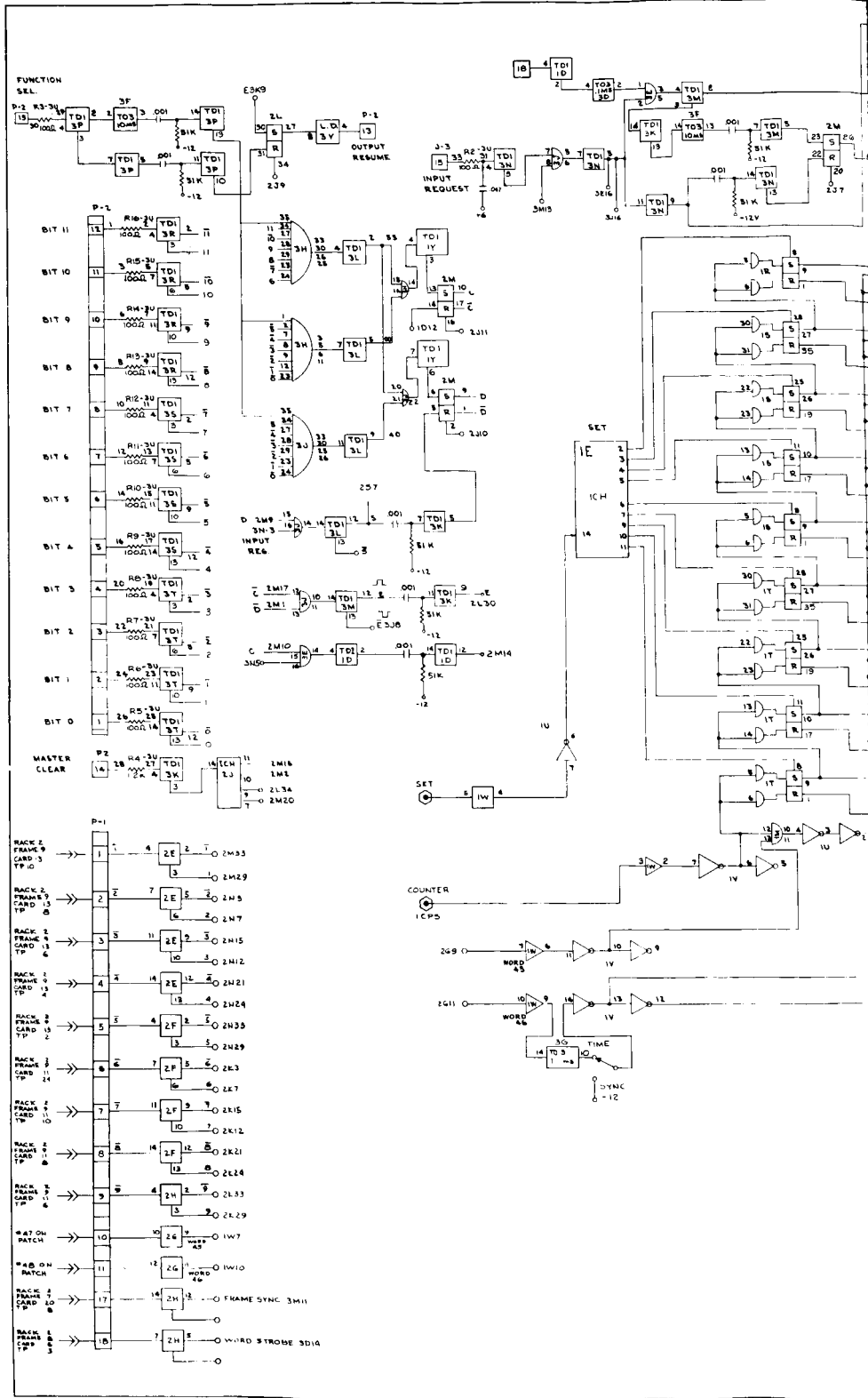


FIGURE B2.—Normal data transfer timing chart.

Then either the output (status) interrogate pulse or the output (data) interrogate pulse is initiated at input request time. The next operation is the insertion of a 2 μ sec word strobe* pulse and the 9 bit data word from the PCM decom station. If data are chosen by the computer, then the output (data) interrogate pulse turns on the 12

*Word strobe pulse is a pulse generated at the telemetry rate by the PCM decom. This pulse occurs during the ninth data bit and insures that the entire data word is transferred from the PCM decom to the computer at the correct time.



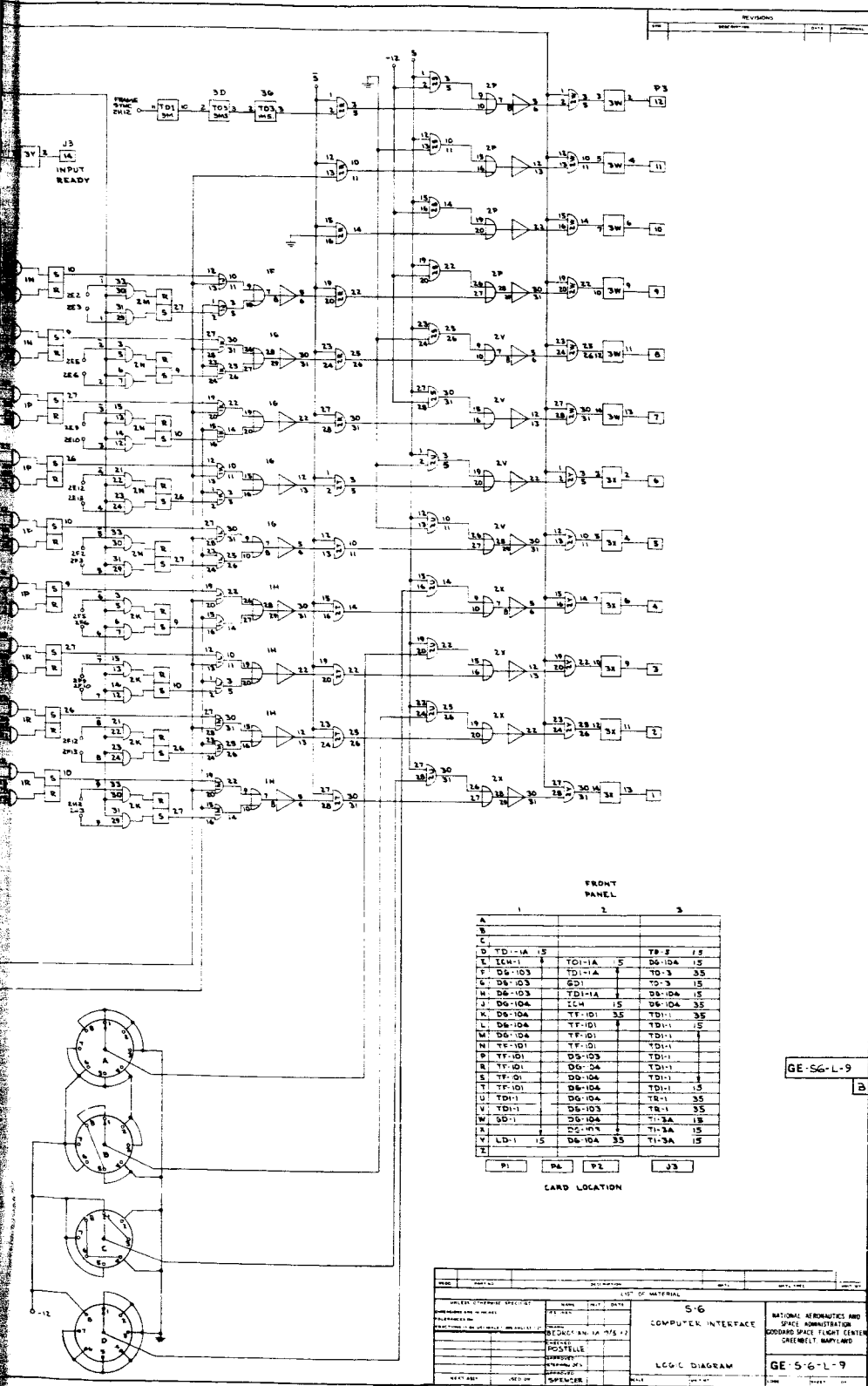


FIGURE B4.—Schematic of Mod II computer interface.

bit data register consisting of 3 flag bits* and the 9 bit data word. At word strobe time a data transfer pulse is developed that transfers the data to the computer via the interfacing amplifiers (line drivers).

The same cycle is followed to return status data. Here the computer requests status and the status coder is interrogated by the output (status) interrogate pulse. The status code pre-selected by the operator is then transferred to the computer upon receipt of the data transfer pulse. The timing for the status operation is shown in Figure B3.

The running-time clock, which consists of a 9 bit binary counter, a 9 bit storage register, and associated gating, provides spacecraft running-time. The binary counter is advanced by a 1 sec pulse obtained from a secondary time standard. A gating arrangement that utilized gating pulses from the PCM decom allows the 9 bit time code to be inserted as a normal data word during one of the sync word times in the format. The time word is then transferred to the computer as a normal data word. The clock is reset at each "on" command sent by the command transmitter or detected automatically during magnetic tape playbacks; then it begins a running-time count of the spacecraft program.

*3 flag bits may be used during data transmissions to indicate any of several functions. In the case covered by this report, 1 bit indicated decom frame sync, 1 bit was used to indicate that the time word was present in the data, and the last bit was not assigned.

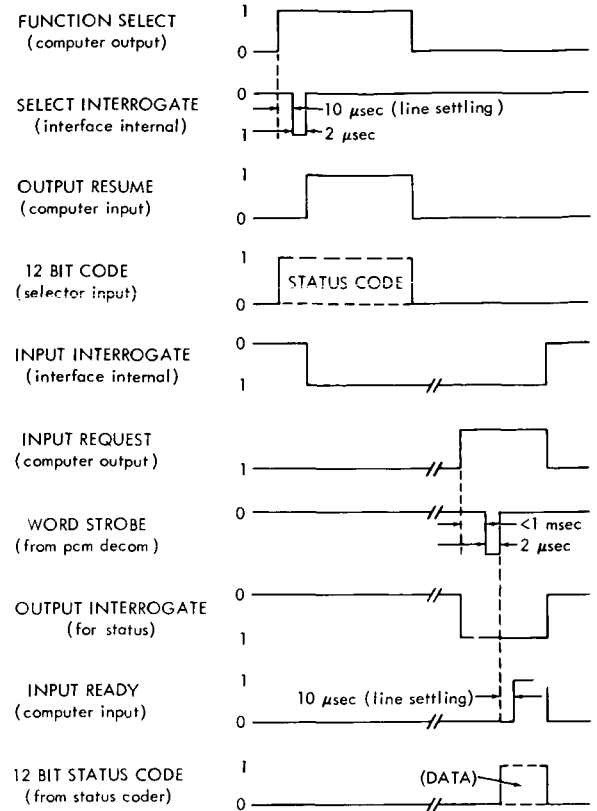


FIGURE B3.—Status data transfer timing chart.

Figure B4 is a schematic drawing of the Mod II interface and gives more information on the gating arrangements for the operations discussed.

APPENDIX D

COMPUTER PROGRAM PRINTOUT FORMATS

The formats as shown in this section are those derived for a CDC 166-2 line printer with a 300 line/min capability. Similar formats were used for a CDC 1612 line printer which printed 1000 lines/min. A dateline (Table D1), which contained the date, time, and pass (or run) number, and command counter and location information, is included for the subcom format only, although this line ordinarily was inserted in any format when selected at initiation of the particular run and program.

SUBCOM FORMAT

Table D1 illustrates the subcommutator (subcom) format. Following the dateline was a level line that identified the particular levels of subcommutation. The subcom main channels (17, 31, and 33) were identified in the left margin. In real time the telemetered information was not received in the sequence of lines as indicated, but took the sequence: channel 17, level 1 (17-1); channel 31, level 1 (31-1); channel 33, level 1 (33-1); channel 17, level 2 (17-2); . . . ; channel 33, level 16 (33-16). At this point, program running-time was computed and formatted at the right margin of the channel 33 printout. Note the 0 and 5 volt references at the beginning or end of the voltage printout for each channel. Unacceptable data, in most cases, were represented by "****" in the appropriate printout location. Temperatures in °C were represented by two-numeral numbers without decimal points, as shown for channel 33, level 3. The staggered format was utilized to facilitate the tracing of a data point from one subcom frame to another.

ELECTRON TEMPERATURE PROBE (ETP) DATA FORMAT

The first block of data in Table D2 is information from ETP II and covers calibration and several live data cycles from that experiment. Data were received, as presented in this format, sequentially from left to right, line by line. A "-1 volt" was indicated by an "*" preceding the decimal point, and unacceptable data by "****".

The second block is again ETP II information and covers calibration and some raw data for the sensitivity (low or high) following that sampled for the first block.

The 4 smaller blocks are ETP I information and take the following sequence: calibration, sensitivity Y; live data, sensitivity Y; calibration, sensitivity Z; and live data, sensitivity Z. Y and Z represent high and low sensitivities but not necessarily in that order. The data sequence for ETP I, which utilized nine main channels, was from left to right, line by line, and, to facilitate plotting of printed values, was formatted in frames, with 2 frames/line.

BAYARD-ALPERT AND REDHEAD PRESSURE GAGE DATA FORMATS

The format shown in Table D3 was that used during the pressure gage calibration period. For each main sample source, 9 successive samples were printed out, 3 samples (1 main frame)/line, 3 frames/block. The sampled sources were identified by name, and the same special indications were used as those for the subcom and ETP formats. The illustrated calibration sequence first presented high calibration for the BA's and RH's, then mid-calibration, and concluded with low calibration. Program running-time for the last sampled frame for each experiment group was printed out in the right margin.

OPTICAL ASPECT FORMAT

The optical aspect system utilized 2 coded words (8 and 32) that relayed information on selected events and counters. This was formatted for printout as shown in Table D4, beginning with word 8 at the extreme left, followed by word 32 to its right, word 8 again, and so on for 4 frames, which completed 1 printout line. The following frame was presented in the next printout line; and, except for periods of prolonged loss of frame synchronization, all optical aspect data were printed out sequentially. "No sync" was indicated by "00000000," an illegal code for the aspect system.

Table D2

Electron Temperature Probe Printout Format.

1.07	1.12	1.16	1.24	1.30	1.35	1.37	1.39	1.47	1.53	1.63	1.69	1.75	1.77	1.81	1.87	2.08	2.41
0.73	5.36	5.22	5.10	5.01	4.88	4.78	4.69	4.54	4.35	4.22	4.13	4.03	3.99	3.79	3.67	3.59	3.46
3.24	3.13	3.00	2.79	2.72	2.54	2.46	2.35	2.25	2.11	2.00	1.93	1.79	1.68	1.59	1.45	1.35	1.27
1.02	0.97	0.83	0.75	0.67	0.58	0.51	0.54	0.53	0.53	0.56	0.53	0.51	0.44	0.40	0.31	0.25	0.20
0.20	0.20	0.27	0.32	0.39	0.48	0.58	0.70	0.81	0.89	0.98	1.10	1.18	1.30	1.49	1.62	1.75	1.84
2.05	2.18	2.27	2.40	3.35	5.37	5.21	5.91	5.02	4.92	4.78	4.73	4.56	4.48	4.34	4.25	4.03	3.89
3.09	3.54	3.48	3.39	3.22	3.18	3.03	2.91	2.79	2.70	2.55	2.47	2.34	2.25	2.12	2.02	1.89	1.68
1.68	1.36	1.25	1.17	1.06	0.98	0.86	0.77	0.65	0.54	0.54	0.50	0.54	0.54	0.53	0.50	0.49	0.45
0.30	0.26	0.21	0.20	0.17	0.22	0.25	0.34	0.41	0.50	0.57	0.68	0.78	0.89	1.10	1.00	0.90	0.83
0.93	0.93	0.93	0.93	0.93	0.93	0.92	0.92	0.93	0.93	0.92	0.93	0.93	0.93	0.93	0.93	0.93	0.93
0.63	0.61	0.59	0.60	0.59	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.60
0.58	0.60	0.55	0.56	0.58	0.60	0.55	0.55	0.58	0.54	0.53	0.49	0.47	0.46	0.51	0.51	0.54	0.58
0.59	0.59	0.54	0.57	0.55	0.55	0.46	0.30	0.26	0.71	1.18	1.67	2.17	2.88	3.41	4.08	4.45	5.35
0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
0.65	0.66	0.64	0.61	0.59	0.60	0.58	0.59	0.61	0.62	0.59	0.61	0.60	0.64	0.60	0.62	0.63	0.61
0.59	0.63	0.61	0.61	0.57	0.59	0.61	0.63	0.56	0.60	0.56	0.56	0.54	0.57	0.56	0.51	0.53	0.59
0.60	0.59	0.59	0.59	0.56	0.59	0.79	0.83	0.82	0.43	0.05	0.21	0.10	0.05	0.15	0.22	0.32	0.45
0.68	0.79	0.86	0.89	0.88	0.86	0.88	0.89	0.89	0.93	0.98	1.10	1.21	1.35	1.46	1.59	1.87	2.00
2.25	2.42	2.42	2.82	2.97	3.09	3.25	3.36	3.53	3.94	3.78	3.92	4.07	4.24	4.34	4.45	4.60	4.87
0.91	4.67	4.51	4.40	4.24	4.12	3.97	3.82	3.69	3.42	3.30	3.15	3.02	2.87	2.75	2.60	2.48	2.32
2.06	1.91	1.79	1.68	1.53	1.40	1.26	1.13	1.00	0.88	0.74	0.63	0.49	0.39	0.26	0.17	0.07	0.05
0.29	0.41	0.51	0.64	0.74	0.85	0.88	0.91	0.87	0.90	0.88	0.89	0.89	0.96	1.03	1.13	1.26	1.41
1.79	1.94	2.06	2.21	2.32	2.49	2.60	2.75	2.88	3.03	3.15	3.31	3.43	3.56	3.70	3.86	3.97	4.11
4.51	4.68	4.78	4.93	4.69	4.52	4.45	4.34	4.17	4.06	3.91	3.78	3.63	3.49	3.35	3.24	3.08	2.81
2.54	2.41	2.26	2.15	2.00	1.88	1.39	0.97	0.97	0.98	0.99	1.00	0.97	0.97	0.97	0.98	0.97	0.99
0.97	0.98	0.97	0.98	0.98	1.00	0.98	1.01	1.00	1.01	0.99	1.01	0.97	0.98	0.97	1.00	0.98	1.01
1.00	1.03	1.10	1.20	1.24	1.31	1.35	1.45	1.51	1.62	1.67	1.65	1.78	1.78	1.76	1.69	1.65	1.68
1.81	1.87	1.96	2.02	2.01	2.06	2.19	2.33	2.48	2.66	2.87	3.07	3.27	3.48	3.66	3.82	3.99	4.17
0.97	0.98	0.97	0.98	0.97	0.98	0.96	0.97	0.96	0.97	0.97	0.96	0.96	0.97	1.00	0.97	0.98	0.98
0.97	1.00	0.98	1.00	0.97	0.98	0.97	1.00	0.98	1.00	0.98	1.00	0.97	1.00	0.97	0.98	0.97	0.98
0.97	1.00	0.98	1.02	1.02	1.10	1.15	1.24	1.31	1.41	1.45	1.52	1.53	1.55	1.68	1.70	1.63	1.54
1.66	1.71	1.86	1.92	1.89	1.98	2.05	2.08	2.15	2.24	2.07	2.07	0.97	0.98	0.97	0.98	0.97	1.00
1.10	1.12	1.16	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05
0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
5.12	5.41	5.03	5.03	5.03	5.03	5.03	5.03	5.03	5.03	5.03	5.03	5.03	5.03	5.03	5.03	5.03	5.03
4.68	5.07	3.40	3.40	3.40	3.40	3.40	3.40	3.40	3.40	3.40	3.40	3.40	3.40	3.40	3.40	3.40	3.40
1.13	1.16	1.25	1.36	1.46	1.56	1.66	1.70	1.82	1.92	2.03	2.13	2.23	2.33	2.43	2.53	2.63	2.73
1.29	1.36	1.46	1.56	1.70	1.82	1.92	2.03	2.13	2.23	2.33	2.43	2.53	2.63	2.73	2.83	2.93	3.03
1.27	1.32	1.37	1.48	1.51	1.54	1.60	1.64	1.68	1.72	1.76	1.80	1.84	1.88	1.92	1.96	2.00	2.04
1.17	1.17	1.22	1.26	1.27	1.30	1.32	1.37	1.40	1.44	1.48	1.52	1.56	1.60	1.64	1.68	1.72	1.76
0.94	0.94	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
2.43	2.65	2.89	3.16	3.37	3.60	3.87	4.08	4.29	4.53	4.72	4.93	5.14	5.35	5.56	5.77	5.98	6.19
2.17	2.40	2.60	2.91	3.11	3.35	3.62	3.83	4.02	4.30	4.50	4.69	4.89	5.09	5.29	5.49	5.69	5.89
1.91	2.13	1.12	1.15	1.17	1.21	1.24	1.26	1.28	1.30	1.32	1.34	1.36	1.38	1.40	1.42	1.44	1.46
1.19	1.14	1.26	1.27	1.27	1.31	1.32	1.32	1.36	1.40	1.44	1.48	1.52	1.56	1.60	1.64	1.68	1.72
1.13	1.14	1.16	1.17	1.18	1.19	1.22	1.25	0.12	1.32	0.96	1.02	1.00	1.03	1.03	1.03	1.04	1.04
1.08	1.10	1.14	1.18	1.20	1.23	1.24	1.27	1.30	1.31	1.34	1.36	1.37	1.40	1.40	1.40	1.40	1.40
1.01	1.01	1.03	1.05	1.08	1.11	1.16	1.21	1.22	1.30	1.32	1.37	1.41	1.40	1.40	1.40	1.40	1.40

Table D3

Bayard-Alpert and Redhead Printout Format During Calibration.

BA I UC	BA I AC	BA II DC	BA II AC	TIME	
4.62 4.63 4.63	2.44 2.45 2.45	4.24 4.24 4.24	2.50 2.50 2.50		124
4.63 4.64 4.64	2.45 2.44 2.44	4.24 4.25 4.24	2.50 2.50 2.50		
4.63 4.63 4.64	2.45 2.45 2.45	4.24 4.24 4.24	2.51 2.50 2.51		
RH I AC	RH II AC			TIME	
4.73 4.73 4.70	4.83 4.83 4.81				126
4.73 4.73 4.70	4.83 4.83 4.81				
4.70 4.73 4.72	4.81 4.83 4.83				
RH I AC	RH II AC			TIME	
3.47 3.46 3.46	3.48 3.46 3.46				129
3.47 3.46 3.47	3.48 3.46 3.46				
3.47 3.47 3.46	3.48 3.46 3.46				
BA I DC	BA I AC	BA II DC	BA II AC	TIME	
2.42 2.42 2.42	1.42 1.42 1.42	2.18 2.18 2.18	1.77 1.76 1.77		135
2.42 2.41 2.42	1.40 1.40 1.40	2.19 2.19 2.18	1.77 1.76 1.76		
2.41 2.42 2.42	1.40 1.40 1.40	2.18 2.18 2.19	1.77 1.77 1.77		
RH I AC	RH II AC			TIME	
2.01 2.01 2.01	1.96 1.96 1.96				137
2.01 2.01 2.01	1.96 1.96 1.96				
2.01 2.01 2.01	1.95 1.96 1.96				

APPENDIX E

TYPICAL TEST STAND AND CDC 160A COMPUTER PROGRAM

Special storage areas in bank "1":

0101-0200	Program dateline
0201-0300	Subcom level line
0301-0400	BA channel identification line
0401-0500	RH channel identification line
0600-0641	Printer codes for numerics and sign
1001-2000	Printout load locations
2000-2060	Raw data storage
3001-5400	Printout load locations
5401-6200	Reference storage
6600-6677	Thermistor conversion constants

bank "O"

0000	7123	0060	0000	0140	4010	0220	4010	0300	4010	0360	6753
0001	0000	0061	0000	0141	0001	0221	0001	0301	2015	0361	0503
0002	0000	0062	0000	0142	0001	0222	0001	0302	0703	0362	5355
0003	0000	0063	0000	0143	0001	0223	0001	0303	4015	0363	0601
0004	0000	0064	0000	0144	7500	0224	0401	0304	0001	0364	4355
0005	0000	0065	0000	0145	0740	0225	4241	0305	2205	0365	7101
0006	0000	0066	0000	0146	7600	0226	4017	0306	4067	0366	2600
0007	0000	0067	2055	0147	6503	0227	0400	0307	2206	0367	5623
0010	0000	0070	7774	0150	7500	0230	4005	0310	4070	0370	3432
0011	0000	0071	0000	0151	0720	0231	2023	0311	6207	0371	6026
0012	0000	0072	4210	0152	7101	0232	4015	0312	2055	0372	7500
0013	0000	0073	0013	0153	6631	0233	7500	0313	2053	0373	5500
0014	0000	0074	7703	0154	0020	0234	5500	0314	2005	0374	7600
0015	2000	0075	0001	0155	7666	0235	7600	0315	2052	0375	4050
0016	0000	0076	0076	0156	0021	0236	1060	0316	2056	0376	0110
0017	0000	0077	7715	0157	7104	0237	3461	0317	2002	0377	0207
0020	0000	0100	0061	0160	3000	0240	0717	0320	7100	0400	0702
0021	0120	0101	7100	0161	0200	0241	6506	0321	6175	0401	6707
0022	7020	0102	6373	0162	0600	0242	7500	0322	2024	0402	7101
0023	3500	0103	0712	0163	0152	0243	5500	0323	0001	0403	7000
0024	0535	0104	6703	0164	1200	0244	7600	0324	0001	0404	7500
0025	1301	0105	7101	0165	2757	0245	1031	0325	4100	0405	5500
0026	0000	0106	4450	0166	4056	0246	4115	0326	1004	0406	7600
0027	0000	0107	7500	0167	3500	0247	5405	0327	0001	0407	4024
0030	0000	0110	5500	0170	7500	0250	0720	0330	0001	0410	1060
0031	0000	0111	7600	0171	0740	0251	6707	0331	0001	0411	3461
0032	0000	0112	0102	0172	7600	0252	6013	0332	2025	0412	0677
0033	0000	0113	4024	0173	6503	0253	0716	0333	0001	0413	6524
0034	0000	0114	0201	0174	7101	0254	6712	0334	0001	0414	4005
0035	0000	0115	0701	0175	5002	0255	6010	0335	4100	0415	7101
0036	0000	0116	6507	0176	7101	0256	0702	0336	1005	0416	0242
0037	0000	0117	2024	0177	0170	0257	6715	0337	0001	0417	0060
0040	0000	0120	0111	0200	0020	0260	6005	0340	0001	0420	2200
0041	0000	0121	0217	0201	2200	0261	0720	0341	0001	0421	1004
0042	0000	0122	4163	0202	0152	0262	6720	0342	0001	0422	4100
0043	0000	0123	2200	0203	4100	0263	4005	0343	2200	0423	0326
0044	0000	0124	0101	0204	0575	0264	6422	0344	0101	0424	0601
0045	0000	0125	0105	0205	0021	0265	5415	0345	5317	0425	4100
0046	0000	0126	0123	0206	7101	0266	6207	0346	0601	0426	0336
0047	4424	0127	2200	0207	2620	0267	3452	0347	4311	0427	0001
0050	0000	0130	0177	0210	2200	0270	6726	0350	0001	0430	0001
0051	0000	0131	0106	0211	0277	0271	2200	0351	0001	0431	0001
0052	0000	0132	0127	0212	0106	0272	6207	0352	0001	0432	0001
0053	0000	0133	7500	0213	0210	0273	4305	0353	0001	0433	0001
0054	0000	0134	0740	0214	7100	0274	6746	0354	5415	0434	0001
0055	0000	0135	7600	0215	0175	0275	5410	0355	5750	0435	0061
0056	0000	0136	6503	0216	0001	0276	0703	0356	5747	0436	2200
0057	0000	0137	7100	0217	7100	0277	6735	0357	3445	0437	0677

bank "O"

0440	4326	0520	2200	0600	0020	0660	0061	0740	0002	1020	0001
0441	7100	0521	1001	0601	2200	0661	7500	0741	0003	1021	0001
0442	6626	0522	4332	0602	0152	0662	5500	0742	0004	1022	0001
0443	7100	0523	2200	0603	4100	0663	7600	0743	0005	1023	0001
0444	6343	0524	1077	0604	1147	0664	1036	0744	0022	1024	0001
0445	2024	0525	4331	0605	0021	0665	0102	0745	0023	1025	0001
0446	4100	0526	6202	0606	0001	0666	0701	0746	0024	1026	0001
0447	1267	0527	0000	0607	0001	0667	6506	0747	0025	1027	5725
0450	2025	0530	2200	0610	2200	0670	7500	0750	0042	1030	0277
0451	4100	0531	2000	0611	0301	0671	5500	0751	0043	1031	0744
0452	1270	0532	4015	0612	0105	0672	7600	0752	0044	1032	6042
0453	7101	0533	5604	0613	0610	0673	4024	0753	0045	1033	0412
0454	0467	0534	0723	0614	2200	0674	5406	0754	0001	1034	5324
0455	0000	0535	6015	0615	0377	0675	0715	0755	4005	1035	0601
0456	0000	0536	6102	0616	0106	0676	6103	0756	4006	1036	4323
0457	0000	0537	0000	0617	0614	0677	2024	0757	2200	1037	0001
0460	0000	0540	7500	0620	7100	0700	4121	0760	3625	1040	0001
0461	0000	0541	5540	0621	4010	0701	2006	0761	4346	1041	0001
0462	0000	0542	7600	0622	7100	0702	0716	0762	0544	1042	0001
0463	0000	0543	0277	0623	2664	0703	6107	0763	5344	1043	5405
0464	0000	0544	4024	0624	0001	0704	2024	0764	7500	1044	0704
0465	0000	0545	0701	0625	7100	0705	4116	0765	5500	1045	6544
0466	0000	0546	6021	0626	0175	0706	2021	0766	7600	1046	4005
0467	2200	0547	2024	0627	0001	0707	4070	0767	4050	1047	2600
0470	1001	0550	0702	0630	0001	0710	2016	0770	0110	1050	0045
0471	0105	0551	6016	0631	7101	0711	4067	0771	0207	1051	5341
0472	0467	0552	0400	0632	0647	0712	2006	0772	0702	1052	0601
0473	2200	0553	4314	0633	1001	0713	3625	0773	6507	1053	4340
0474	1077	0554	0001	0634	5701	0714	6524	0774	4006	1054	0001
0475	0106	0555	0001	0635	3600	0715	2024	0775	7100	1055	0001
0476	0473	0556	0001	0636	2000	0716	4100	0776	6343	1056	0001
0477	7100	0557	0001	0637	6506	0717	2000	0777	7100	1057	0001
0500	0175	0560	7500	0640	6007	0720	5701	1000	1175	1060	0277
0501	7100	0561	0740	0641	0000	0721	0277	1001	2200	1061	0713
0502	4010	0562	7600	0642	0000	0722	0744	1002	2000	1062	6561
0503	0001	0563	6503	0643	0000	0723	6032	1003	4015	1063	2200
0504	0001	0564	7500	0644	0000	0724	5711	1004	7100	1064	0067
0505	0001	0565	0727	0645	0000	0725	0277	1005	6175	1065	5355
0506	2312	0566	7103	0646	0000	0726	0741	1006	2024	1066	0601
0507	0602	0567	7101	0647	0060	0727	6537	1007	4100	1067	4354
0510	4320	0570	0227	0650	2200	0730	4006	1010	1001	1070	0001
0511	3200	0571	3431	0651	1001	0731	0514	1011	2025	1071	0001
0512	0076	0572	7103	0652	4100	0732	5317	1012	4100	1072	0001
0513	4317	0573	0000	0653	1111	0733	6552	1013	1002	1073	0001
0514	5613	0574	0000	0654	2200	0734	7101	1014	7101	1074	6573
0515	0703	0575	0152	0655	1077	0735	0764	1015	1027	1075	0544
0516	6527	0576	7101	0656	4100	0736	0001	1016	0001	1076	5374
0517	4210	0577	0205	0657	1115	0737	0001	1017	0001	1077	2200

bank "O"

1100	1001	1160	2024	1240	7500	1320	7104	1400	7500	1460	5617
1101	4371	1161	4100	1241	5500	1321	0701	1401	5500	1461	0702
1102	0601	1162	1067	1242	7600	1322	6003	1402	7600	1462	6552
1103	4370	1163	0001	1243	0110	1323	7101	1403	4050	1463	4214
1104	0001	1164	0001	1244	0207	1324	1247	1404	0110	1464	5406
1105	0001	1165	0001	1245	0704	1325	4006	1405	0207	1465	0703
1106	0001	1166	0001	1246	6506	1326	5734	1406	0702	1466	6106
1107	0001	1167	0001	1247	7500	1327	5727	1407	6507	1467	4006
1110	2200	1170	0001	1250	5500	1330	2024	1410	7100	1470	7100
1111	1001	1171	2025	1251	7600	1331	4221	1411	6175	1471	6343
1112	0105	1172	4100	1252	4024	1332	2213	1412	2024	1472	7100
1113	1110	1173	1070	1253	5406	1333	4115	1413	4100	1473	1175
1114	2200	1174	7101	1254	0711	1334	5415	1414	1001	1474	7101
1115	1077	1175	0000	1255	6104	1335	5703	1415	2025	1475	3200
1116	0106	1176	7101	1256	2024	1336	3600	1416	4100	1476	0000
1117	1114	1177	1160	1257	4266	1337	2221	1417	1002	1477	0000
1120	7100	1200	0060	1260	7144	1340	6506	1420	7101	1500	0400
1121	4010	1201	2200	1261	0701	1341	2200	1421	1433	1501	4006
1122	0001	1202	0152	1262	6104	1342	2213	1422	0001	1502	0060
1123	0001	1203	4100	1263	2024	1343	4311	1423	0001	1503	4100
1124	0001	1204	3321	1264	4264	1344	6110	1424	0001	1504	1650
1125	7100	1205	0061	1265	7137	1345	0000	1425	0001	1505	0061
1126	0175	1206	2200	1266	0703	1346	0000	1426	0001	1506	2200
1127	2200	1207	0401	1267	6105	1347	0000	1427	0001	1507	2000
1130	0100	1210	0105	1270	2024	1350	0000	1430	0001	1510	4015
1131	5320	1211	1206	1271	4100	1351	0000	1431	0001	1511	2016
1132	2200	1212	2200	1272	6300	1352	0000	1432	0001	1512	4067
1133	0100	1213	0477	1273	7131	1353	0000	1433	5415	1513	2021
1134	5317	1214	0106	1274	0701	1354	5701	1434	0403	1514	4070
1135	3600	1215	1212	1275	6105	1355	0703	1435	5321	1515	7500
1136	1377	1216	7100	1276	2024	1356	6003	1436	0601	1516	5500
1137	6527	1217	0175	1277	4100	1357	7101	1437	4320	1517	7600
1140	7101	1220	7100	1300	6000	1360	1240	1440	0001	1520	4047
1141	3340	1221	4010	1301	7123	1361	4306	1441	0001	1521	0110
1142	0701	1222	0060	1302	0713	1362	0503	1442	0001	1522	0207
1143	6007	1223	2200	1303	6104	1363	5371	1443	0001	1523	0704
1144	7500	1224	6300	1304	2024	1364	0503	1444	5632	1524	6507
1145	0602	1225	4245	1305	4241	1365	5365	1445	0703	1525	2047
1146	7101	1226	2200	1306	7116	1366	2200	1446	6536	1526	0217
1147	0000	1227	6000	1307	0701	1367	2000	1447	4227	1527	0705
1150	7101	1230	4250	1310	6104	1370	4015	1450	0402	1530	6513
1151	0610	1231	0061	1311	2024	1371	2100	1451	5335	1531	6012
1152	7101	1232	0400	1312	4237	1372	6000	1452	0601	1532	7500
1153	0647	1233	4006	1313	7111	1373	4167	1453	4334	1533	5500
1154	0000	1234	7100	1314	0717	1374	2100	1454	0001	1534	7600
1155	0000	1235	2664	1315	6104	1375	6300	1455	0001	1535	4047
1156	0000	1236	2023	1316	2024	1376	4170	1456	0001	1536	0110
1157	0000	1237	4015	1317	4230	1377	0001	1457	0001	1537	0207

bank "O"

1540	0704	1620	6300	1700	1001	1760	0720	2040	0061	2120	4337
1541	6002	1621	2200	1701	4213	1761	6562	2041	7500	2121	2200
1542	6136	1622	0000	1702	0001	1762	6003	2042	5500	2122	3604
1543	7500	1623	4115	1703	0001	1763	0000	2043	7600	2123	4345
1544	5500	1624	5415	1704	0601	1764	0000	2044	4047	2124	2200
1545	7600	1625	5710	1705	4220	1765	4302	2045	0110	2125	5401
1546	4024	1626	5706	1706	0001	1766	5702	2046	0207	2126	4314
1547	5406	1627	0001	1707	0001	1767	0702	2047	0704	2127	7100
1550	0715	1630	0001	1710	2126	1770	6003	2050	6507	2130	2460
1551	6104	1631	0001	1711	0001	1771	7101	2051	7500	2131	7500
1552	2024	1632	0001	1712	0001	1772	2026	2052	5500	2132	5500
1553	4170	1633	0001	1713	4100	1773	4307	2053	7600	2133	7600
1554	7123	1634	0001	1714	1001	1774	7101	2054	4024	2134	4047
1555	0701	1635	7100	1715	0001	1775	0000	2055	5604	2135	0110
1556	6104	1636	2460	1716	0001	1776	7101	2056	3604	2136	0207
1557	2024	1637	5604	1717	0001	1777	1500	2057	6114	2137	0704
1560	4167	1640	0712	1720	5426	2000	0400	2060	6030	2140	6577
1561	7116	1641	6103	1721	2126	2001	4006	2061	0000	2141	2047
1562	0701	1642	6004	1722	0001	2002	2170	2062	0001	2142	0220
1563	6104	1643	0000	1723	0001	2003	0060	2063	0006	2143	6014
1564	2024	1644	7101	1724	4100	2004	4100	2064	0013	2144	7101
1565	4223	1645	1532	1725	1002	2005	1610	2065	0021	2145	2051
1566	7111	1646	4303	1726	0001	2006	4100	2066	0026	2146	7500
1567	0721	1647	2200	1727	0001	2007	1622	2067	0033	2147	5500
1570	6104	1650	0000	1730	0001	2010	0061	2070	0041	2150	7600
1571	2024	1651	6112	1731	0001	2011	7101	2071	0046	2151	0110
1572	4261	1652	2200	1732	5426	2012	1607	2072	0053	2152	0207
1573	7104	1653	0000	1733	0403	2013	5612	2073	2312	2153	0704
1574	0707	1654	3600	1734	5320	2014	0724	2074	0715	2154	6003
1575	6007	1655	0160	1735	0601	2015	6003	2075	6105	2155	7101
1576	7101	1656	6203	1736	4311	2016	7101	2076	2024	2156	2242
1577	1543	1657	7101	1737	0001	2017	1710	2077	4100	2157	7500
1600	7101	1660	2566	1740	0001	2020	4205	2100	5520	2160	5500
1601	2000	1661	0401	1741	0001	2021	7100	2101	6114	2161	7600
1602	0001	1662	4312	1742	0001	2022	0175	2102	0701	2162	4024
1603	0001	1663	5605	1743	7101	2023	7101	2103	6112	2163	5604
1604	4006	1664	0720	1744	2013	2024	1745	2104	2024	2164	3604
1605	2024	1665	6004	1745	2200	2025	0000	2105	4100	2165	6114
1606	4214	1666	7101	1746	1001	2026	7100	2106	5530	2166	6030
1607	2200	1667	1532	1747	0105	2027	0175	2107	6106	2167	0000
1610	0000	1670	0000	1750	1745	2030	7500	2110	2024	2170	0001
1611	4115	1671	4301	1751	2200	2031	0721	2111	4100	2171	0006
1612	5415	1672	7100	1752	1077	2032	7101	2112	5401	2172	0013
1613	2167	1673	2664	1753	0106	2033	1500	2113	5701	2173	0021
1614	4100	1674	7101	1754	1751	2034	0000	2114	5736	2174	0026
1615	6000	1675	2500	1755	7100	2035	0000	2115	2334	2175	0033
1616	2170	1676	0001	1756	4010	2036	0000	2116	0753	2176	0041
1617	4100	1677	2200	1757	5604	2037	0000	2117	6546	2177	0046

bank "O"

2200 0053	2260 5752	2340 0001	2420 4100	2500 2200	2560 4026
2201 2312	2261 5745	2341 0001	2421 2307	2501 2500	2561 0400
2202 0715	2262 6530	2342 0403	2422 0610	2502 4026	2562 4324
2203 6105	2263 2200	2343 5316	2423 4100	2503 2200	2563 7101
2204 2024	2264 5412	2344 0601	2424 2312	2504 2000	2564 1677
2205 4100	2265 4345	2345 4315	2425 7103	2505 4015	2565 7600
2206 5521	2266 2200	2346 0001	2426 0000	2506 2100	2566 0060
2207 6114	2267 5521	2347 0001	2427 0000	2507 6000	2567 2324
2210 0701	2270 4362	2350 0001	2430 2463	2510 4167	2570 4100
2211 6112	2271 0610	2351 0001	2431 5610	2511 2100	2571 1615
2212 2024	2272 4356	2352 5735	2432 0707	2512 6300	2572 2321
2213 4100	2273 2200	2353 5653	2433 6104	2513 4170	2573 4100
2214 5531	2274 1001	2354 0711	2434 4205	2514 7100	2574 1620
2215 6106	2275 4230	2355 6541	2435 7101	2515 6175	2575 0061
2216 2024	2276 0601	2356 4250	2436 2263	2516 2024	2576 7101
2217 4100	2277 4231	2357 0402	2437 7101	2517 4126	2577 1505
2220 5412	2300 0001	2360 5333	2440 2236	2520 5426	2600 2451
2221 5701	2301 0001	2361 0601	2441 0000	2521 2025	2601 0060
2222 5736	2302 0001	2362 4332	2442 7500	2522 4126	2602 5100
2223 2334	2303 0001	2363 0001	2443 5500	2523 5426	2603 0326
2224 0753	2304 7100	2364 0001	2444 7600	2524 5415	2604 0601
2225 6546	2305 2664	2365 0001	2445 0110	2525 3724	2605 4100
2226 4337	2306 2100	2366 0001	2446 0207	2526 6011	2606 0336
2227 2200	2307 5520	2367 5760	2447 0702	2527 5607	2607 0061
2230 3604	2310 4170	2370 5756	2450 6506	2530 0702	2610 7101
2231 4345	2311 2100	2371 5636	2451 7500	2531 6515	2611 0367
2232 7100	2312 5530	2372 0702	2452 5500	2532 4204	2612 0001
2233 2460	2313 4167	2373 6565	2453 7600	2533 5724	2613 0001
2234 7101	2314 2100	2374 4233	2454 7500	2534 5722	2614 0001
2235 2431	2315 5401	2375 0406	2455 5500	2535 6527	2615 0001
2236 5730	2316 4115	2376 5351	2456 7600	2536 0000	2616 0001
2237 5723	2317 7100	2377 0601	2457 7101	2537 2200	2617 0001
2240 6572	2320 6175	2400 4350	2460 1637	2540 2000	2620 7100
2241 0000	2321 2024	2401 0001	2461 7101	2541 4015	2621 2664
2242 2322	2322 0001	2402 0001	2462 2442	2542 2200	2622 2200
2243 4204	2323 0001	2403 0001	2463 0061	2543 6000	2623 0107
2244 2100	2324 4100	2404 0001	2464 7101	2544 4335	2624 4100
2245 5520	2325 1001	2405 3600	2465 2702	2545 0060	2625 1001
2246 4100	2326 2025	2406 1402	2466 0000	2546 4100	2626 2200
2247 0000	2327 4100	2407 6003	2467 0117	2547 1615	2627 0301
2250 5606	2330 1002	2410 7101	2470 0000	2550 2200	2630 4100
2251 0711	2331 7101	2411 2306	2471 0000	2551 6300	2631 1101
2252 6005	2332 2342	2412 0060	2472 0000	2552 4100	2632 2200
2253 5704	2333 0001	2413 2200	2473 0000	2553 1620	2633 0303
2254 4334	2334 0001	2414 5401	2474 0000	2554 0061	2634 4100
2255 6511	2335 0001	2415 4100	2475 0000	2555 4343	2635 1201
2256 0000	2336 0001	2416 2315	2476 0000	2556 2200	2636 2200
2257 4301	2337 0001	2417 3250	2477 0000	2557 2500	2637 0201

bank "O"

2640	0105	2720	0602	3000	0400	3060	0736	3140	7102	3220	0175
2641	2640	2721	4313	3001	4217	3061	6506	3141	4355	3221	2200
2642	7101	2722	3200	3002	4230	3062	4203	3142	7101	3222	1001
2643	0210	2723	0076	3003	2163	3063	2024	3143	3760	3223	0105
2644	0001	2724	4312	3004	0704	3064	7102	3144	7100	3224	3221
2645	0001	2725	3600	3005	6204	3065	0000	3145	3175	3225	2200
2646	0001	2726	1477	3006	0401	3066	3071	3146	6203	3226	1077
2647	1000	2727	6524	3007	4211	3067	3107	3147	7100	3227	0106
2650	0061	2730	2200	3010	6105	3070	3123	3150	0575	3230	3225
2651	0400	2731	1001	3011	0705	3071	7100	3151	7101	3231	7100
2652	4100	2732	4324	3012	6103	3072	6175	3152	4056	3232	4010
2653	1000	2733	2200	3013	0401	3073	2024	3153	7100	3233	0060
2654	5701	2734	1077	3014	4216	3074	0111	3154	0575	3234	5500
2655	3600	2735	4323	3015	7100	3075	0277	3155	7100	3235	1372
2656	2000	2736	7100	3016	0575	3076	0702	3156	1147	3236	0601
2657	6506	2737	0175	3017	2200	3077	6540	3157	7100	3237	5500
2660	2311	2740	7500	3020	0000	3100	5714	3160	3321	3240	1375
2661	4306	2741	0721	3021	6106	3101	7101	3161	7100	3241	0061
2662	6105	2742	5614	3022	7100	3102	3740	3162	3175	3242	5624
2663	7101	2743	0704	3023	2775	3103	0001	3163	7100	3243	0703
2664	1674	2744	6544	3024	7101	3104	0001	3164	0575	3244	6003
2665	7101	2745	4211	3025	3031	3105	7101	3165	7101	3245	7101
2666	2650	2746	7100	3026	0001	3106	3037	3166	3500	3246	1371
2667	7101	2747	0175	3027	7100	3107	7100	3167	0000	3247	4217
2670	4740	2750	7500	3030	0575	3110	6175	3170	7100	3250	0060
2671	0000	2751	0721	3031	2200	3111	2024	3171	0175	3251	2200
2672	0000	2752	7101	3032	0000	3112	0111	3172	7500	3252	6000
2673	0000	2753	0000	3033	6004	3113	0277	3173	0727	3253	4100
2674	0000	2754	7101	3034	7101	3114	0701	3174	7101	3254	1372
2675	0000	2755	2700	3035	3147	3115	6556	3175	0000	3255	2200
2676	0000	2756	0000	3036	0001	3116	5732	3176	7101	3256	6300
2677	0000	2757	2200	3037	7500	3117	7101	3177	3170	3257	4100
2700	7101	2760	0152	3040	5500	3120	3750	3200	0060	3260	1375
2701	4014	2761	4214	3041	7600	3121	7101	3201	2200	3261	2200
2702	0001	2762	7100	3042	4047	3122	3037	3202	1001	3262	2000
2703	7100	2763	1775	3043	0110	3123	7100	3203	4100	3263	0061
2704	0175	2764	7100	3044	0207	3124	6175	3204	1414	3264	4015
2705	2200	2765	3475	3045	0704	3125	2024	3205	0601	3265	2200
2706	1001	2766	7100	3046	6507	3126	0111	3206	4100	3266	0000
2707	0105	2767	2753	3047	2047	3127	0277	3207	1417	3267	4006
2710	2705	2770	7100	3050	0217	3130	0712	3210	0001	3270	7500
2711	2200	2771	0175	3051	0710	3131	6572	3211	0001	3271	5540
2712	1077	2772	7500	3052	6513	3132	2025	3212	0001	3272	7600
2713	0106	2773	0727	3053	7500	3133	0111	3213	0001	3273	4024
2714	2711	2774	7101	3054	5500	3134	0277	3214	0001	3274	0210
2715	7100	2775	0000	3055	7600	3135	0710	3215	0001	3275	6002
2716	4010	2776	7101	3056	4115	3136	6677	3216	0061	3276	7130
2717	2305	2777	2762	3057	5606	3137	2200	3217	7100	3277	2024

bank "O"

3300	0207	3360	7101	3440	2200	3520	3024	3600	0012	3660	2024
3301	0701	3361	1146	3441	4241	3521	4201	3601	0000	3661	0207
3302	6106	3362	0702	3442	4100	3522	2272	3602	0007	3662	0706
3303	7100	3363	6103	3443	0225	3523	4001	3603	0010	3663	6513
3304	0175	3364	7101	3444	2200	3524	5701	3604	0004	3664	7101
3305	7500	3365	0647	3445	2650	3525	3600	3605	0000	3665	0100
3306	0721	3366	7100	3446	4100	3526	4005	3606	0003	3666	0000
3307	7112	3367	0175	3447	2666	3527	6515	3607	0011	3667	0000
3310	0704	3370	7500	3450	0061	3530	0504	3610	0005	3670	0000
3311	6103	3371	0727	3451	7113	3531	5306	3611	0000	3671	0000
3312	7101	3372	7101	3452	0060	3532	0001	3612	0002	3672	0000
3313	1223	3373	1146	3453	0401	3533	7500	3613	0000	3673	0000
3314	7100	3374	0000	3454	4100	3534	4220	3614	0006	3674	0000
3315	0175	3375	0000	3455	0225	3535	7600	3615	0000	3675	0000
3316	7500	3376	0000	3456	2200	3536	0751	3616	0001	3676	0000
3317	0727	3377	0000	3457	2663	3537	6504	3617	0000	3677	0000
3320	7101	3400	2013	3460	4100	3540	7500	3620	0000	3700	7500
3321	0000	3401	0277	3461	2666	3541	4220	3621	0000	3701	5500
3322	7101	3402	0712	3462	0061	3542	7600	3622	0000	3702	7600
3323	1206	3403	6102	3463	7101	3543	4024	3623	0000	3703	4226
3324	0000	3404	5412	3464	0572	3544	0422	3624	0000	3704	0110
3325	0000	3405	2012	3465	0000	3545	0111	3625	0000	3705	0207
3326	3500	3406	0277	3466	0000	3546	3024	3626	0012	3706	0702
3327	0000	3407	0712	3467	0000	3547	4201	3627	0000	3707	6507
3330	0000	3410	6314	3470	7100	3550	2254	3630	0007	3710	7500
3331	0000	3411	6004	3471	0175	3551	4005	3631	0000	3711	5500
3332	0000	3412	2012	3472	7500	3552	5701	3632	0004	3712	7600
3333	0000	3413	0712	3473	0721	3553	6107	3633	0000	3713	4206
3334	0000	3414	4012	3474	7101	3554	4011	3634	0003	3714	0110
3335	0000	3415	5411	3475	0000	3555	0010	3635	0000	3715	0207
3336	0000	3416	0277	3476	7101	3556	0000	3636	0005	3716	0702
3337	0000	3417	0712	3477	3470	3557	0000	3637	0000	3717	6507
3340	7500	3420	6304	3500	0061	3560	0000	3640	0002	3720	2200
3341	5540	3421	2011	3501	7500	3561	0011	3641	0000	3721	0000
3342	7600	3422	0712	3502	4240	3562	3706	3642	0006	3722	0702
3343	4024	3423	4011	3503	7600	3563	6523	3643	0000	3723	3606
3344	0210	3424	0400	3504	6503	3564	0504	3644	0001	3724	6714
3345	6003	3425	4004	3505	7500	3565	5314	3645	0000	3725	7101
3346	7101	3426	0473	3506	4220	3566	7111	3646	0000	3726	3724
3347	3500	3427	7101	3507	7600	3567	7500	3647	0000	3727	7101
3350	2024	3430	6152	3510	0704	3570	5540	3650	7500	3730	3700
3351	0207	3431	7500	3511	6504	3571	7600	3651	5500	3731	0000
3352	0701	3432	5540	3512	7500	3572	0210	3652	7600	3732	0000
3353	6107	3433	7600	3513	4220	3573	0710	3653	4024	3733	0000
3354	7100	3434	0277	3514	7600	3574	6405	3654	0110	3734	0000
3355	0175	3435	0702	3515	4024	3575	7101	3655	0207	3735	0000
3356	7500	3436	6014	3516	0422	3576	3650	3656	0702	3736	0000
3357	0721	3437	0060	3517	0111	3577	4700	3657	6506	3737	0000

bank "O"

3740	7100	4020	2200	4100	0061	4160	6105	4240	0502	4320	2200
3741	3726	4021	2047	4101	0400	4161	2024	4241	5306	4321	5601
3742	7100	4022	4100	4102	4100	4162	4211	4242	0602	4322	4214
3743	1147	4023	2141	4103	5675	4163	7101	4243	4305	4323	2200
3744	7100	4024	0061	4104	0412	4164	4151	4244	2024	4324	5677
3745	3321	4025	7111	4105	4100	4165	0730	4245	4100	4325	4215
3746	7101	4026	4211	4106	5676	4166	6515	4246	5601	4326	6107
3747	3037	4027	0060	4107	0400	4167	4005	4247	5701	4327	2200
3750	7100	4030	2200	4110	4100	4170	2024	4250	5611	4330	5701
3751	3726	4031	0400	4111	5775	4171	4207	4251	0705	4331	4205
3752	7100	4032	4100	4112	0412	4172	2200	4252	6517	4332	2200
3753	3321	4033	2141	4113	4100	4173	5220	4253	4206	4333	5777
3754	7101	4034	0061	4114	5776	4174	4207	4254	5706	4334	4206
3755	3037	4035	7101	4115	0001	4175	7101	4255	5605	4335	2200
3756	0000	4036	2040	4116	7500	4176	4202	4256	0702	4336	5701
3757	0000	4037	0000	4117	5540	4177	2200	4257	7101	4337	0105
3760	7100	4040	5615	4120	7600	4200	5220	4260	4427	4340	4337
3761	3726	4041	0703	4121	0210	4201	4202	4261	0000	4341	2200
3762	7100	4042	6104	4122	6504	4202	2200	4262	0000	4342	5777
3763	1147	4043	7500	4123	7500	4203	0522	4263	0000	4343	0106
3764	7100	4044	0720	4124	5500	4204	0102	4264	4302	4344	4343
3765	3321	4045	7107	4125	7600	4205	4302	4265	7101	4345	0001
3766	7101	4046	0701	4126	0110	4206	0201	4266	4040	4346	0001
3767	3144	4047	6104	4127	0207	4207	6102	4267	6007	4347	0001
3770	0000	4050	4205	4130	0704	4210	0412	4270	2007	4350	0001
3771	0000	4051	7101	4131	6506	4211	4207	4271	6003	4351	0001
3772	0000	4052	4277	4132	4005	4212	5701	4272	7101	4352	0001
3773	0000	4053	7101	4133	4007	4213	3600	4273	4123	4353	7500
3774	0000	4054	4270	4134	6015	4214	4220	4274	7101	4354	0710
3775	0000	4055	0000	4135	7500	4215	6513	4275	4410	4355	0120
3776	0000	4056	7500	4136	5500	4216	6013	4276	4313	4356	7300
3777	0000	4057	5540	4137	7600	4217	0000	4277	2200	4357	4356
4000	7500	4060	7600	4140	0110	4220	0000	4300	5701	4360	5605
4001	0710	4061	0060	4141	0207	4221	0000	4301	4333	4361	0702
4002	0120	4062	4100	4142	0704	4222	0000	4302	5611	4362	6126
4003	7300	4063	4370	4143	6006	4223	0000	4303	0702	4363	4202
4004	4000	4064	0061	4144	4007	4224	0000	4304	6111	4364	6002
4005	7500	4065	7101	4145	0400	4225	0000	4305	4206	4365	0000
4006	0720	4066	4100	4146	4225	4226	0000	4306	6021	4366	0001
4007	7101	4067	0000	4147	4231	4227	0000	4307	7101	4367	2200
4010	1757	4070	0000	4150	6022	4230	0000	4310	4116	4370	0000
4011	7101	4071	0000	4151	7500	4231	0511	4311	7101	4371	0207
4012	4000	4072	0000	4152	5500	4232	5321	4312	4410	4372	0701
4013	7300	4073	0000	4153	7600	4233	2314	4313	0000	4373	6103
4014	5623	4074	0000	4154	0110	4234	0111	4314	4301	4374	7101
4015	0702	4075	0000	4155	4024	4235	4024	4315	2200	4375	4760
4016	6010	4076	0000	4156	5405	4236	2316	4316	5601	4376	0706
4017	0060	4077	0000	4157	0707	4237	5024	4317	4351	4377	6103

bank "O"

4400	7101	4460	2047	4540	0000	4620	5610	4700	2001	4760	5612
4401	4116	4461	0217	4541	0000	4621	2024	4701	0111	4761	0750
4402	7100	4462	0712	4542	0000	4622	6605	4702	3002	4762	6003
4403	3175	4463	6513	4543	0000	4623	0612	4703	4100	4763	7101
4404	7101	4464	0400	4544	2200	4624	6102	4704	0126	4764	4116
4405	0142	4465	4005	4545	4206	4625	0412	4705	2003	4765	4205
4406	7101	4466	7500	4546	4330	4626	4240	4706	0111	4766	7100
4407	4100	4467	5500	4547	2323	4627	2200	4707	3004	4767	3175
4410	7101	4470	7600	4550	0702	4630	0000	4710	4100	4770	7101
4411	4672	4471	4024	4551	6004	4631	6102	4711	0127	4771	3153
4412	0001	4472	5405	4552	0501	4632	0412	4712	2005	4772	0000
4413	0110	4473	0736	4553	5304	4633	4230	4713	0111	4773	0000
4414	0207	4474	6506	4554	6505	4634	0400	4714	3006	4774	0000
4415	0702	4475	4005	4555	4025	4635	4305	4715	4100	4775	0000
4416	6506	4476	2024	4556	4005	4636	7101	4716	0134	4776	0000
4417	7500	4477	3600	4557	2310	4637	4662	4717	2007	4777	0000
4420	5500	4500	0260	4560	0604	4640	7101	4720	0111	5000	0000
4421	7600	4501	6204	4561	4201	4641	0107	4721	3010	5001	0000
4422	7500	4502	0400	4562	2332	4642	0060	4722	4100	5002	0000
4423	5500	4503	4024	4563	5025	4643	0400	4723	0135	5003	0000
4424	7600	4504	6011	4564	0102	4644	4100	4724	0400	5004	0000
4425	7101	4505	3600	4565	4025	4645	4524	4725	4001	5005	0000
4426	4135	4506	0270	4566	5405	4646	5701	4726	4002	5006	0000
4427	4220	4507	6204	4567	0707	4647	3600	4727	4003	5007	0000
4430	0060	4510	0401	4570	6003	4650	4544	4730	4004	5010	0000
4431	2200	4511	4024	4571	5707	4651	6506	4731	4005	5011	0000
4432	2314	4512	6103	4572	6510	4652	2200	4732	4006	5012	0000
4433	4100	4513	0402	4573	4005	4653	4524	4733	4007	5013	0000
4434	4233	4514	4024	4574	2200	4654	4307	4734	4010	5014	0000
4435	0602	4515	2024	4575	2323	4655	0061	4735	7101	5015	0000
4436	4100	4516	4206	4576	4327	4656	2025	4736	3567	5016	0000
4437	4236	4517	5701	4577	6143	4657	0114	4737	0000	5017	0000
4440	0061	4520	3600	4600	3455	4660	7101	4740	0412	5020	0000
4441	2206	4521	4226	4601	6306	4661	4600	4741	4100	5021	0000
4442	6103	4522	6022	4602	4024	4662	2200	4742	1076	5022	0000
4443	7101	4523	6553	4603	0401	4663	0000	4743	2200	5023	0000
4444	4264	4524	0000	4604	4100	4664	0111	4744	0100	5024	0000
4445	7101	4525	0000	4605	0111	4665	3200	4745	5303	5025	0000
4446	4177	4526	0000	4606	6106	4666	0000	4746	3600	5026	0000
4447	0000	4527	0000	4607	3055	4667	4100	4747	2076	5027	0000
4450	7500	4530	0000	4610	4024	4670	0112	4750	6510	5030	0000
4451	5500	4531	0000	4611	0412	4671	6531	4751	2200	5031	0000
4452	7600	4532	0000	4612	4100	4672	7500	4752	1076	5032	0000
4453	4047	4533	0000	4613	0111	4673	5500	4753	4311	5033	0000
4454	0110	4534	0000	4614	2024	4674	7600	4754	7101	5034	0000
4455	0207	4535	0000	4615	0712	4675	4050	4755	2663	5035	0000
4456	0704	4536	0000	4616	6305	4676	7101	4756	0000	5036	0000
4457	6507	4537	0000	4617	4024	4677	4413	4757	0000	5037	0000

bank "O"

6000	2167	6060	3433	6140	6606	6220	0000	6300	2037	6360	7500
6001	3570	6061	6306	6141	0612	6221	0000	6301	4011	6361	5500
6002	4033	6062	4024	6142	5015	6222	0000	6302	4012	6362	7600
6003	2115	6063	5401	6143	5701	6223	0000	6303	4013	6363	4050
6004	3570	6064	2024	6144	5707	6224	0000	6304	2050	6364	0110
6005	4024	6065	6013	6145	5404	6225	0000	6305	3455	6365	0207
6006	6031	6066	6515	6146	0703	6226	0000	6306	4024	6366	0702
6007	6206	6067	3033	6147	6725	6227	0000	6307	6305	6367	6507
6010	1422	6070	4024	6150	7101	6230	0000	6310	5411	6370	2050
6011	4024	6071	5706	6151	3400	6231	4000	6311	2024	6371	1031
6012	2041	6072	0207	6152	4024	6232	0000	6312	6014	6372	7101
6013	4011	6073	0704	6153	2111	6233	0000	6313	6606	6373	0103
6014	6226	6074	6003	6154	0111	6234	0000	6314	3055	6374	7101
6015	3433	6075	2024	6155	5024	6235	0000	6315	0712	6375	6360
6016	6305	6076	6626	6156	2112	6236	0000	6316	4024	6376	0000
6017	4024	6077	0503	6157	0111	6237	0000	6317	6305	6377	0000
6020	2034	6100	5315	6160	5113	6240	0000	6320	5412	6400	2200
6021	4011	6101	2035	6161	4025	6241	0000	6321	2024	6401	6600
6022	6220	6102	6012	6162	0400	6242	0000	6322	6004	6402	4210
6023	3033	6103	0411	6163	4001	6243	0000	6323	6606	6403	0620
6024	3440	6104	3401	6164	4002	6244	0000	6324	0612	6404	4211
6025	6306	6105	4001	6165	4003	6245	0000	6325	5013	6405	0620
6026	2167	6106	0411	6166	4035	6246	0000	6326	2111	6406	4056
6027	3115	6107	3402	6167	2013	6247	0000	6327	0001	6407	0620
6030	4024	6110	4002	6170	3437	6250	0000	6330	4024	6410	4057
6031	5435	6111	0412	6171	4113	6251	0000	6331	2112	6411	2100
6032	6205	6112	3403	6172	0412	6252	0000	6332	0111	6412	6620
6033	3040	6113	4003	6173	4137	6253	0000	6333	3113	6413	4053
6034	4024	6114	2043	6174	7101	6254	0000	6334	0001	6414	2100
6035	0001	6115	4207	6175	2516	6255	0000	6335	4025	6415	6640
6036	0001	6116	0503	6176	0001	6256	0000	6336	0400	6416	4054
6037	0406	6117	5216	6177	0001	6257	0000	6337	4011	6417	2154
6040	0111	6120	0503	6200	2167	6260	0000	6340	4012	6420	0277
6041	4011	6121	5221	6201	3570	6261	0000	6341	4013	6421	4024
6042	0406	6122	0400	6202	3600	6262	0000	6342	7101	6422	0001
6043	0111	6123	4024	6203	0550	6263	0000	6343	0445	6423	7101
6044	4012	6124	2004	6204	6305	6264	0000	6344	2050	6424	7054
6045	4013	6125	4024	6205	0400	6265	0000	6345	1031	6425	0612
6046	2024	6126	0103	6206	4004	6266	0000	6346	4050	6426	0705
6047	6045	6127	5024	6207	7101	6267	0000	6347	7101	6427	6303
6050	0112	6130	5704	6210	6000	6270	0000	6350	6300	6430	0401
6051	4024	6131	2024	6211	2200	6271	0000	6351	0000	6431	4213
6052	1036	6132	0712	6212	5454	6272	0000	6352	0000	6432	2154
6053	6004	6133	6306	6213	4024	6273	0000	6353	0000	6433	0712
6054	2024	6134	4024	6214	4025	6274	0000	6354	0000	6434	6002
6055	3433	6135	5414	6215	7101	6275	0000	6355	0000	6435	0612
6056	6104	6136	2024	6216	6174	6276	0000	6356	0000	6436	3206
6057	2024	6137	6004	6217	0000	6277	0000	6357	0000	6437	0712

bank "O"

6440	6005	6520	2424	6600	6412	6660	6760	6740	3613	7020	0110
6441	0612	6521	3157	6601	5500	6661	0000	6741	7111	7021	0207
6442	5246	6522	4024	6602	6415	6662	0060	6742	0110	7022	0704
6443	6105	6523	6205	6603	0061	6663	0467	6743	7107	7023	6105
6444	0000	6524	0440	6604	2154	6664	0111	6744	4207	7024	2024
6445	5243	6525	4153	6605	0111	6665	4100	6745	0110	7025	0217
6446	0401	6526	2424	6606	4154	6666	6475	6746	3205	7026	7101
6447	4210	6527	4024	6607	0001	6667	4100	6747	7103	7027	0412
6450	0400	6530	2053	6610	0001	6670	6510	6750	0112	7030	0060
6451	4305	6531	0001	6611	0001	6671	5500	6751	7101	7031	2100
6452	7101	6532	4205	6612	0001	6672	6412	6752	6467	7032	0326
6453	6772	6533	4213	6613	0001	6673	5500	6753	0000	7033	0603
6454	6304	6534	4240	6614	0001	6674	6415	6754	0000	7034	4100
6455	7101	6535	0400	6615	0001	6675	0061	6755	0400	7035	C326
6456	6662	6536	4100	6616	5456	6676	7101	6756	6443	7036	0601
6457	0000	6537	1124	6617	5457	6677	6616	6757	0000	7037	4100
6460	0605	6540	4154	6620	3600	6700	6255	6760	4301	7040	0336
6461	3302	6541	2024	6621	6700	6701	6213	6761	0210	7041	0001
6462	5213	6542	3455	6622	6003	6702	6213	6762	0710	7042	0001
6463	0400	6543	4024	6623	7101	6703	6214	6763	6005	7043	0001
6464	4305	6544	6306	6624	6411	6704	6217	6764	2305	7044	0001
6465	0060	6545	5500	6625	7101	6705	6220	6765	0207	7045	0001
6466	2200	6546	1124	6626	0000	6706	6223	6766	7101	7046	0001
6467	6476	6547	2024	6627	7101	6707	6227	6767	0154	7047	0061
6470	4100	6550	6011	6630	6400	6710	6232	6770	7101	7050	0403
6471	6752	6551	6607	6631	7500	6711	6233	6771	3500	7051	5015
6472	0061	6552	3055	6632	5540	6712	6236	6772	2153	7052	7101
6473	2156	6553	0712	6633	7600	6713	7137	6773	0111	7053	0367
6474	7101	6554	4024	6634	4217	6714	7136	6774	0207	7054	2153
6475	6700	6555	6307	6635	7500	6715	0102	6775	0705	7055	3621
6476	0112	6556	5554	6636	5540	6716	7134	6776	7101	7056	6016
6477	4024	6557	2024	6637	7600	6717	4234	6777	6454	7057	2024
6500	0060	6560	6605	6640	4214	6720	0102	7000	7500	7060	6014
6501	2200	6561	0412	6641	3612	6721	3232	7001	5500	7061	2154
6502	6511	6562	4154	6642	6511	6722	7130	7002	7600	7062	0111
6503	4100	6563	6213	6643	2211	6723	0103	7003	4207	7063	0277
6504	6752	6564	0605	6644	4207	6724	7126	7004	7500	7064	4154
6505	0061	6565	6302	6645	5610	6725	4226	7005	5500	7065	2024
6506	2156	6566	5554	6646	0707	6726	0103	7006	7600	7066	0712
6507	7101	6567	2154	6647	6512	6727	3224	7007	4204	7067	6003
6510	6700	6570	6407	6650	4205	6730	7122	7010	7101	7070	7101
6511	5024	6571	0712	6651	6005	6731	0102	7011	7014	7071	6425
6512	0467	6572	6104	6652	0001	6732	4221	7012	0370	7072	7101
6513	0111	6573	5500	6653	0000	6733	0102	7013	0532	7073	6432
6514	4317	6574	1124	6654	0000	6734	3217	7014	7500	7074	7101
6515	4305	6575	6514	6655	0000	6735	7115	7015	5500	7075	6662
6516	0400	6576	0060	6656	2302	6736	4215	7016	7600	7076	1273
6517	4154	6577	5500	6657	7101	6737	0110	7017	4024	7077	0000

bank "O"

bank "1"

7100	2015	7160	0000	0000	0000	0060	7017	0140	0000
7101	3452	7161	0000	0001	0000	0061	4000	0141	0000
7102	6322	7162	0000	0002	0000	0062	0740	0142	0000
7103	2045	7163	0000	0003	0000	0063	0132	0143	0000
7104	0060	7164	0000	0004	0000	0064	0162	0144	0000
7105	4100	7165	0000	0005	0000	0065	6016	0145	0000
7106	0357	7166	0000	0006	0000	0066	0157	0146	0000
7107	0601	7167	0000	0007	0000	0067	2053	0147	0000
7110	4100	7170	0000	0010	0000	0070	7774	0150	0000
7111	0362	7171	0000	0011	0000	0071	0000	0151	0000
7112	2200	7172	0000	0012	0000	0072	4210	0152	0000
7113	0677	7173	0000	0013	0000	0073	0013	0153	0000
7114	4100	7174	0000	0014	0000	0074	7703	0154	0000
7115	0440	7175	0000	0015	2000	0075	0001	0155	0000
7116	0061	7176	0000	0016	2053	0076	0076	0156	0000
7117	2200	7177	0000	0017	0000	0077	7715	0157	0000
7120	2000	7200	0000	0020	0000	0100	0000	0160	0000
7121	4015	7201	0000	0021	2052	0101	4124	0161	0000
7122	7101	7202	0000	0022	7777	0102	4500	0162	0000
7123	0262	7203	0000	0023	2000	0103	0000	0163	0000
7124	7101	7204	0000	0024	0000	0104	0006	0164	0000
7125	0410	7205	0000	0025	0000	0105	0300	0165	0000
7126	7500	7206	0000	0026	0000	0106	0000	0166	0000
7127	5500	7207	0000	0027	0000	0107	6323	0167	0000
7130	7600	7210	0000	0030	0000	0110	5113	0170	0000
7131	7500	7211	0000	0031	0777	0111	0000	0171	0000
7132	5500	7212	0000	0032	0717	0112	0000	0172	0000
7133	7600	7213	0000	0033	0000	0113	0046	0173	0000
7134	7101	7214	0000	0034	0605	0114	5162	0174	0000
7135	1725	7215	0000	0035	0000	0115	7123	0175	0053
7136	7101	7216	0000	0036	7000	0116	0046	0176	0061
7137	1725	7217	0000	0037	0600	0117	2565	0177	0000
7140	0000	7220	0000	0040	0631	0120	5100	0200	0000
7141	0000	7221	0000	0041	0640	0121	6240	0201	6370
7142	0000	7222	0000	0042	1775	0122	4723	0202	0000
7143	0000	7223	0000	0043	2001	0123	0047	0203	0000
7144	0000	7224	0000	0044	0700	0124	6122	0204	0000
7145	0000	7225	0000	0045	2211	0125	2214	0205	0100
7146	0000	7226	0000	0046	0304	0126	0000	0206	0000
7147	0000	7227	0000	0047	0000	0127	0000	0207	0000
7150	0000	7230	0000	0050	0000	0130	0000	0210	0200
7151	0000	7231	0000	0051	0300	0131	4745	0211	0000
7152	0000	7232	0000	0052	2060	0132	0000	0212	0000
7153	0000	7233	0000	0053	0000	0133	0000	0213	0300
7154	0000	7234	0000	0054	0000	0134	0000	0214	0000
7155	0000	7235	0000	0055	0144	0135	0000	0215	0000
7156	0000	7236	0000	0056	0000	0136	6422	0216	0400
7157	0000	7237	0000	0057	0000	0137	2300	0217	0000

bank "1"

0220	0000	0300	0000	0360	0000	0440	0000	0520	0000	0600	0012
0221	0500	0301	0000	0361	0000	0441	0000	0521	0000	0601	0001
0222	0000	0302	0000	0362	0000	0442	0000	0522	0000	0602	0002
0223	0000	0303	6261	0363	0000	0443	0000	0523	0000	0603	0003
0224	0600	0304	0071	0364	0000	0444	0000	0524	0000	0604	0004
0225	0000	0305	0000	0365	0000	0445	0000	0525	0000	0605	0005
0226	0000	0306	6463	0366	0000	0446	0000	0526	0000	0606	0006
0227	0700	0307	0000	0367	0000	0447	0000	0527	0000	0607	0007
0230	0000	0310	0000	0370	0000	0450	0000	0530	0000	0610	0010
0231	0000	0311	0000	0371	0000	0451	2371	0531	0000	0611	0011
0232	1000	0312	0000	0372	0000	0452	4465	0532	0000	0612	0012
0233	0000	0313	0000	0373	0000	0453	0000	0533	0000	0613	0013
0234	0000	0314	0000	0374	0000	0454	0000	0534	0000	0614	0000
0235	1100	0315	6261	0375	0053	0455	0000	0535	0000	0615	0000
0236	0000	0316	0071	0376	0061	0456	0000	0536	0000	0616	0000
0237	0001	0317	0000	0377	0000	0457	0000	0537	0000	0617	0000
0240	1200	0320	6163	0400	0000	0460	0000	0540	0000	0620	0000
0241	0000	0321	0000	0401	0000	0461	0000	0541	0000	0621	0000
0242	0001	0322	0000	0402	0000	0462	0000	0542	0000	0622	0000
0243	0100	0323	0000	0403	5170	0463	0000	0543	0000	0623	0000
0244	0000	0324	0000	0404	0071	0464	0000	0544	0000	0624	0000
0245	0001	0325	0000	0405	0000	0465	0000	0545	0000	0625	0000
0246	0200	0326	0000	0406	6163	0466	0000	0546	0000	0626	0000
0247	0000	0327	6261	0407	0000	0467	0000	0547	0000	0627	0000
0250	0001	0330	0071	0410	0000	0470	0000	0550	0000	0630	0000
0251	0300	0331	7100	0411	0000	0471	0000	0551	0000	0631	0000
0252	0000	0332	6463	0412	0000	0472	0000	0552	0000	0632	0000
0253	0001	0333	0000	0413	0000	0473	0000	0553	0000	0633	0000
0254	0400	0334	0000	0414	0000	0474	0000	0554	0000	0634	0000
0255	0000	0335	0000	0415	5170	0475	0053	0555	0000	0635	0000
0256	0001	0336	0000	0416	0071	0476	0061	0556	0000	0636	0000
0257	0500	0337	0000	0417	7100	0477	0020	0557	0000	0637	0000
0260	0000	0340	0000	0420	6163	0500	0000	0560	0000	0640	0040
0261	0001	0341	6261	0421	0000	0501	0000	0561	0000	0641	0054
0262	0600	0342	0071	0422	0000	0502	0000	0562	0000	0642	0000
0263	0000	0343	7100	0423	0000	0503	0000	0563	0000	0643	0000
0264	0000	0344	6163	0424	0000	0504	0000	0564	0000	0644	0000
0265	0000	0345	0000	0425	0000	0505	0000	0565	0000	0645	0000
0266	0000	0346	0000	0426	0000	0506	0000	0566	0000	0646	0000
0267	2371	0347	0000	0427	0000	0507	0000	0567	0000	0647	0000
0270	4465	0350	0000	0430	0000	0510	0000	0570	0000	0650	0000
0271	0000	0351	2371	0431	0000	0511	0000	0571	0000	0651	0000
0272	0000	0352	4465	0432	0000	0512	0000	0572	0000	0652	0000
0273	0000	0353	0000	0433	0000	0513	0000	0573	0000	0653	0000
0274	0000	0354	0000	0434	0000	0514	0000	0574	0000	0654	0000
0275	0053	0355	0000	0435	0000	0515	0000	0575	0000	0655	0000
0276	0061	0356	0000	0436	0000	0516	0000	0576	0000	0656	0000
0277	0000	0357	0000	0437	0000	0517	0000	0577	0000	0657	0000

bank "1"

6600	1113	6660	1603
6601	1116	6661	2274
6602	1121	6662	2032
6603	1124	6663	1607
6604	1214	6664	1673
6605	1217	6665	1601
6606	1236	6666	1142
6607	1241	6667	1130
6610	1244	6670	1200
6611	1247	6671	1306
6612	1260	6672	1147
6613	1255	6673	1142
6614	1023	6674	1306
6615	1026	6675	1356
6616	1050	6676	1236
6617	1053	6677	0741
6620	1114	6700	0000
6621	1117	6701	0000
6622	1122	6702	0000
6623	1125	6703	0000
6624	1215	6704	0000
6625	1220	6705	0000
6626	1237	6706	0000
6627	1242	6707	0000
6630	1245	6710	0000
6631	1250	6711	0000
6632	1261	6712	0000
6633	1256	6713	0000
6634	1024	6714	0000
6635	1027	6715	0000
6636	1051	6716	0000
6637	1054	6717	0000
6640	0024	6720	0000
6641	0043	6721	0000
6642	0026	6722	0000
6643	0024	6723	0000
6644	0024	6724	0000
6645	0025	6725	0000
6646	0022	6726	0000
6647	0021	6727	0000
6650	0024	6730	0000
6651	0027	6731	0000
6652	0023	6732	0000
6653	0023	6733	0000
6654	0026	6734	0000
6655	0034	6735	0000
6656	0021	6736	0000
6657	0014	6737	0000

AN EFFICIENT PCM ERROR CORRECTION AND SYNCHRONIZATION CODE*

MARVIN S. MAXWELL AND RICHARD L. KUTZ

An efficient method of frame synchronization and error correction is analyzed. The probability of frame synchronization as a function of single bit error probability is shown. Additional error correction information may be obtained if the voltage from the bit integrator is used with an error detecting code. The equivalent gain in signal-to-noise ratio, made by using the additional error correction method, is discussed. On the basis of this analysis, a ground station is proposed to process data by using these error correction techniques.

INTRODUCTION

A communication system is proposed herein which uses biphasic modulation to transmit binary information in serial form. When serial binary data transmission is used, identification is necessary so that the data can be regrouped into words or frames at the receiver. A synchronization code pattern is usually used for word or frame identification. With adequate synchronization assured the problem of errors occurring because of noise in the communication channel still remains. One method of reducing errors in the data output is to use an error correcting code.

This study was made to find an improved synchronization and error correction method for a one-way planetary landing probe communication system. A crossed parity check or iterated code (Reference 1) is used here since it lends itself to an intuitive understanding of the frame synchronization and error correction method. With a planetary landing probe, a limited quantity of data can be transmitted in a short time at a very high cost; therefore the data efficiency of the communication system should be as high as possible, and the amount of planetary probe communication electronics should be kept small to improve reliability. Finally, it is desirable to keep any necessary increase in system complexity in the ground station.

The frame synchronization and error correction codes are combined in this study for greater com-

munication efficiency; however increased data processing is required at the receiver. A computer simulation is used to study the modified frame synchronization method and to make a set of curves for frame synchronization versus single-bit error probability. When an error detection code is used, additional error correction information is available at the receiver. The equivalent signal-to-noise ratio gain and the method of using this information is discussed.

FRAME SYNCHRONIZATION

Since measurements accurate to 1 percent of full scale are desired, the data are coded in 7 bit binary words. A horizontal parity bit is added to each 7 bit data word, making the total length 8 bits. A number of 8 bit words are arranged in a matrix, as shown in Figure 1, with an 8 bit vertical parity word added on the bottom to complete the frame. The error correcting code is linear both vertically and horizontally, thus making an iterated code. In the computer simulation, frame sizes varying from 2 to 10 total words are studied.

Frame synchronization can be obtained by using the parity checks as if they were ones and zeros in a synchronization word (Reference 2). Parity checks may be assigned to be either even or odd, as desired, in the frame format. The data enters the receiver frame synchronizer in serial form. While the data shifts through the synchronizer, the data is checked for the formatted horizontal (X) and vertical (Y) parity. A frame

*Published as *NASA Technical Note D-2317*, June 1964.

NOISE VOLTAGES READ FROM INTEGRATOR AT END OF BIT TIME WITH NO DATA BEING TRANSMITTED
(obtained from table of Gaussian deviates with $\mu = 0$ and $\sigma = 1/2.05$)

-0.225v	-0.094v	-0.006v	-0.590v	+1.049v	+0.652v	-0.960v	-0.862v
-0.711v	+0.431v	+0.488v	-0.082v	+0.408v	-0.603v	+0.796v	-0.693v
-0.936v	-0.608v	-0.102v	+0.381v	-0.161v	-1.441v	-0.218v	-0.046v

DATA WITH PROPER PARITY (to be transmitted) THE SYMBOL p SHOWS LOCATION OF PARITY BITS

0	1	0	1	0	0	1	1	p
1	1	0	0	1	1	0	1	p
0	p	1	p	0	p	0	p	1
ODD	ODD	EVEN	ODD	ODD	EVEN	EVEN	EVEN	ODD

SIGNAL PLUS NOISE VOLTAGE AS RECEIVED WITH THE CORRESPONDING PROBABILITY OF EACH BIT BEING IN ERROR

-1.225v 3.6X10 ⁻⁵	+0.906v 4.7X10 ⁻⁴	-1.006v 2.2X10 ⁻⁴	+0.410v 3.0X10 ⁻²	+0.049v 3.9X10 ⁻¹	-0.348v 4.8X10 ⁻²	+0.040v 4.1X10 ⁻¹	+0.138v 2.3X10 ⁻¹
+1.711v 5.6X10 ⁻⁷	+1.431v 5.8X10 ⁻⁶	-0.512v 1.3X10 ⁻²	-1.082v 1.1X10 ⁻⁴	-1.408v 7.2X10 ⁻⁶	+0.397v 3.3X10 ⁻²	-0.294v 1.4X10 ⁻¹	+0.307v 6.6X10 ⁻²
-1.936v 8.6X10 ⁻⁸	+0.392v 3.6X10 ⁻²	-1.102v 9.2X10 ⁻⁵	-0.619v 5.2X10 ⁻³	-1.161v 5.6X10 ⁻⁵	-0.441v 2.3X10 ⁻²	+0.782v 1.3X10 ⁻³	-1.046v 1.5X10 ⁻⁴
			A	B			C

FIGURE 1.—Three word frame with a signal-to-noise ratio of 6.2 db in a bandwidth of one-half the bit rate.

synchronizer for a 3 word frame is shown in Figure 2. The number of X and Y parity disagreements or errors per serial bit shift are recorded. If the parity error count is searched to find low numbers of errors spaced one frame apart, then synchronization can be obtained reliably at single bit error probabilities less than 10⁻².

To determine the synchronization capabilities of this code a computer simulation was performed. A constant bit rate and a random error probability were assumed. The effect of bit synchronization

extending into noise on either or both ends of the data were considered.

The computer program generates up to forty frames of serial bits as if they were the output of a bit synchronizer. The bits are produced by a random number generator. Frames of data differ from frames of noise only in that bits are arranged in the right-hand column and bottom row to satisfy the parity requirements. Both X and Y parity structures must be specified for each test as if they were synchronization words. Data frames followed and preceded by noise frames are produced with the number in each group being a specified integer. Random errors are then added, as if the errors had occurred because of additive white noise in the transmission path, by randomly complementing bits at a specified error rate.

Any combination of four output routines may be called for in the computer simulation. In routine one the sequence of frames is printed out before and after the errors are added; in addition a listing of the error positions is printed out. Routine two analyzes the number of X and Y parity errors and prints a line each time the number of errors in X or Y or their sum is less than a specified threshold. This threshold is used to limit the amount of data output from the computer. Each line of printout gives the bit position measured from the first bit in the series, the number and location of the X and Y parity errors, and the number of bit positions between this match and the preceding fifteen lines of printout. This routine was used to establish the synchronization rule to be used in the next section. Each potential series of frames in routine three is given an identity number and a point count on the output. The series with the highest point is assumed to be correct synchronization series. Figure 3 also shows a "no synchronization" curve which corresponds to two or more series having the same point count. The fourth routine complements each bit before it enters the bit synchronizer and has the same output as routine three. Inversion of the data can only be discovered by using an odd number of words in a frame and checking the normal and inverted Y parity sequences.

COMPUTER SYNCHRONIZATION RULE

The synchronization rule used in routines three and four requires that the starting frame have less

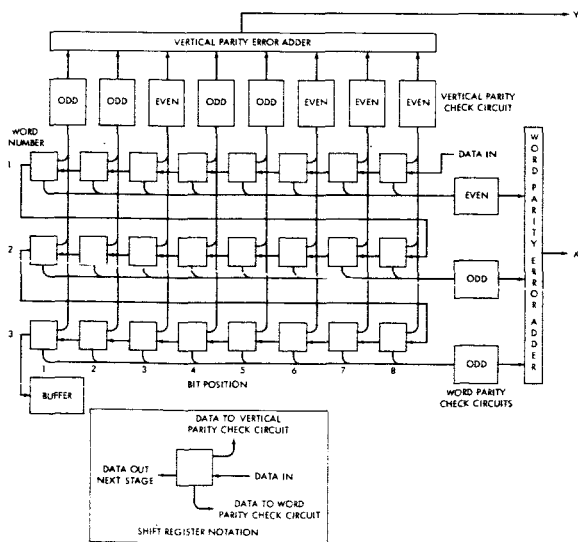


FIGURE 2.—Frame synchronization system.

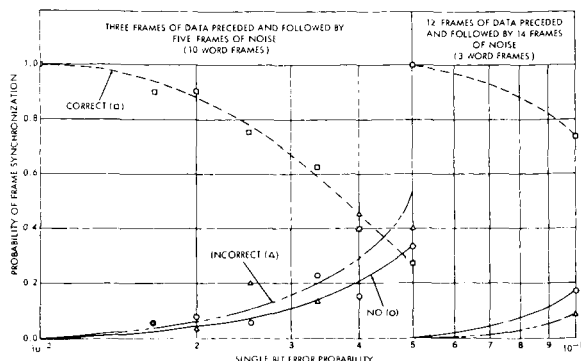


FIGURE 3.—Frame synchronization probability.

than three X and Y errors without having been used in any previous series. If the starting frame is the only one in a series, then it will be deleted from the output. Once a starting frame has been found, the two preceding frames, spaced one and two frame lengths away, as well as all succeeding frames in this synchronization pattern, multiplies of one frame length away, are checked for the number of parity errors. When two succeeding frames having more than two X or Y errors are found, the series is ended. All frames in a series having less than three X and three Y errors are given one of the following point counts:

1. Number of X and Y errors unequal, -1
2. Three X and three Y errors, 1
3. Two X and two Y errors, 2
4. One X and one Y error, 3
5. Zero X and zero Y errors, 4

If a frame having more than three X or Y errors is encountered between two frames in a series, the point count is reduced by 1. The point count is summed for all the frames in a series to give the total point count which appears on the output.

With an error probability of 2×10^{-2} , the computer simulation showed a 90 percent probability of correct synchronization on a group of three 10 word frames preceded and followed by five frames of noise (see Figure 3). When the same amount of data is sent using 3 word frames, the probability of finding the synchronization series at an error rate of 10^{-1} is 75 percent; however, it is more difficult to determine where the data starts and stops with relation to the noise.

Using an odd number of words in a frame proved to be a reliable method of telling whether the data bits had been complemented during transmission.

Since data with a single bit error probability of less than 10^{-3} are expected and frame synchronization is acquired reliably at an error probability of 10^{-2} , it is felt that an adequate frame synchronization margin has been provided. In the section on error correction, perfect frame synchronization is assumed. This assumption is invalid at error probabilities greater than 10^{-2} . A single bit error probability of 10^{-2} is therefore used as an upper bound in the error correction study.

ERROR CORRECTION

Error Correction Advantages

The crossed parity structure used in this study for frame synchronization and error correction has a minimum Hamming distance of 4. With a Hamming distance of 4 an even number of errors having a distance of 4 or greater can go undetected in a frame. Fortunately, a frame having the number of errors necessary to produce an undetectable error condition seldom has the errors lying in an undetectable pattern. Within the error probability range of interest—single bit error probabilities less than 10^{-2} —the binomial probability of increasing numbers of errors in a frame approaches zero with sufficient rapidity to cause the probability of 4 or more errors in a 10 word frame to be less than 10^{-2} . If all the frames with 4 or more errors were falsely corrected, less than 1 percent of the frames would be affected.

Since the iterated code is being used for single error correction, there is the danger of changing some of the odd numbers of errors into an undetectable even number of errors. When error correction is used, the gain in correct data must not be offset by a comparable gain in bad data. Since the tables of binomial probability distribution show that no difficulty will result from using single error correction or even dual and triple error correction, it is safe to use the code under discussion. Figures 4 and 5 show the probability of a frame being in error after single and dual error correction, respectively. The single frame error probability is plotted against signal-to-noise-power ratio in a bandwidth equal to the information bit rate for three different frame sizes. The relationship between bit rate and information bit rate is determined as follows:

The bit rate $BR = N/T$ and the information bit rate $IBR = K/T$, where N is the total number of

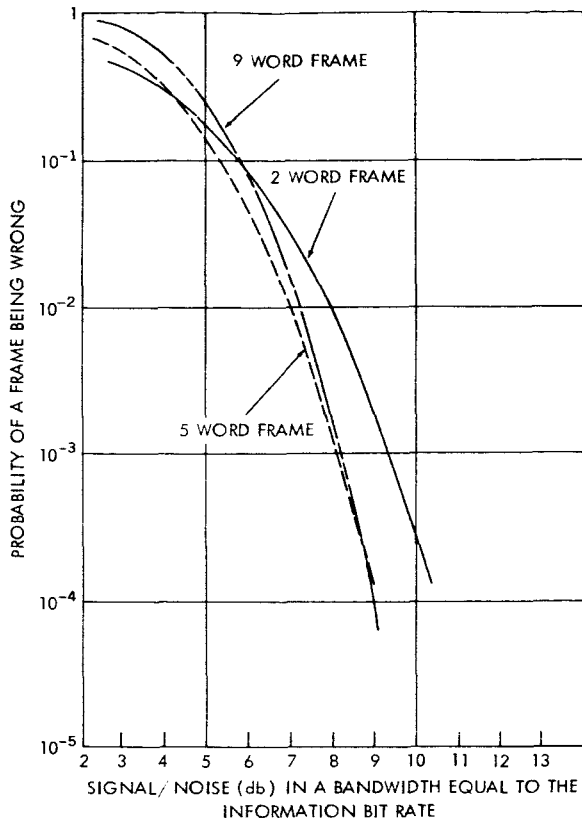


FIGURE 4.—Single frame error probabilities, single error correction. The information rate is held constant.

bits in a frame, T is the total frame transmission time, and K is the number of information bits in a frame. Thus

$$\frac{IBR}{BR} = \frac{K}{N}$$

When frame sizes are compared, a constant information bit rate is used to find the highest probability or correctly receiving data. This is done in order to receive the maximum amount of information per unit time without regard to the code efficiency K/N . The efficiency and consequently the bit rate is varied by changing the frame size. A point on Figure 4 may be determined in the following manner. Since odd values are not found in the binomial probability distribution tables (Reference 3) start with a convenient single bit error probability. If a single bit error probability of 10^{-2} is chosen as the starting value, then a cumulative binomial probability of 0.16229 is found for two or more errors in a 9 word frame which corresponds to single bit error correction.

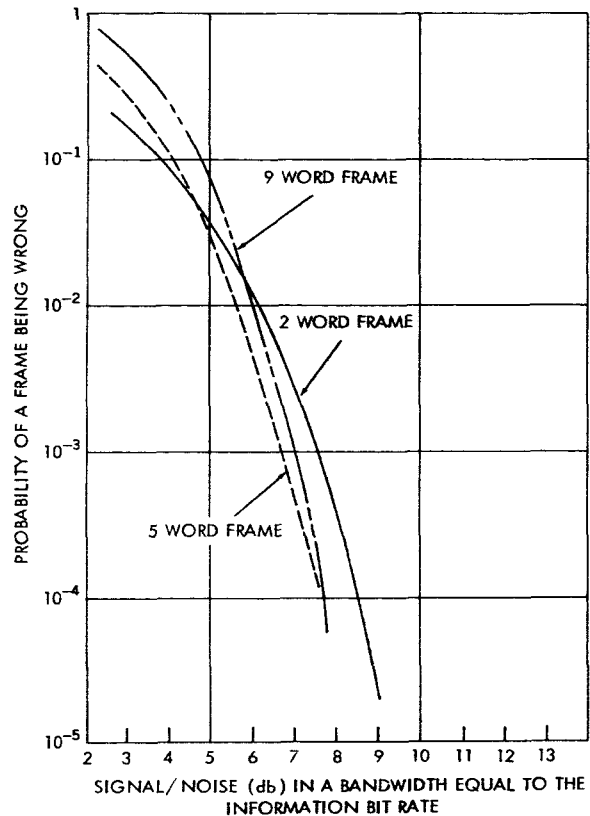


FIGURE 5.—Single frame error probabilities, dual error correction. The information rate is held constant.

The binomial probability value is plotted on the ordinate of Figure 4, but the abscissa value still remains to be found. For a matched filter bit detector a single bit error probability of 10^{-2} corresponds to a 7.35 db signal-to-noise ratio in a bandwidth equal to half the bit rate of a 4.34 db signal-to-noise ratio in a bandwidth equal to the bit rate. Going from the bit rate bandwidth to the information bit rate bandwidth produces a gain in signal-to-noise ratio equal to N/K . For the 9 word frame the N/K ratio is 1.286 which corresponds to a gain of 1.09 db. The single frame error probability of 0.16229 for the 9 word frame corresponds to a 5.43 db signal-to-noise ratio in a bandwidth equal to the information bit rate. In Figure 4 all the frames have been adjusted to have the same information bit rates with the same noise density in the communication channel. This means that the 2 word frame must have a bit rate 1.78 times that of the 9 word frame and, consequently, a higher single bit error proba-

bility. Even though the 2 word frame has the highest single bit error probability associated with it, fewer bits are in the 2 word frame, decreasing the probability of an error occurring in the frame. The 2 word frame also has the highest ratio of bit correction capability to data bits in the frame. These two effects work together to make the single frame probability of an error for the 2 word frame less than the error probabilities of the larger frames at low signal-to-noise ratios. The superior error correction capability of the 2 word frame is not as valuable at high signal-to-noise ratios and, therefore, the single frame error probability curve of the 2 word frame drops below the larger-frame error probability curves. It must be kept in mind that Figures 4 and 5 show single frame error probabilities, and it takes eight 2 word frames or two 5 word frames to equal the same amount of data contained in one 9 word frame.

Determination of Conditioned Bit Error Probability

More correction capability is available in a system using a matched filter bit detector with an error detecting code than can be achieved by using a simple one-zero decision corresponding to a plus or minus voltage polarity from the bit detector. The added information available from the bit detector may be obtained by recording not only the polarity but also the magnitude of the voltage from the integrator at the end of the bit time. The probability of a bit being in error would be expected to decrease as the bit voltage increases. This is true because it is unlikely that a large integrator voltage is due to a bit voltage polarity having been changed, which causes an error, and further increased due to noise alone. The probability of interest, $P(e/|v|)$, is the conditional probability of an error having occurred if the absolute value of the integrator voltage is given. The absolute value of the integrator voltage is used because both polarities of voltage will produce the same result in the derivation. The Gaussian probability densities, due to white noise, of ones $p^1(v)$ and zeros $p^0(v)$ are symmetrically placed about zero volts and have the same standard deviation (Figure 6). Ones and zeros are assumed to have equal probabilities of occurrence in the following discussion. $P(e/|v|)$ is

$$P(e/|v|) = \frac{P(e|v|)}{P(|v|)} \quad (1)$$

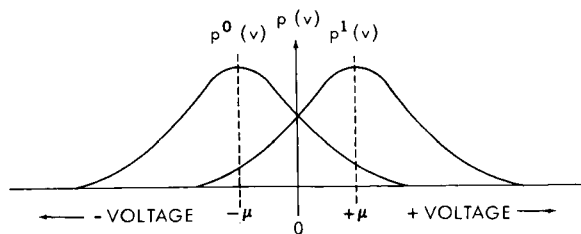


FIGURE 6.—Noise pulse signal probability densities.

where $P(e|v|)$ is the joint probability of having an error and having the voltage $|v|$ present due to the probability density $p^E(v)$ (Figure 7),

$$P(e|v|) = p^E(v) \Delta v \quad (2)$$

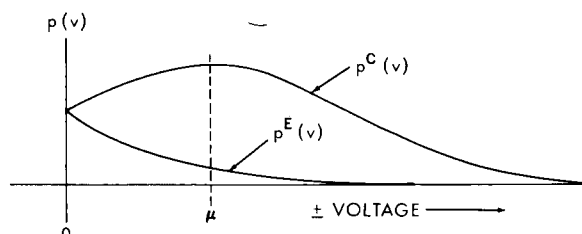


FIGURE 7.—Correct bit and error bit probability densities.

$P(|v|)$ is the probability of finding the voltage $|v|$ due to probability densities $p^E(v)$ and $p^C(v)$ combined,

$$P(|v|) = p^E(v) \Delta v + p^C(v) \Delta v \quad (3)$$

If $|\mu|$, which is both the mean of the noise distribution and the amplitude of the signal level, is found, then all the voltage readings obtained can be normalized and no loss of generality results from using a normalized mean of ± 1 in the Gaussian probability density distributions. The distribution $p^E(v)$ is the tail of the density function and indicates an error condition; $p^C(v)$ is the main body of the density function associated with a correct bit,

$$p^E(v) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(v+1)^2/2\sigma^2} \quad (4)$$

$$p^C(v) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(v-1)^2/2\sigma^2} \quad (5)$$

Substituting Equations 2 and 3 into Equation 1

shows

$$P(e/|v|) = \frac{p^E(v) \Delta v}{p^E(v) \Delta v + p^C(v) \Delta v} \quad (6)$$

Then by substituting Equations 4 and 5 into Equation 6 and simplifying,

$$P(e/|v|) = \frac{e^{-v/\sigma^2}}{e^{-v/\sigma^2} + e^{v/\sigma^2}} = \frac{e^{-v/\sigma^2}}{2 \cosh \frac{v}{\sigma^2}} \quad (7)$$

If $v/\sigma^2 > 4$ this simplifies to:

$$P(e/|v|) \approx e^{-2v/\sigma^2} \quad (8)$$

A plot of $P(e/|v|)$ is shown as a function of bit detector voltages for several different values of signal-to-noise ratio in Figure 8.

Computation of μ and σ

A significant sample of bit voltages must be chosen but not so large a sample that a changing μ and σ will be averaged. The mean μ of the Gaussian density function must be found first. In a binary system the density functions for ones and zeros add together in the receiver and cannot be separated. The Gaussian mean cannot be found directly, but a good approximation results if the absolute value of the integrator voltage is used to compute a mean:

$$\mu' = \frac{2\sigma}{\sqrt{2\pi}} e^{-\mu^2/2\sigma^2} + 2\mu \left(\frac{1}{\sqrt{2\pi}} \int_0^{\mu/\sigma} e^{-y^2/2} dy \right) \quad (9)$$

If μ'/μ is calculated through the range of interest of σ , it is found that $\mu \approx \mu'$. The standard deviation can be calculated by standard methods since it is not affected by the change in the density function if the Gaussian mean is known. With the mean equal to 1, the standard deviation σ is equal to the reciprocal of the signal-to-noise rms voltage ratio (for a bandwidth equal to half the bit rate for a matched filter bit detector).

The Multiple Error Correction Procedure

Multiple bit error correction can be achieved with the iterated code under discussion in conjunction with the conditional bit error probability $P(e/|v|)$. The position and number of parity errors must be inspected to determine all the

possible configurations of errors for each possible number of errors in a frame. For example, two errors may occur in three basic ways: both may be in the same row (two Y errors, no X errors), the same column (two X errors, no Y errors), or different rows and columns (two X and two Y errors). If any of the above parity error conditions are found, two errors are assumed to exist in the frame. This assumption is not necessarily valid but, as previously explained, will not result in a significant amount of the data being degraded.

A specific example showing how two errors could occur in a 3 word frame is shown in Figure 1. The parity used for the data of Figure 1 is the same as that shown in the frame synchronizer of Figure 2. The noise voltages without data, shown in Figure 1, are obtained by using a table of Gaussian deviates having a mean of zero and a standard deviation of one. The standard deviation is then adjusted to the case of interest by multiplying each value obtained from the Gaussian deviate table by the desired standard deviation. In Figure 1 the standard deviation σ was set equal to $1/2.05$, which is the reciprocal of the signal-to-noise rms voltage ratio in the bandwidth shown. This signal-to-noise ratio leads to a single bit error probability of 2×10^{-2} . A high error probability is used so that it is not necessary to wait long for two errors to occur.

If the data of Figure 1 were transmitted without noise, the voltage appearing on the integrator at the end of a bit time would be +1 volt for a zero. The receiver bit detector is assumed to be linear so that superposition may be used to find the result of signal plus noise. The bottom chart of Figure 1 shows the data bit voltages as received. The voltage polarity indicates whether a one or zero has been received, + for one and - for a zero. By comparing the bits to be transmitted with those received, it is found that errors have occurred at AC and BD.

In using the proposed error correction method, first the probability of each bit being in error is found from Figure 8 and then placed with its corresponding voltage. The parity errors indicate a dual error is most likely and lies either in AC and BD or in AD and BC. The conditional probability of each error is found:

$$P(e/0.049V) = 3.9 \times 10^{-1}, \text{ position AC} \quad (10)$$

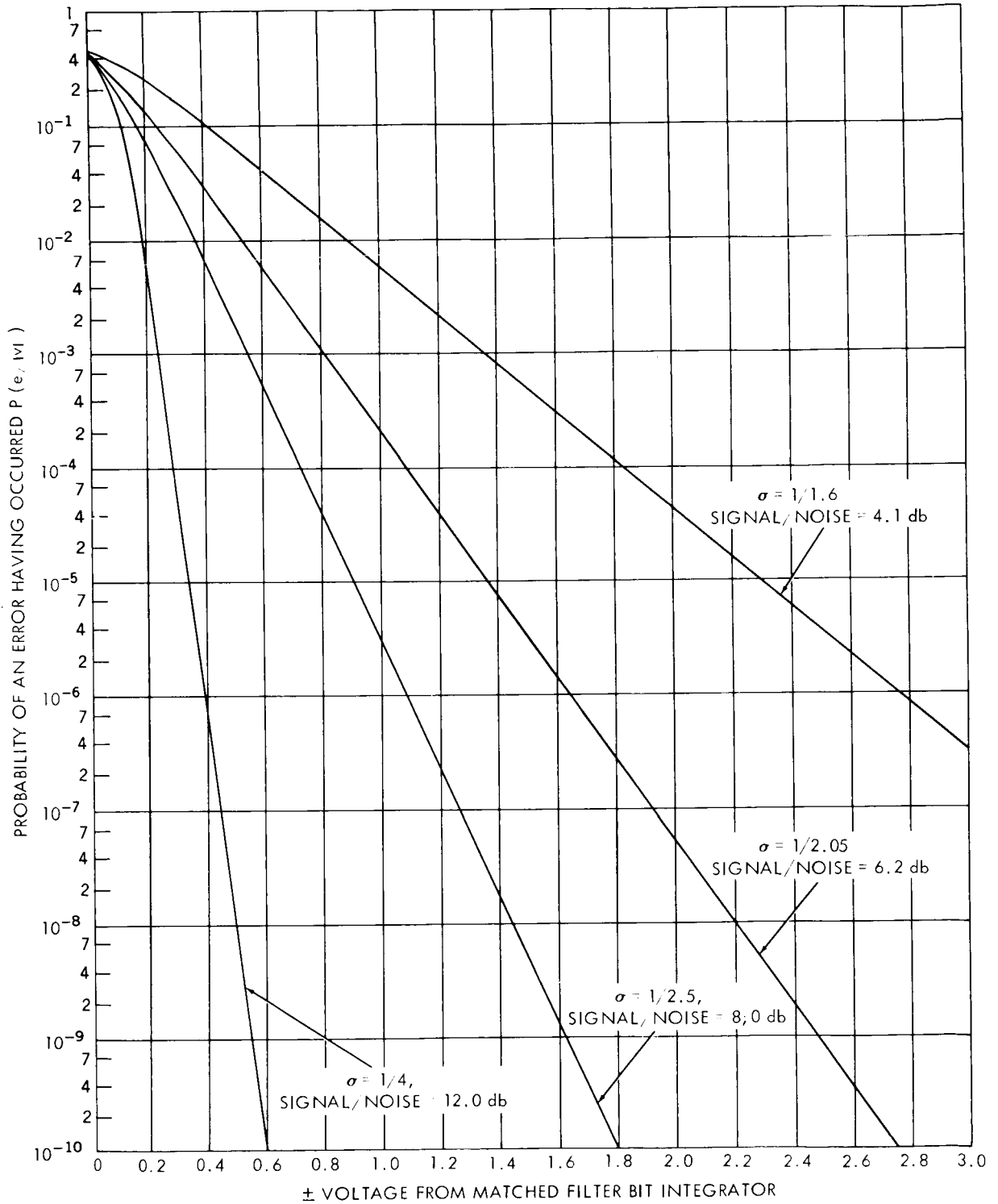


FIGURE 8.—Voltage-conditioned error probability. $P(e/|v|) = e^{-v/\sigma^2} / 2 \cosh v/\sigma^2$. The voltage from the integrator has been divided by μ . The value of the signal-to-noise ratio is for a bandwidth of one-half the bit rate.

$$P(e/0.348V) = 4.8 \times 10^{-2}, \text{ position BC} \quad (11)$$

$$P(e/1.161V) = 5.6 \times 10^{-5}, \text{ position AD} \quad (12)$$

$$P(e/0.441V) = 2.3 \times 10^{-2}, \text{ position BD} \quad (13)$$

The joint probability of errors AC and BD is found by using Equations 10 and 13

$$P(\text{AC and BD}) = 8.97 \times 10^{-3} \quad (14)$$

The joint probability of errors BC and AD is found by using Equations 11 and 12

$$P(\text{BC and AD}) = 2.69 \times 10^{-6} \quad (15)$$

The results of Equations 14 and 15 show that the dual error is most probably in AC and BD. This method of error correction has an error probability as a function of the signal-to-noise ratio associated with it. The probability of incorrectly "correcting" a frame may be calculated in a manner similar to that used for a multiple channel or character error probability calculation (Reference 4). Figure 9 shows that the equivalent gain in signal-to-noise ratio for equal error probabilities between single and dual error correction for the 2 word frame is approximately 1.8 db. If all errors detected by the crossed parity check are corrected, the maximum possible gain in equivalent signal-to-noise ratio is about 6 db.

GROUND STATION CHARACTERISTICS

The proposed organization of a ground station to process this type of data is shown in Figure 10. The output of the receiver is a noisy biphase modulated coherent subcarrier. Conventional techniques of doubling, filtering, halving, and again filtering are used to reconstruct a subcarrier which is either in phase or 180 degrees out of phase with the received signal. The output of the multiplier is then noise plus a square wave PCM signal. Data is encoded with the Manchester Code in which the phase of the subcarrier is reversed in the center of each bit time. This leads to a high transition density in the bit stream so that the bit timing detector can rapidly establish a bit timing rate equal to twice the information bit rate. The voltage on the matched filter integrator is measured by the analog-to-digital converter just prior to the capacitor being discharged, and the value is

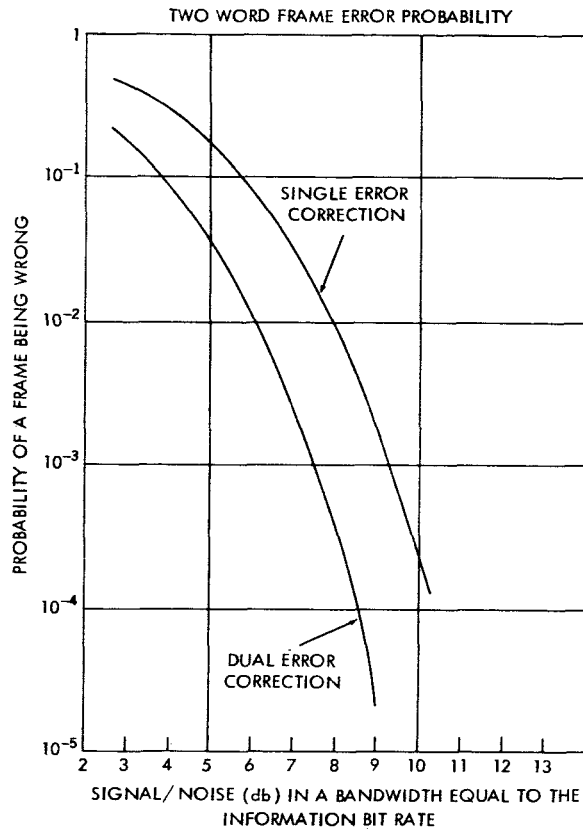


FIGURE 9.—Two word frame error probability.

transferred into a small computer. As soon as a few transitions between ones and zeros are detected in the data, the computer can establish the proper phasing of the information bit rate data within the double bit rate data. To establish the voltage that would have been on the integrator if more conventional coding had been used, the voltage on the integrator at the center of the bit time is added to the complement of the voltage on the integrator at the end of the bit time. No penalty is incurred by running the bit synchronizer at twice the information bit rate, since the signal is integrated in the proper phase over the full information bit time. If this voltage is positive, an "A" has been detected, and if negative, a "Not A" has been detected. It is not possible at this stage in the processing to state if A is a one or zero since the method of generating the local subcarrier destroys this information.

The computer is used to establish frame synchronization and to determine whether an A is a

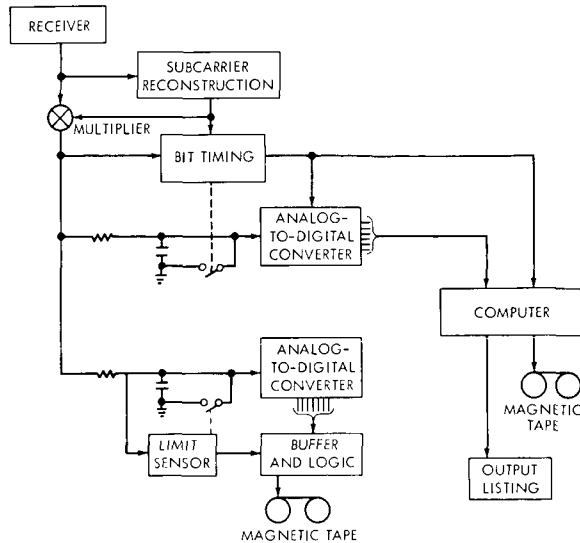


FIGURE 10.—Proposed ground station.

one or zero. There will only be a parity correlation for one of these cases since a frame with an odd number of words will be used and complementing all of the bits destroys the specified parity structure. The printed output of this computer will be frame-synchronized data indicating the rows and columns in which parity errors occur. The output voltage of the integrator for each bit is presented so that the previously described error correction procedures may be used if more than one error occurs. This output will be made available in real time and will lag the incoming data by about three frame transmission times. A small computer with a memory of 4096 words of 12 bits each, operating at a rate of about 10,000 additions per second could easily perform the computations necessary to synchronize the data. The operational rate would be about 2 bits per second with a bench test rate of 10 bits per second in a good signal-to-noise ratio environment.

Data will be lost in the system described above during the time that the bit timing detector is establishing the proper timing. In order not to lose this data and to be able to operate in a noise environment where it may be excessively difficult to establish bit timing rapidly, another system will be used to acquire data for nonreal time processing. The output of the multiplier will go to an integrator which continuously integrates this voltage. A limit sensor detects whether this integrator is approaching its voltage limit ranges

and resets the voltage to zero. The voltage on the integrator will be digitized at about 20 samples per bit time along with information about the voltage reset pulses that may have been used. In an off-line process the contents of the integrator can be reconstructed if the voltage reset times are known. Bit timing may be established by cross correlation techniques. After the subcarrier reconstruction unit establishes phase lock between the noisy incoming signal and its output, all of the data can be recovered in an optimum manner. The ratio between the subcarrier frequency and bit rate will be over 250 to 1; therefore very few bits will be lost during the time that the subcarrier reconstruction unit is acquiring phase lock.

CONCLUSIONS

The crossed parity structure is effective in establishing frame synchronization of the data and in providing an error detecting and correcting code. Frame synchronization can be established at error rates such that the majority of the data could not be corrected. The extension of the error correction process, by measuring the voltage on the integrator at the sample time (from single bit to multiple bit error correction) is effective and is a powerful tool with application to many other codes. This technique allows a significant increase in the probability of getting correct data after the error correction process is completed without increasing the complexity of the data encoder. The ground station used to recover data in this form is best implemented with the aid of a small computer.

ACKNOWLEDGMENTS

The authors are indebted to Joseph D. Barksdale who programmed the IBM 7090 to test the frame synchronization method and to obtain data for Figure 3.

REFERENCES

1. PETERSON, W. W., "Error Correcting Codes," Cambridge, Mass., M.I.T. Press, 1961.
2. WILLARD, M. W., "Optimum Code Patterns for PCM Synchronization," Paper 5-5, *Proc. National Telemetry Conference*, 1962, vol. 1.
3. Harvard University Computation Laboratory, "Tables on the Cumulative Binomial Probability Distribution," Harvard Univ. Press, 1955.
4. VITERBI, A. J., "On Coded Phase-Coherent Communications," *IRE Trans. SET-7(1):3-14*, March 1961.

DEVELOPMENT OF A 1200 FOOT ENDLESS-LOOP TAPE TRANSPORT FOR SATELLITE APPLICATIONS*

KENNETH W. STARK

The tape recorder development described herein is a low power, two speed recorder whose unique endless-loop cartridge stores 1200 ft of $\frac{1}{4}$ in. wide lubricated magnetic tape. A maximum mechanical power of 1.29 watts is required to drive the transport at 30 ips at 0°C. During operation in the record mode at $3\frac{3}{4}$ ips tape speed only 0.083 watt of mechanical power is required at 0°C to operate the transport. Speed reductions are obtained through the use of accurately machined pulleys and seamless polyester film belts. Although the optimum motor has not been obtained at this writing, wow and flutter measurements have been obtained at 1.14 percent p-p from 0-1000 cps bandwidth.

Included in this report is a discussion on the problems that arose with the tape during tests and the resultant solutions. The exceptional performance is made possible by the use of the unique tape cartridge and the accurately machined capstan assemblies which have less than 50×10^{-6} in. total indicated runouts. Duplex preloaded bearings are utilized throughout the transport. The mechanical portions of this transport have survived sinusoidal vibrations at 10 g from 5 to 2000 cps for 18 min and random vibrations at 20 g rms from 20-2000 cps for 4 min.

This paper describes in detail the design of the various components of this advanced tape recorder.

INTRODUCTION

Tape recorders are an essential part of any spacecraft instrumentation where data gathered by sensors must be stored until a position is reached in orbit when the spacecraft can be conveniently interrogated. At this point the recorder sends the information to the ground station through the spacecraft transmitter.

The main reason an endless-loop recorder is preferred to a two reel system is that multiple interrogations can be made without exceeding the required playback time to a ground station. That is, a satellite can be interrogated several times before an orbit is completed without the information loss which occurs in a two reel system. Additional advantages of an endless-loop system are that tape reversing mechanisms are not necessary, end-of-tape sensing devices are not required, momentum compensation is relatively simple, the single reel provides compact storage, and the tape velocity is in the same direction for both record and playback modes, minimizing errors due to tape skew and alignment.

The advent of interplanetary capsules and large satellites where considerable data are to be stored necessitates large capacity endless-loop tape recorders. Endless-loop recorders with storage capacities of 200 ft. of tape† are no longer satisfactory for some designs. Recorders with capacities of 1200 ft. of tape or more are required. In addition, a rigid set of design specifications must be employed to assure long life and reliability. New design considerations, with respect to obtaining a functionally operational recorder, are required to achieve the flutter, low power, long life, and low weight necessary for the recorder's increased storage.

One design parameter that must be investigated is tape life. Included in this term "tape life" are the deterioration of magnetic properties, abrasion loss of the lubricant, edge wear, and the basic characteristics of the Mylar base material and oxide binder. Each one of these considerations is affected by one or more of the following: temperature, tape velocity, mean pack diameter (which determines the slippage between outer and inner

†Falwell, R. C., Stark, K. W., and White, A. F., "A Precision Endless Loop Magnetic Tape Recorder for Space Applications," NASA Technical Note D-1542, February 1963.

*Published as NASA Technical Note D-2316, June 1964.

tape diameters), and total recorder operational life required. Because of these tentative trouble areas, just a larger cartridge to accommodate the larger tape lengths will not solve the problem.

Another area to be considered in the design is the selection of the proper motor. Again, simply utilizing existing motors as used in other tape recorders will not work. The power requirements are different, and motors must be matched to the recorder to prevent undesirable hunting which produces detrimental flutter and wow.

Capstan design (including bearing preload techniques, machining tolerances, material, and assembly and mounting arrangements) determines to a great extent how the instrument's performance will be affected by such disturbances as flutter and wow, skew and amplitude variations, etc.

Last but not least of the major design considerations is environmental testing. A 1200-ft. endless-loop recorder is much more sensitive to temperature and vibration testing than a small compact 200 ft. endless-loop recorder is. The large loop tends to be very erratic under random and sinusoidal vibration tests if left in a free condition in a cartridge.

The objective of this paper is to describe in detail the problems that arose and the solutions that were applied in the "in-house" design and development of a 1200-foot endless-loop tape recorder. The contents of this paper will be concerned with the initial development phase of the recorder.

OPERATIONAL DESCRIPTION

The 1200 ft. endless-loop tape recorder is a lightweight, low power, two speed system. The overall external dimensions are $13\frac{3}{8}$ -in. in diameter and $3\frac{1}{2}$ in. in depth; it weighs 10 lbs.

Figure 1 shows how an endless-loop cartridge operates. The tape is first wrapped onto the reel (g) and rollers (l) by starting the first wrap at the edge of the reel and the tapered portion of the rollers. Successive wraps are built up progressively towards the edge of the rollers until 1200 ft. are contained. The two ends of the tape are brought up through the tape guide plate (k) and are spliced to form an endless loop. When the cartridge is in operation, the tape is pulled from

the first wrap on the reel by the capstans (m) and it wraps back onto the outside of the tape pack. As the tape moves at constant velocity, each layer of tape slips upon the other providing tension to pull the tape back down into the cartridge on the return side (j). The operation of this recorder can be broken down into two basic modes, record and reproduce. When the recorder is operating in the record mode, the dual capstans are driven at 287 rpm. The rotating capstans impart linear velocity to the magnetic tape, drawing it across the face of the heads. In the reproduce mode, the capstans operate at 2296 rpm yielding a record-reproduce ratio of 8:1.

The $\frac{1}{4}$ -in. lubricated Mylar base magnetic tape accepts four signal channels. During each orbit, one information and one timing channel are recording a signal; therefore, with a tape record speed of $3\frac{3}{4}$ ips, 64 min. of information can be recorded. The stored information is played back in a total time of 8 min., completing one cycle of operation. It is also possible to utilize existing digital heads which make it possible to store eight channels of data on the $\frac{1}{4}$ in wide tape.

The use of dual capstans provides a constant tension across the heads, thus making possible the elimination of pressure pads and the consequent reduction in amplitude modulation, flutter, and skew.

MECHANICAL DESIGN

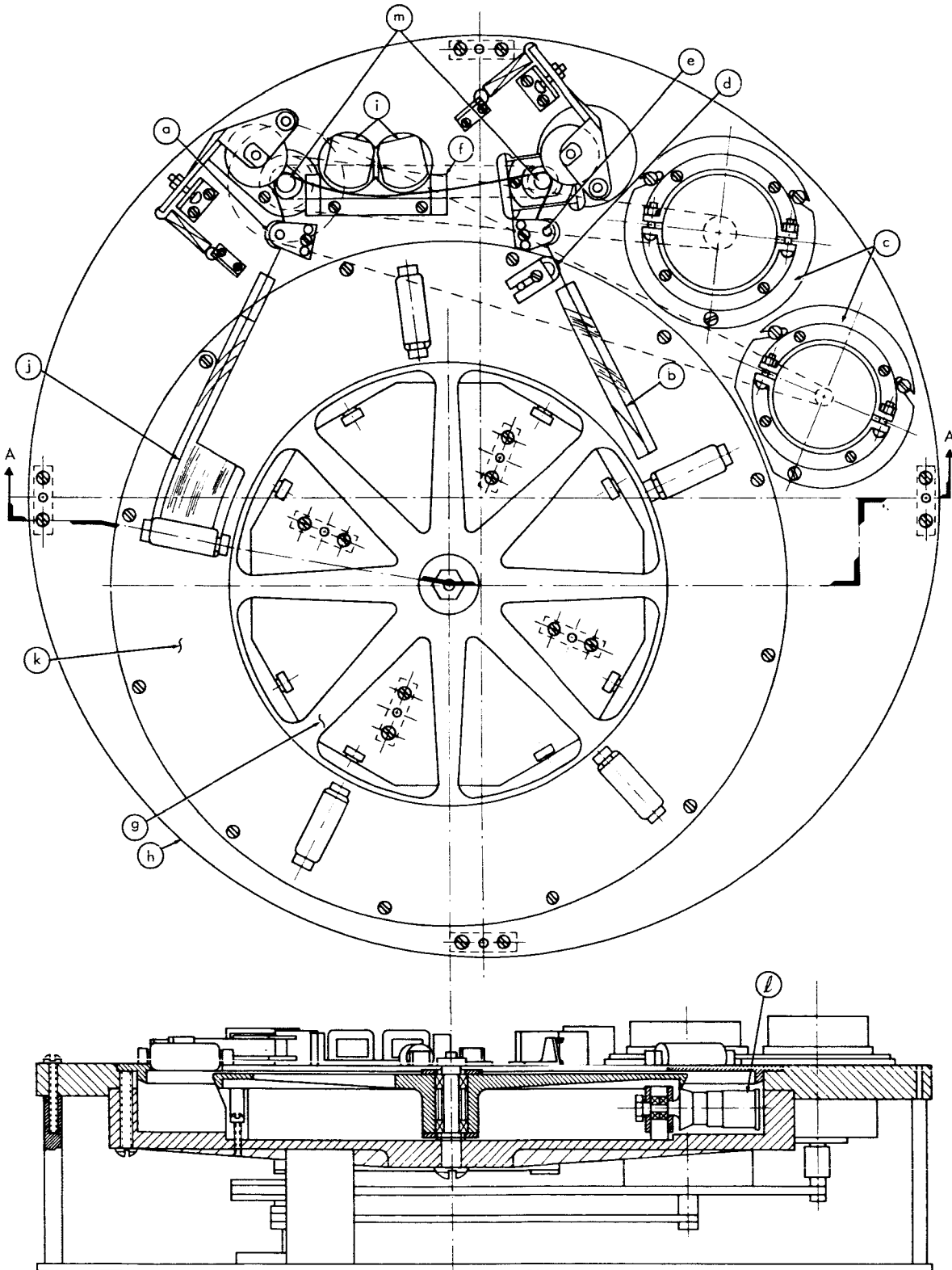
The mechanical design of the tape recorder is predicated on the following requirements:

1. An 8 to 1 speed ratio, i.e., tape speeds of $3\frac{3}{4}$ ips on record and 30 ips on playback,
2. A tape length of 1200 ft.,
3. An operational life of 9 months,
4. Low flutter for both two speed and single speed record and playback,
5. Low mechanical power for both record and playback modes,
6. Survival of the required environmental testing.

These requirements led to major design advancements in records of this type and capacity.

Tape Transport

The transport is designed to conform to the tape cartridge (g) for optimum space utilization and



SECTION A-A

FIGURE 1.—The 1200 ft. endless-loop tape transport assembly.

maximum performance (Figure 1). Essentially, two cartridges were designed and tested. A detailed description is given later.

Two capstans (m), accurately machined for high precision, provide a constant tension on the tape between them and filter out minor disturbing pulses which could possibly be generated from the slippage of each layer of tape in the tape pack. The constant tension on the tape is produced by making one capstan diameter slightly smaller than the other. A speed difference of about 1 percent is used.

The magnetic heads (i) are placed between the capstan, inserted far enough to allow for positive head gap to tape contact. No pressure pads are used, although guiding the tape across the heads (f) improved alignment. Tape guide rollers (a), illustrated in Figure 2, are employed on both the

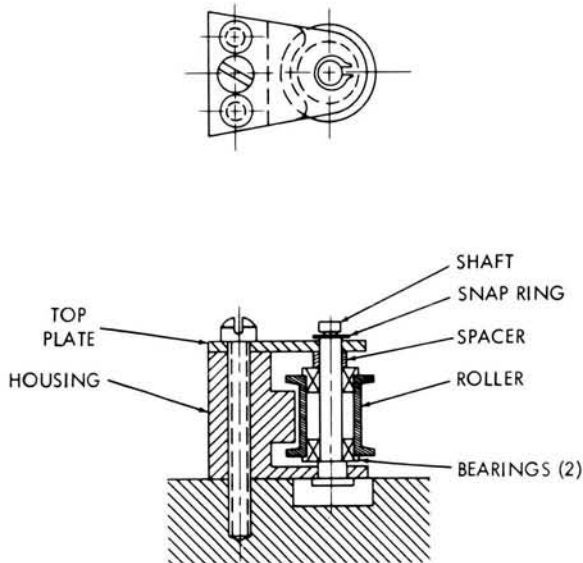


FIGURE 2.—Tape guide roller assembly.

tape entrance and exit sides to confine the tape to its proper position during performance and environmental testing. The motors (c) are located at the maximum distance from the magnetic heads. This is done to minimize the possibility of the heads being influenced by stray magnetic fields emanating from the motors, and to allow any mechanical vibrations set up by the motors to be damped by the top plate before reaching the heads. Although the exact motor or motors have not been selected at this writing, the positions

indicated were established as the most likely locations.

The use of K-1A magnesium for the top plate (h) construction provides internal damping which efficiently damps vibration induced by the rotating components.

Motors

It was not until a breadboard model of the 1200 ft. recorder (Figure 3) and later a preprototype (Figures 1 and 4) were tested that the actual mechanical power required became known. The measurements were made by using both a calibrated dc motor and a tensiometer which measured the tensile force between capstans. At room temperature, for the record speed of $3\frac{3}{4}$ ips a

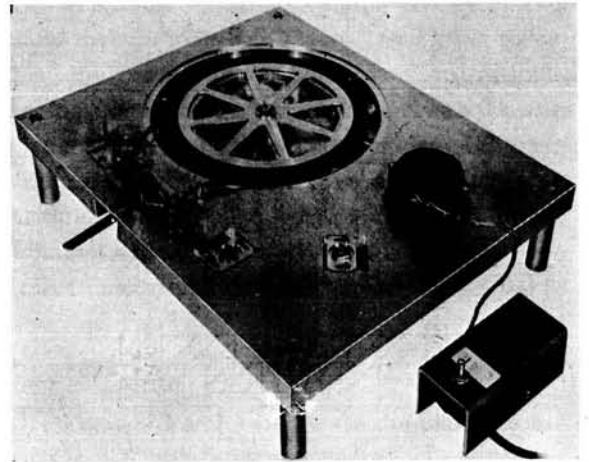


FIGURE 3.—Breadboard model of the recorder.

power of 0.0717 watt is required, and for a playback speed of 30 ips a power of 0.897 watt is required. These values vary slightly, depending upon tape type, tape pack looseness, and temperature. The powers could easily be converted to torques required at the capstans—0.34 in.-oz. at record speed and 0.53 in.-oz. at playback speed. An attempt was made to use hysteresis synchronous motors, available from a Tiros recorder, as record mode motors on the preprototype model. It immediately became evident that the torque capacity of the motor was marginal and that severe hunt oscillations of 20 cps were causing unacceptable wow and flutter values. Several other models of hysteresis synchronous motors were tried, but all contributed significant flutter

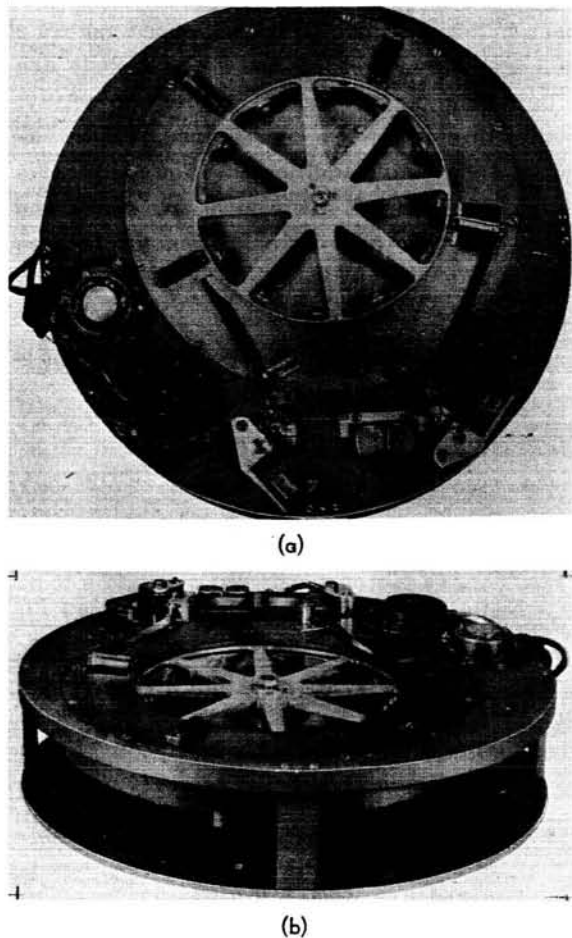


FIGURE 4.—Two views of the 1200 ft. tape transport.

because of inherent motor oscillations. The use of flywheels and friction dampeners on some of the motors did not eliminate the pulsations. To solve the immediate problem of testing the preprototype recorder, a 3600 rpm motor, used for testing of the breadboard model, was used for the high speed mode. A previously mentioned hysteresis motor with the lowest flutter contribution was used for the record mode.

However, from the motor tests it was obvious that no motor was immediately available which could be utilized as either a record or playback motor. The requirements for a good tape recorder motor are sufficient driving torque and low instantaneous speed fluctuations. In addition, the tests indicated that the motors were possibly being affected by the slippage fluctuations in the tape pack being reflected back to the motor.

Future testing will help to determine an optimum motor design.

Using power and performance data obtained from the preprototype tests, investigations have been initiated on obtaining dual speed motors of two types—hysteresis synchronous and brushless dc. The advantages of dual speed motors are elimination of one motor, higher reliability, since slip clutches are not required, and the reduction of possible additional flutter components which would result from an extra motor and extra pulleys. A brushless dc motor is preferred to a hysteresis synchronous motor for the following reasons:

1. The starting torques are similar to those of a dc motor,
2. The speed stability is the same as that of a hysteresis synchronous motor,
3. The efficiencies approach those of a dc motor,
4. The motor contains no brushes (a dc motor does), eliminating a wear and noise problem.

Belt Drives

Endless polyester film belts are used for the speed reduction and power transmission systems. These belts are efficient and accurate, and they minimize speed variations because of their uniform thickness. Their expected life is extremely high when designed properly, and their thickness (0.0015 in.) allows the use of small diameter pulleys.

Capstan Assemblies

Capstan assemblies (Figure 5) which contribute no more than 50×10^{-6} in. of total indicated runouts (TIR) at the tape driving surface are required to minimize the flutter contribution due to velocity differences. The capstan diameter is 0.250 in., a size considered adequate to maintain low percentage velocity change due to runouts.

An attempt was made to obtain the low TIR by using extremely low runout duplex bearing pairs and grinding the shafts and housings to very low tolerances. If this method were successful, bearings damaged through vibration testing could be replaced with new ones of the same tolerances without discarding or reworking a complete assembly. After manufacturing, assembling, and

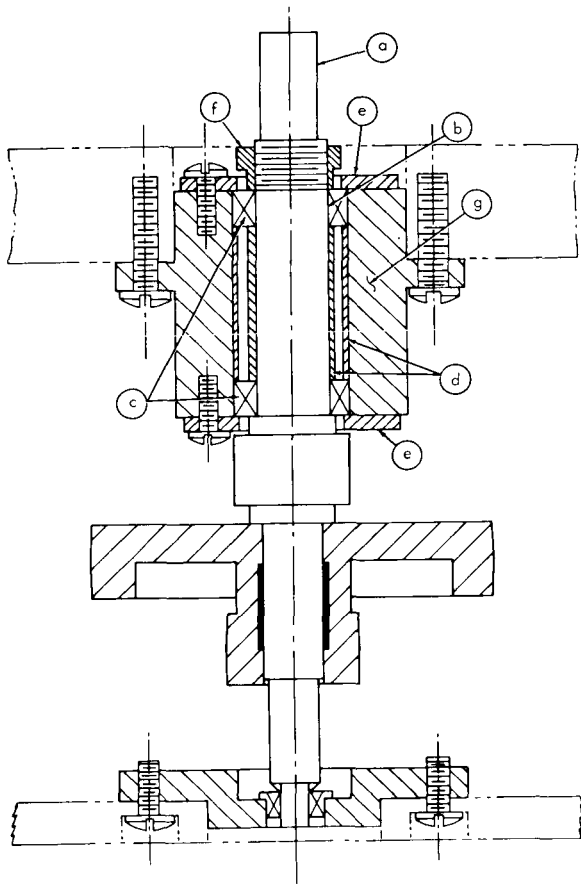


FIGURE 5.—Capstan assembly.

testing several capstan assemblies in the prototype recorder, it was found that it was not possible to consistently obtain and hold the required capstan runouts by this method. Although the tolerances with regard to diameter fit and the TIR on both the capstan diameter (a) and bearing inner races (b) were held extremely tight, there was a finite clearance in all fits. This allowed the shafts and inner race to be misaligned each time the preload was applied and, together with the machined TIR tolerances, the total radial runouts exceeded that required.

A grind-in-place technique was adopted to solve the problem. This consists, essentially, of assembling the duplex bearing pairs (c), bearing spacers (d), end caps (e), and capstan shaft (a) into the housing. The preload nut (f) is applied with enough force to preload the bearings and to prevent the inner race from rotating with respect to the capstan shaft. Caution must be exercised not

to deform the races and damage the bearing. The final operation consists of shielding the bearings from external debris and grinding the capstan to the required final diameter and TIR. This is accomplished by holding the capstan housing (g) fixed and externally driving the capstan shaft while grinding in place.

Pressure Roller Assembly

The assembly's function is to maintain intimate contact between the tape and capstan. However, a preload system has been developed whereby a variable preload could be applied to the roller bearings without resorting to special duplex pairs, end caps, spacers, and captivated nuts. The preload was determined by applying just enough force to remove bearing play (Figure 6). This setting could then be locked in place. The spring force is applied through the center of the roller. By doing this, the possibility of cocking the roller and producing skew problems in the tape is minimized. An additional area of improvement was the elimination of a bulky viscous damper that existed on a Tiros damper. The original purpose of the damper was to prevent the roller from leaving contact with the capstan when the resonant frequency of the load spring was reached during vibration testing. At this point, the tape would ride up and off the capstan. To retain this function, the roller arm was adjusted so that, during vibration, the maximum distance the roller could back off was half the total roller compression on the capstan under the 24 oz. roller force against the capstan.

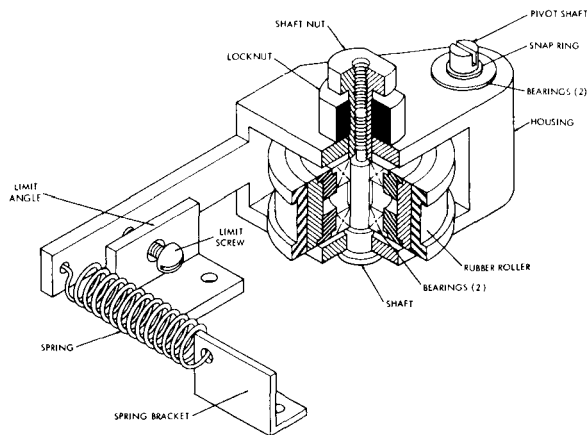


FIGURE 6.—Rubber roller arm.

Tape Path

Figures 1 and 4 illustrate the path that the tape follows during operation of the recorder. The tape is 1.3 mil thick and coated with oxide on one side and lubricant on the reverse side; it has a polyester film base.

The tape is withdrawn from the inside of the tape pack (b in Figure 1), passing over the tape guide ramp (d) and through the guide roller (e) to the capstan (m). The capstans draw the tape through the tape guide (f) and across the erase and record-reproduce heads (i). The tape then passes through the final guide roller (a) returning onto the outer layer of the tape pack (j).

The differential angular velocity between the inner and outer layer of the tape pack results in interlayer tape slippage. This means that the individual layers of tape are slipping with respect to one another, maintaining tensions throughout the pack. This tension provides a means for drawing the tape back onto the pack after it leaves the capstan. Lubrication of the tape reduces the interlayer friction, thus reducing the flutter imparted to the tape and the amount of drive power required.

Tape Cartridge

The item which required the longest development time is the tape cartridge. The only cartridges in common usage had tape capacities from 200 to 300 ft. For designing a cartridge for 1200 ft. of tape, various design considerations must be taken into account such as tape life, as defined earlier, transport operational life, performance requirements, and environmental testing. A 9 month life was a design requirement which allowed approximately 3 months of satellite testing on the ground and 6 months of operation in orbit. The life requirement and environmental testing were very severe on the tape; however, this will be discussed in another section.

The tape cartridge had to be designed so that it could effectively allow 1200 ft. of magnetic tape to function properly. Such parameters as power, flutter, signal-to-noise ratio, and operation at two speeds for record and playback had to be taken into consideration in addition to the life requirement. The 9 months of continuous operation necessitated that the cartridge design impose a

minimum wear condition on the tape. This meant that the tape must not suffer deterioration of the oxide, lubricant, and base material to the extent that its operation would be detrimental to the system as a whole.

The first approach to a cartridge design was to obtain an optimum mean diameter of the cartridge with minimum pack thickness. It is important to have a narrow pack thickness to minimize tape wear induced by the angular velocity difference between inner and outer tape layers. In addition, the mean tape diameter should be within reasonable limits to prevent an excessively heavy and large recorder design. From a curve of mean tape pack diameter vs. pack thickness for various lengths of tape (Figure 7) an optimum mean diameter of 8 in. was chosen for a pack thickness of 0.8 in. for 1200 ft. of tape.

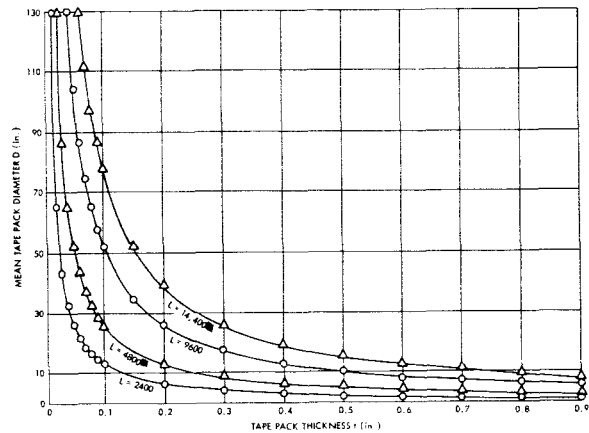


FIGURE 7.—Mean diameter of the tape reel vs tape loop thickness. $D = WL/\pi t$ where W is the tape thickness and L is the tape length.

After the basic size of the cartridge was determined, there were two design possibilities. One was to expand the Tiros type into a larger size. Also, as a result of preliminary tests, it was found that flutter was appreciably reduced when the tape was supported on rollers during operation. Thus the second design consisted of utilizing rollers to support the tape in the cartridge.

A large Tiros type of cartridge (Figure 8) was built and installed in the breadboard model for testing. Performance tests were run at various speeds. Although the cartridge appeared to function well, severe tape pulsations occurred which

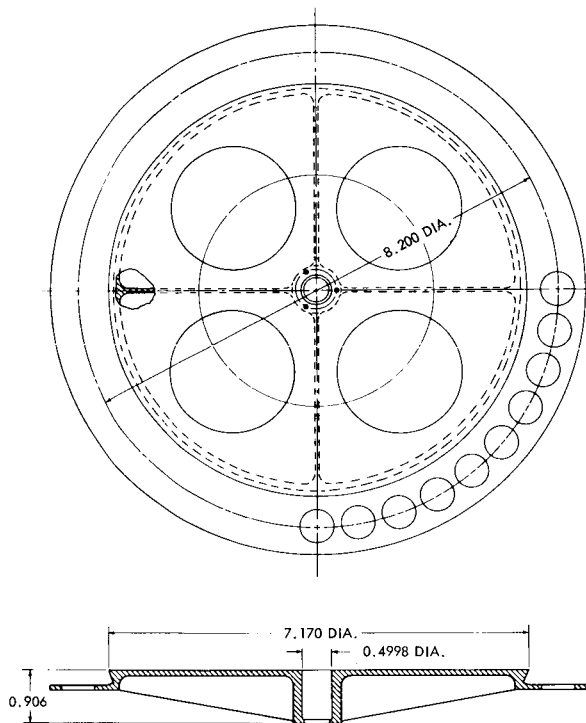


FIGURE 8.—Tape reel and cross section.

increased with tape speed. Therefore, this type of cartridge was eliminated from further consideration because the pulsations resulted from a condition of excessive tape friction between each layer and between the tape edges and reel flange. This causes accelerated tape wear and is detrimental to overall recorder reliability. Secondly, because of the magnitude of the pulsations filtering could not be easily accomplished with simple devices. This would result in large periodic flutter and wow disturbances.

At this point, the roller-reel tape cartridge was put into the breadboard model (Figure 3) and testing proceeded. The main problem area developed with the tape support rollers. The original rollers were straight cylinders. During testing, it was observed that upon acceleration the tape would form a loop on the return side which would not pull in and which remained at a constant length. This situation resulted because of the lack of friction between the outer layers of tape and the tape rollers. Friction did exist between the outer layers of tape and the flange of the larger Tiros reel. With the Tiros type of cartridge, the flange which supports the tape has a

higher linear velocity than the tape itself at every point; therefore, there is a constant frictional force pulling the tape into the cartridge on the return side. This effect is absent with the rollers and actually works in reverse where the tape must drive the rollers. To correct this situation, a series of tests was conducted that resulted in two cylindrical steps being added to the straight roller section. It was found that the rpm of the rollers was determined at the smallest diameter which provides a higher circumferential velocity at the two larger steps. The resultant relative velocity between the roller and tape edge was sufficient to prevent the tape from forming and maintaining a loop on the return side of the cartridge. With this problem solved, the cartridge functioned uniformly; and it was decided that further recorder development would be centered around this cartridge.

The cartridge (Figure 9) consists of a tape reel (a), 12 tape support rollers (b), a roller mounting

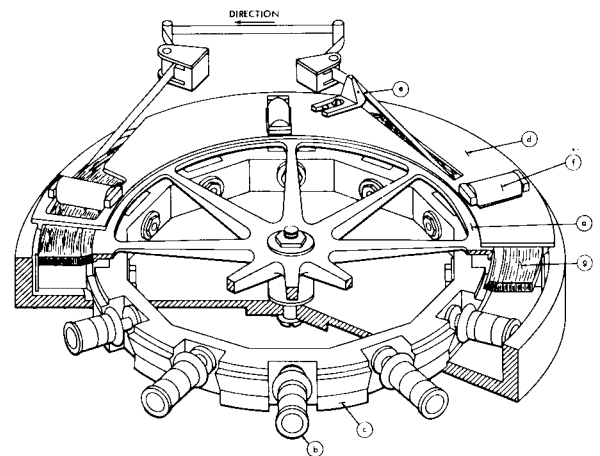


FIGURE 9.—Tape cartridge.

ring (c), a tape guide plate (d), a tapered guide (e), cover plate rollers (f), and, of course, 1200 ft. of lubricated Mylar base magnetic tape (g):

1. Tape reel—Figure 10. It was found that if other than a cylindrical surface to which the tape pack conforms is used, the failure rate of the tape pack is very high. This is because wear increases at the areas where bending of the pack occurs (e.g., the 4 corners in a square pack). In addition, the cylindrical surface, with a profile such as this reel has, allows the tape to emerge from the

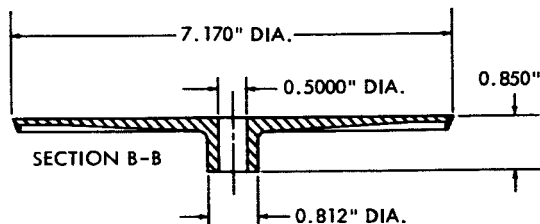
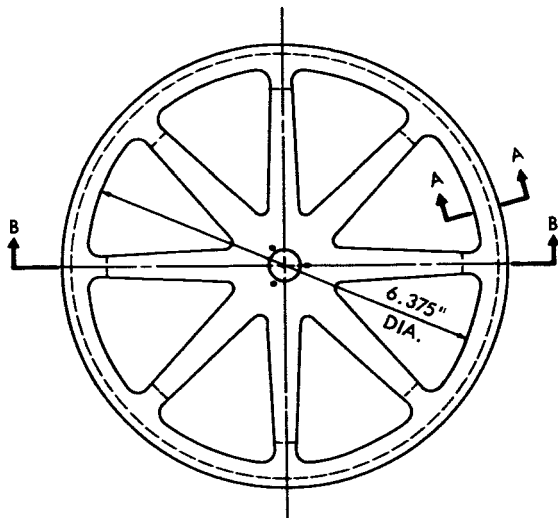


FIGURE 10.—Tape reel.

cartridge with minimum tape disturbance due to flexure, edge wear, and oxide abrasion. This provides maximum performance and minimum wear. The reel is mounted in duplex bearings to insure a stable and smooth reel rotation and provide protection of the bearings during environmental vibration testing.

2. **Tape support rollers**—Figure 11. These rollers essentially comprise the “heart” of the tape cartridge. Twelve rollers were selected to prevent sagging of the tape between them and to obtain a circular tape pack.

As testing proceeded, it was determined that the two steps added to the straight cylindrical roller required vapor honing to increase the pull in force on the tape on the return side. Each roller is supported in two bearings in the roller mounting

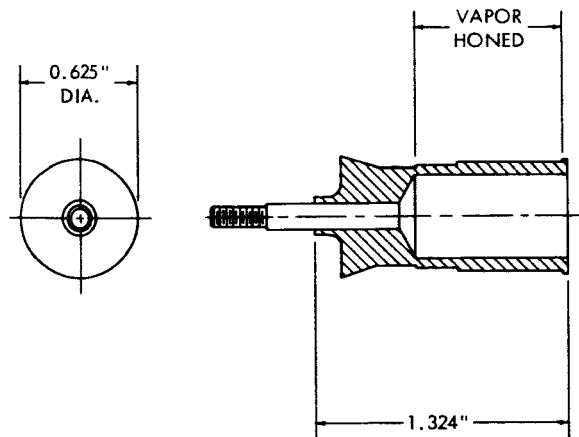


FIGURE 11.—Tape support roller.

ring (Figure 12) which are not preloaded; however, the radial play of the bearings is limited to between 0.0001 and 0.0002 in. The bearings are lubricated with MIL-L 6085A oil. Preloaded bearings cannot be used; they cause intermittent rotation of the rollers. Each roller is made with the tapered portion of its cross section nearest the

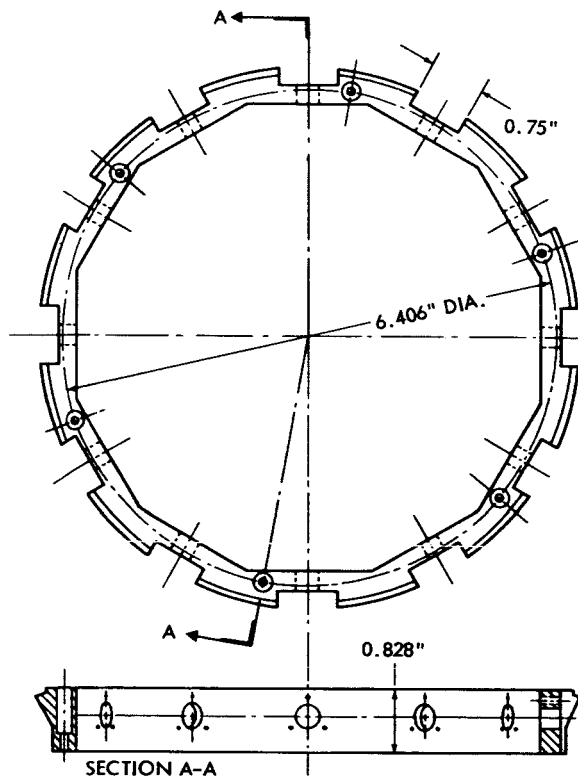


FIGURE 12.—Roller mounting ring.

reel. On both the reel edge slope and the roller taper, as seen in the profile section when assembled, the resulting cross section represents the attempt made to bring each layer of tape into an ideal position before it exits the cartridge. Thus, minimum strain and abrasion on the tape during the cartridge operation is exerted.

3. Roller mounting ring—Figure 12. The ring provides mounting of the tape support rollers and a meshing fit of the outer surface of the tape reel. This type of fit is a safety design to prevent the tape from possibly working its way between the conical portion of the tape support rollers and the tape reel edge.

4. Tape guide plate—k on Figure 1. The plate provides for mounting of the guide rollers which prevent the tape from rubbing on the underside of the plate when extraneous disturbances cause the tape pack to rise. In addition, this plate acts as a baffle to prevent debris from falling into the tape and cartridge.

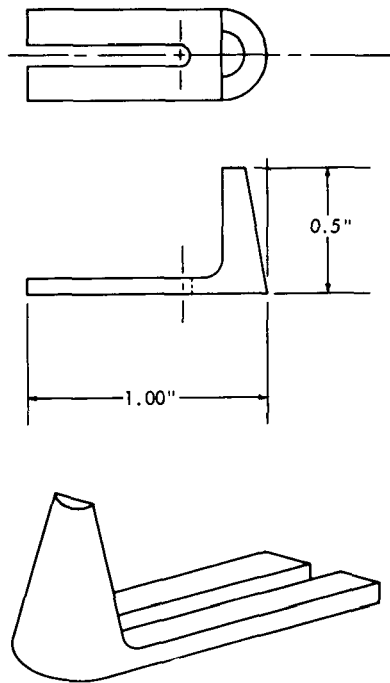


FIGURE 13.—Tapered guide.

5. Tapered Guide—Figure 13. The guide aids the tape in making the transition from an inclined to a horizontal position as the tape emerges from the cartridge.

TAPE SELECTION

Although the problem of obtaining a functionally operating endless-loop tape cartridge for 1200 ft. of tape had been solved, an area which required investigation still remained. This area concerned the magnetic tape itself. The tape which was used for smaller endless-loop cartridges on the Tiros and Nimbus recorders was not satisfactory for use on the larger cartridge. Initial testing with this tape revealed deficiencies in the mechanical stability of the tape during temperature testing. Primarily, there were two temperature problems. The oxide binder became soft at 70°C and the tape had a very high shrinkage rate.

Temperature tests were run on the 1200 ft. endless-loop cartridge using Minnesota Mining and Manufacturing Company's LR 1220 tape at 30 ips and temperatures of 0°C, 25°C, and 60°C. At 0°C, the transport operated when the recorder was placed in a plastic bag, flushed with dry air, and sealed before being brought down in temperature. This was necessary to prevent the cartridge from jamming. It is felt that this occurrence is due to ice condensing on the transport and tape.

At 25°C the cartridge performed satisfactorily. However, at 60°C the tape exhibited excessive interlayer friction. This interlayer interference caused the cartridge to tighten up and eventually jam. A series of tests was conducted to determine whether the oxide, lubricant, or both were causing the trouble. At about 70°C the sliding friction of oxide against oxide increased, but the sliding friction of lubricant against lubricant remained constant. At 100°C the same effect was noted. This fact was brought to the attention of the tape manufacturer, who conducted his own tests and found that the oxide binder was stable only up to about 80°C.

Although the recorder temperature test was run at 60°C, it is possible that the interlayer angular velocity difference causes a localized heating effect which raises the tape surface temperature above 70–80°C.

As a result of tests, the manufacturer provided new samples of tape (LR1259 and 8943) which were considered identical in magnetic and lubrication properties to the LR 1220 but had a higher temperature binder on the oxide side. LR 1259 was tested at 60°C in the cartridge and operated considerably longer than LR 1220 at the same

temperature, indicating the new binder was the solution to this problem.

However, at this point the second problem arose. When LR 1220 tape was used, the cartridge never operated long enough at the high temperature to observe other characteristics. As the LR 1259 continued to run at 60°C, it was observed that the tape pack was gradually tightening up. Finally, the point was reached where the pack became so tight that the cartridge jammed. This tightening process was attributed to the shrinkage of Mylar at high temperatures. This characteristic of Mylar is well known, but the exact rates of shrinkage were not as well known for various temperatures. Tests were run on LR 1220, LR 1259, 8943, and a tensilized tape sample to determine their shrinkage rates (Figure 14).

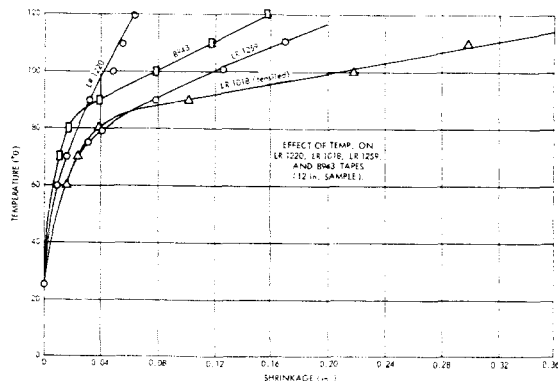


FIGURE 14.—Effect of temperature on LR 1220, LR 1018, LR 1259, and 8943 tapes (12 in. sample).

The percent shrinkage for LR 1259 at 60°C is 0.13 percent; however, if localized heating raises the temperature to 80°C, the shrinkage becomes 0.34 percent. For 1200 ft. of tape and 0.34 percent shrinkage, the tape becomes shorter by 49 in., thus increasing tape tension excessively. The percent shrinkage for LR 1220 at 60°C is 0.075 percent and at 80°C it becomes 0.19 percent. For 1200 ft. of tape with 0.19 percent shrinkage, the tape will lose 27.4 in. or about half that of LR 1259. The 8943 tape had a shrinkage rate of 0.13 percent. Base Mylar (0.001 in. thick) had a shrinkage rate of 0.284 percent.

Because these results showed large shrinkage rates and the variation of rates with tape type,

it was decided to preshrink the LR 1259 tape at 100°C for 1 hour and then repeat the test in the cartridge at 60°C. After this the tape ran continuously at 60°C without any indication of tightening or jamming; however, during life testing a wear problem became apparent with the lubricant. Again the tape manufacturer studied the problem and supplied a new sample which passed the required 9 month accelerated life test. At present, this is the only tape which was found to satisfy the rigid temperature and life requirements. There were not enough samples to determine the recording characteristics although it was specified that the oxide should have the same magnetic properties as the tape it replaced. This new tape has the designation LR 1353.

PERFORMANCE CHARACTERISTICS

The specifications in Table 1 were taken with both the breadboard and preprototype models utilizing two motors, one for record and one for reproduce, not the flight model motors. In addition, the signal-to-noise ratio taken was obtained by observing the average signal level of a reproduced signal and then observing the amplitude noise associated with this signal. The standard technique of reading the residual noise level after the signal is erased was not used because the method utilized describes the functioning of the transport more accurately. That is, any disturbance due to skew, capstan misalignment, or pressure rollers that affects the tape contact with the heads will be noticed. Measurements on the preprototype model were taken only at room temperature because the motor arrangement did not allow enough room and flexibility to permit operation in the available environmental temperature chamber.

From the data in Table 1 a distinct improvement is seen from the breadboard stage to the preprototype model. Although temperature tests were not taken at this time on the preprototype, it can be seen that temperature variations did not affect the breadboard excessively.

ENVIRONMENTAL TESTING

This transport has survived operation at 0°C and 60°C; however, when vibration testing was initiated, a problem of tape spew and tightening arose. Normally, in recorders of this type the

Table 1. Performance Specifications.

Measurement	Model				
	Breadboard			Preprototype	
Temperature (°C)	0	25	60	25	25
Record Speed (ips)	40	40	40	3-3/4	30
Reproduce Speed (ips)	40	40	40	30	30
Flutter (percentage p-p), 0-1000 cps	0.945	0.96	1.12	1.14	0.64
Power to Drive Entire System (watts)	1.29	1.00	0.640	0.0717 record	0.897
				0.897 reproduce	
Power to Drive Tape Alone (watts)	0.120	0.150	0.135	0.0207 record	0.170
				0.170 reproduce	
Signal-to-Noise Ratio (db)	22.4	22.5	24.4	40	-
Amplitude Modulation (percentage p-p)	-	-	-	-	3.8

motor is run to allow the major rotating assemblies which are mounted in ball bearings to rotate during vibration testing. This technique was attempted at various tape speeds with the transport isolation mounted, hard mounted, or with an assortment of tape guiding methods. Tests were also conducted with the motor not operating with negative results. Finally, the tape pack was held stationary by placing snubbers 180 degrees apart around the outside of the tape pack. They each exerted 3 lb. of force on the pack. This method coupled with isolation mounting enabled the tape pack to remain stable during vibration testing; however, additional testing is required to prove the system. The isolation mounts limit the amplification at resonance to 4. If this method is used in the final models, the snubbing action will be applied utilizing solenoids which will make possible the release of the snubbing force after launch, to allow the recorder to become operational.

The following specifications are the vibration levels to which the recorder was tested:

1. Sinusoidal: 10 g, (limit to ¼-in. single amplitude), 5 to 2000 cps, 18 min. duration.
2. Random: 20 g rms, 20 to 2000 cps, 0.2 g²/cps spectral density, 4 min. duration.

CONCLUSIONS

The development of this transport, as described, does not contain equipment, designed for flight, such as electronics, motors, and a pressurized container, which would allow the transport to fly as a system in a satellite. As of this writing, no known recorders of this type have flown in a spacecraft. Future "in house" activities will consist of:

1. Testing the transport for possible application as a TV recorder in the meteorological program,
2. Reducing the overall cartridge and recorder design configuration to obtain minimum volume and weight, and
3. Designing a cartridge and transport to operate with 1200-ft. of ½ in. tape.

Design and development efforts are continuing on an integrated transport system to include transport, electronics, enclosure, mounting techniques, etc. capable of surviving prototype environmental testing.

ACKNOWLEDGMENTS

The author wishes to express his appreciation for the major contributions of Messrs. William Burton and John Humphreys which aided in the design and development of this tape transport.

SURVEY OF CONTINUOUS LOOP RECORDERS DEVELOPED FOR AND FLOWN ON METEOROLOGICAL SATELLITES*

KENNETH W. STARK AND ARTHUR F. WHITE, JR.

Meteorological satellites are used to sense and collect data of the atmosphere. The most economical and efficient method is to store data taken by the satellite by a magnetic tape recorder. For meteorological satellites a continuous-loop magnetic tape cartridge recorder best fulfills this requirement. The most important advantage for using this type of recorder is that the latest data are always transmitted to the ground station.

When initiating a design program for the use of magnetic tape recorders in satellites, a severe set of design parameters is imposed. Rocket vibrations, space environment, and electro-mechanical interfaces are reflected in the design. The first continuous loop recorder was designed to fly on the Vanguard satellite. The next generation of recorders had more sophisticated requirements. These were for the TIROS and Nimbus satellites. For these recorders 200 feet of tape was satisfactory.

Finally, the requirement to store large quantities of information with multiple interrogation capabilities necessitated the development of continuous loop recorders containing 1200 feet of magnetic tape. This resulted in new design problems involving the magnetic tape in temperature variation and wear characteristics. Vibration problems arose because of the larger tape pack.

This paper describes the recorder design concepts, products evolved, and their utilization in meteorological satellites.

INTRODUCTION

Meteorological satellites are used to sense and collect data of the atmosphere. If there were no storage capability within the satellite, an exorbitant number of ground stations, to which the data could be transmitted without loss, would be required. Considering the expense involved per ground station, the total cost would be prohibitive. To circumvent this problem, a means by which this data can be stored until a ground station is within range must be found. The most economical and efficient method is to store the data on magnetic tape until readout is required. For meteorological satellites a continuous loop magnetic tape cartridge recorder best fulfills this requirement. The most important advantage for using this type of recorder is that the latest data is always transmitted to the ground station. With this unique capability multiple interrogations are made possible. That is, the recorder can be interrogated at any time during its orbit and

always transmit the latest information to the ground station within the time allowed. Additional advantages for using this type recorder are that:

- a. The single reel provides compact storage,
- b. reversal mechanisms are not required for record and playback functions,
- c. the tape can drive beyond one pass without requiring safety cutoff devices for the motors, and
- d. momentum compensation is simple when required.

Designing, fabricating, and testing magnetic tape recorders for satellite application is unique.

Satellite recorders must operate and perform under unattended conditions in a spacecraft. Once the recorder is in orbit magnetic heads and tape cannot be cleaned, bearings cannot be relubricated or replaced, components cannot be adjusted or changed; in essence, periodic maintenance is impossible. Thus, it becomes necessary to design the recorder to survive the launch and space environment and then expect it to perform

*Published as *Goddard Space Flight Center Document X-652-64-172*, June 1964.

satisfactorily under unattended conditions. As a result, it is subjected to a rigorous test program to qualify the design and workmanship on a prototype and flight models. The major qualifying tests are random and sinusoidal vibration tests for the rocket launch, vacuum thermal tests for the orbit environment, humidity tests for storage and transportation conditions on earth, and operational and life performance tests to demonstrate its maintenance free capability. Thus, rocket booster vibration, orbital environments, and critical electro-mechanical interfaces with the total satellite system are reflected in the recorder design.

During the launch phase, a combination of steady state acceleration, and sinusoidal and random vibrations are transmitted from the rocket to the satellite and consequently to the recorder. These vibrations result because of rapid changes in vehicle weight due to fuel consumption, sudden maneuvers, engine pulsation, erratic burning of fuel, acoustical impingements, etc. The recorder must be designed to survive the initial launch environment.

Once in orbit, it must then perform satisfactorily under the zero gravity field conditions and temperature changes as the satellite experiences the sunlight and dark portions of the orbit.

To eliminate space vacuum effects on the bearings, rotating parts, and many dynamic surface contacts, recorders are pressure sealed. However, thermal radiation in a vacuum environment to the unit must be considered in the design.

The electro-mechanical interface between the recorder and the rest of the spacecraft system is a significant factor in its overall ability to survive the launch and orbital environment. Of extreme importance is consideration of magnification factors during resonant modes in vibration and the thermal heat transfer effects.

The orbit obtained by a satellite is a function of vehicle thrust and total weight of vehicle and payload. As a result, in the earlier Meteorological programs, it was necessary to design each component with strong emphasis on minimum weight without sacrificing performance and reliability.

Power consumption was and is still a significant factor in a design. Power availability is determined by the batteries and solar cells contained

on the spacecraft. Power systems are now much larger. Although the total number of systems flown on a single spacecraft has increased, the unit power availability is now greater and has alleviated what was once stringent power constraints on recorders.

THE VANGUARD RECORDER

The design and development of continuous loop recorders for Meteorological satellites originated in the Vanguard II program which was part of the United States' space program for the International Geophysical Year. The Vanguard II was a 21 inch diameter satellite (Figure 1), designed



FIGURE 1.—Vanguard Satellite.

to contain an Infra-Red system for observing the cloud cover of the earth. The design of such an IR system was a challenging task when initiated in 1956. The objective was a total satellite weight of less than 21 pounds, consisting of approximately 7 pounds of batteries, 8 pounds of structure and sensors and 6 pounds of electronics. An endless-loop recorder was required which could store 50 minutes of data, reproduce it in 1 minute, survive the vehicle steady state accelerations and vibrations during launch, be designed

for minimum power and weight, and survive the temperature variations in outer space.

Experience in designing endless loop recorders to survive the various rocket and orbit environments was nonexistent. In addition, the vibrations of the vehicle transmitted to the recorders was known to be random in nature. Design techniques for response of endless loop recorders to random vibration was unknown. Therefore testing of units was rigorous and empirical. Success could only be assured by a painstaking quality selection of a few samples which underwent extreme environmental and performance tests. As a result, a total of nine (9) recorders was required to yield three (3) which could be actually considered flight hardware. The performance of the three units selected, however, remained consistent throughout tests and proved adequate for the intended mission.

The Vanguard II was launched on 17 February 1959 and its tape recorder was the first endless loop recorder to be flown in an earth orbiting Meteorological satellite. Although the satellite was to be spin stabilized it went into a tumbling mode resulting in an extremely difficult data analysis task. However, the highly successful operation of the system itself for its full two-week life substantiated the design and test approach which had been taken and was a significant step forward in the design of endless loop recorders for satellite instrumentation.

This recorder was packaged as part of a cylindrical instrumentation can (Figure 1), consisting of batteries, the recorder, and an electronics deck. The entire instrumentation package slipped into a cylindrical can under compression which was part of the satellite structure. An "O" ring seal was used to maintain atmospheric pressure in the can.

The Vanguard recorder (Figure 2) contained 75 feet of lubricated tape. It had a record speed of .3 IPS and a reproduce speed of 15 IPS. Because the recorder was not vibration isolated, to minimize amplification factors, the recorder was placed at the bottom of the Instrumentation package and compressed slightly when secured. Spring preloading of bearings in the recorder itself was used wherever possible.

The recorder was tested through a temperature range of 0°C to 60°C and random vibration at 30 G's rms from 20-2000 cps.

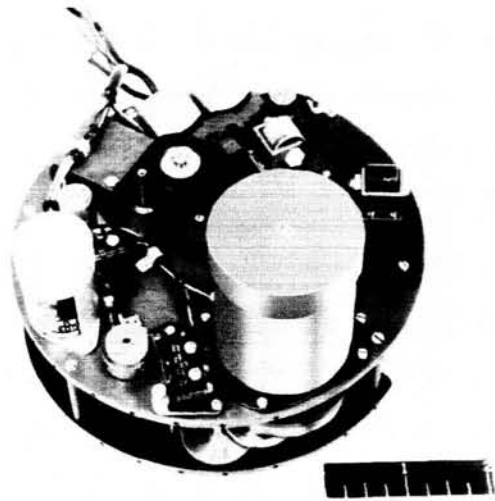


FIGURE 2.—Vanguard II recorder.

To avoid the detrimental effects of space vacuum on the recorder components the entire instrumentation package was pressurized and sealed.

Because the booster used had a relatively low thrust it was necessary to minimize weight wherever possible in order to obtain the desired orbit. Magnesium was used extensively throughout the recorder. It weighed 1.5 pounds and was 3 inches high and 5.5 inches in diameter.

In addition, the satellite's only source of power was storage batteries with no recharging capability. To produce a minimum drain on this power supply the recorder utilized two high efficiency low power dc motors without speed regulators. The record motor utilized .09 watts and the reproduce motor utilized 2.0 watts of power. Flutter was about 6 percent P-P.

The emphasis placed on this recorder was for operation at minimum power and weight rather than accurate performance.

THE SCORE RECORDER

It should be noted at this point, however, that a modified version of this recorder (Figure 3) was flown in a communications satellite before Vanguard was launched. This satellite was part of the Atlas-Score project. It was launched in December 1958 and is believed to be the first known endless loop recorder flown in an earth orbiting satellite. The instrumentation package

THE TIROS RECORDER

The TIROS (Television Infra-Red Observation Satellite) program (Figure 5) produced the 2nd generation of endless loop recorders for Meteorological satellites.* An endless loop recorder was included as part of the Infra-Red system in this spin stabilized satellite series. For the

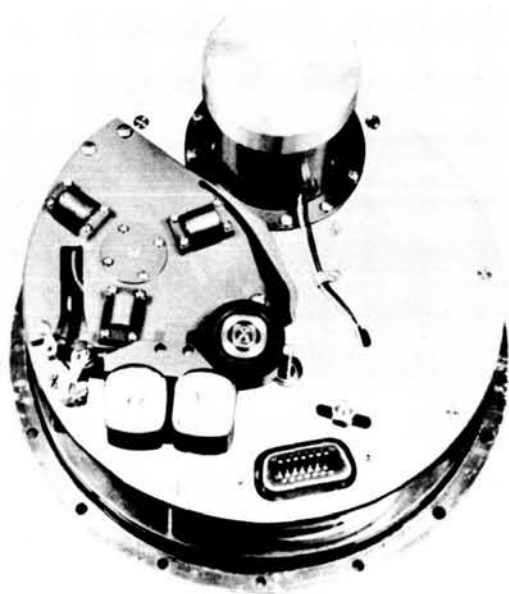


FIGURE 3.—Courier recorder.

is shown in Figure 4. Although the Vanguard recorder was completed and qualified before the modified Score recorder, delays in the Vanguard launch schedule resulted in the Score satellite being launched first. The Score recorder also proved highly successful in performance during its orbit lifetime.

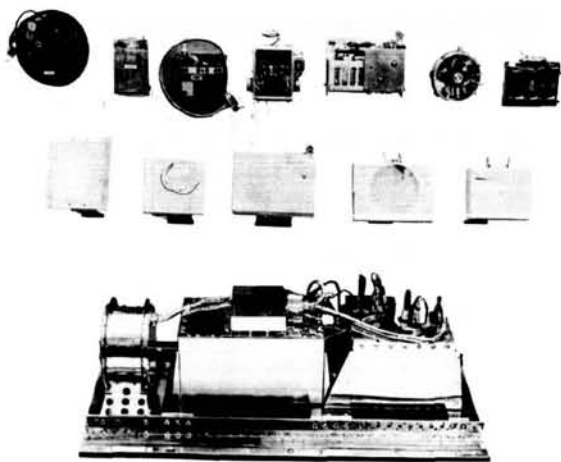


FIGURE 4.—Courier instrument package.



FIGURE 5.—Tiros Satellite.

TIROS program performance requirements were more stringent and considerable design changes were necessary in the recorder. Orbit life of 6 months rather than two weeks was required, wow and flutter was reduced to 2.5 percent P-P rather than 6 percent, and 200 feet of tape instead of 75 feet was required.

Experience gained in the test program of Vanguard II and the increased knowledge of rocket launch environments enabled more thorough design analysis to be utilized. More design emphasis had to be applied to the preloading of ball bearings, polyester film belt drivers,† and to the precision fabrication of the recorder compo-

*NASA TN D-1542 "A Precision Endless Loop Magnetic Tape Recorder for Space Application"—R. C. Falwell, K. W. Stark, A. F. White.

†NASA TN D-668 "Polyester Film Belts"—J. Licht and A. White.

nents in order to obtain the required life and performance and to survive the environmental conditions.

The basic packaging technique of the IR system recorder and electronics components in a cylindrical pressurized can used successfully in Vanguard II, was retained (Figure 6) in the IR system in TIROS.

The TIROS endless-loop recorder (Figure 7) required 200 feet of lubricated tape with a record

speed of .4 IPS and a reproduce speed of 12 IPS. Low power and weight requirements combined with accurate performance and reliability demanded much greater precision fabrication techniques than those used for the Vanguard recorder.

The recorder was "hard-mounted" in the bottom of the instrumentation can under a slight compressive force. The tape recorder (Figures 7 and 8) is a two speed recorder utilizing an endless loop tape cartridge which stores the 200 feet of $\frac{1}{4}$ inch lubricated magnetic tape. The record system employs a 2 phase 137.5 cps, 14 VAC hysteresis loop synchronous motor requiring less than .300 watts to operate and drive the tape at a speed of 0.4 inches per second. The recorder utilizes a 5000 RPM motor requiring less than 1 watt running power, during the playback mode, driving the tape at a speed of 12 inches per second. A transistorized dc speed control unit maintains better than 1 percent regulation of this motor. Speed reduction of the motors to the capstan drive are obtained through reduction pulleys incorporating polyester film belts. Frictional drag from components not being used during certain operating cycles is minimized by using spring clutches. This prevents power drain on the motor during the normal operating cycles. Because it was desired that the record motor operate continuously, the use of spring clutches allow the motor to be over-ridden during the playback mode. Flutter and wow is maintained below 2.5 percent peak to peak from 0 to 1000 cps bandwidth. This exceptional performance is made



FIGURE 6.—Tiros instrumentation components.



FIGURE 7.—Tiros tape recorder.

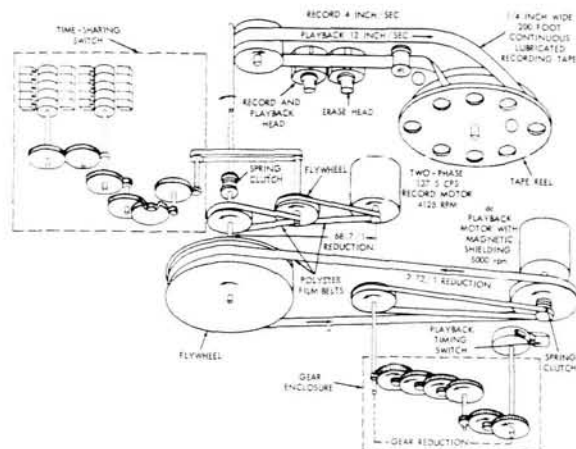


FIGURE 8.—Tape recorder schematic.

possible by use of an extremely accurate gyro-type capstan assembly. The capstan has maximum runouts of 50 millionths of an inch and the assembly utilized the duplex bearing and integral race technique. The bearings in this assembly are preloaded by a fixed center distance which requires very accurate measuring and machining.

Special precautions had to be taken in the duplex preloading arrangement to prevent damage due to the temperature extremes of environmental testing. Wherever thermal expansion, due to differences in materials, caused preloading inconsistencies, materials which had the same coefficient of expansion as the bearings were used. Usually this resulted in compromising the weight requirements. This method of preloading was extended for use in the motor assemblies where shaft torque output had to be a maximum while still obtaining the advantages of preloaded bearings.

The entire recorder was sealed in the instrumentation package under one atmosphere of dry air to protect and prevent oil loss from bearings and other detrimental effects on components and materials due to outgassing.

Magnesium was used extensively for weight reduction. K-1A magnesium was used for the main plate because of its high internal damping characteristics. This aided in reducing the drumming effect due to rocket vibrations and in reducing the mechanical vibrations transmitted to the magnetic heads due to motor and shaft rotations. The resultant recorder weighed 4 pounds and was 2.3 inches high and 6.25 inches in diameter.

A low power requirement was imposed on this recorder resulting in a special motor development which produced a hysteresis synchronous motor with about 25 percent efficiency for the record mode. Because of higher power required for the playback mode a high efficiency dc motor was used with a speed control.

These recorders have flown in the IR system of TIROS II, TIROS III, and TIROS IV, with launch dates respectively on November 23, 1960, July 12, 1961, and February 8, 1962.

At the time of the last complete interrogation on TIROS II on May 2, 1962 when it had been in orbit for over 17 months, the entire IR electronic package with the tape recorder was performing perfectly. The only degradation noticed

occurred on the tape which had run continuously for the 17 months. The TIROS II battery pack (not part of the IR package) had degenerated so that further interrogations of the satellite were no longer possible.

In TIROS III, the electronic package was functioning perfectly at the time of its last interrogation 7 months after launch. The successful launch of TIROS IV with a new IR system resulted in the stopping of interrogations on TIROS III.

In TIROS IV, the electronic package performed perfectly up to its latest interrogation on September 24, 1962, more than 7 months after launch. It was, however, the failure of a tape recorder in the record mode after 7 months which caused the data acquisition to be stopped.

While the facts show that the recorders survived more than six months operating continuously in an orbital environment, there is still a question of shelf life. Although only a single example, the IR package flown in TIROS IV was built at the same time as that for the TIROS II and survived on a shelf for more than a year before launch, plus seven months in orbit.

THE NIMBUS RECORDER

The successful performance of the recorders in the TIROS II, III, and IV series influenced the decision to use this same design for the PCM telemetry system of the yet to be launched Nimbus spacecraft* (Figure 9). The Nimbus Spacecraft is an earth oriented 800 pound advanced Meteorological satellite to be launched in a near polar orbit by a Thor-Agena B rocket. It consists of an active controls system, to maintain the satellite orientation with the earth, a solar array, and a doughnut-type sensory ring structure (Figure 10) to house instrumentation which can be packaged in one of a selected standard sizes of modular boxes. As a result, the basic recorder had to be packaged in a box type package of 8 inches by 6 inches by 6½ inches. In addition a momentum compensation flywheel system had to be added to alleviate momentum effects induced into the spacecraft when flywheel speeds changed during record to playback modes.

*NASA TN D-1422 "The Nimbus Spacecraft and its Communication System as of September 1961"—R. A. Stampfl.

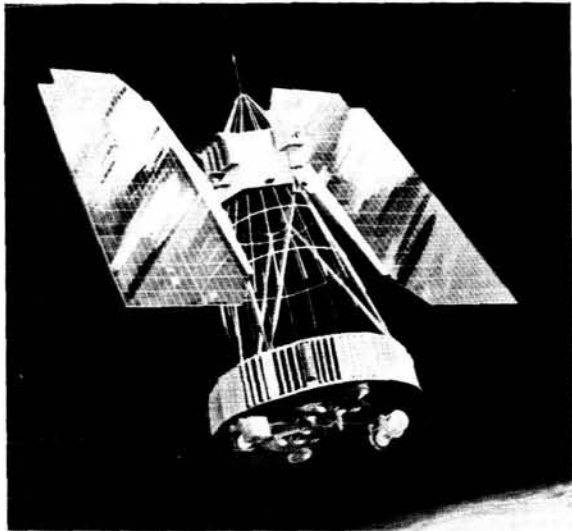


FIGURE 9.—Nimbus Satellite

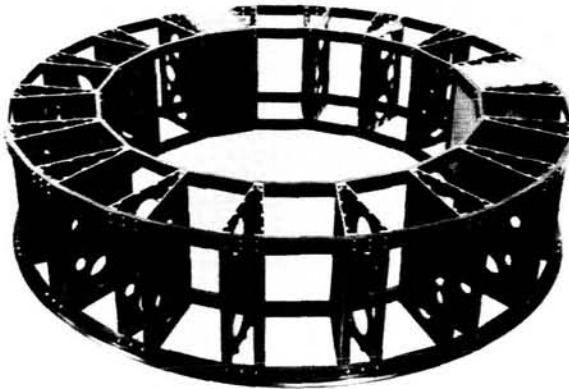


FIGURE 10.—Nimbus sensory ring.

Most design changes were considered straight forward and all future effort was directed toward building recorders in the box form factor. However, serious vibration problems arose when packaging the design on a square plate into a box and considerable effort was expended in solving the vibration problems by means of an isolation system without major redesign of the transport form factor.

Upon qualification of the design the endless loop recorder, (Figure 11) packaged into a box-structure, was applicable for the Nimbus PCM telemetry system and the TIROS IR system. This latest design was flown on TIROS VII, launched June 19, 1963, more than 1 year ago is

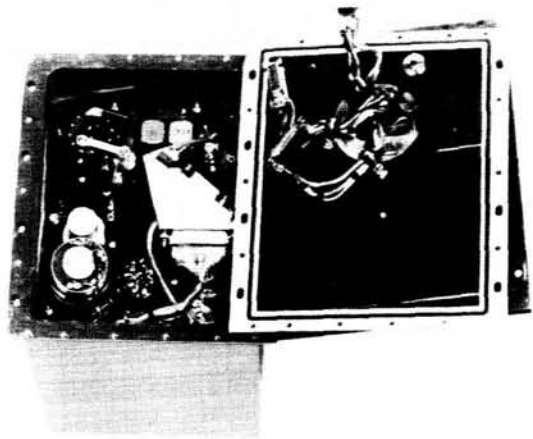
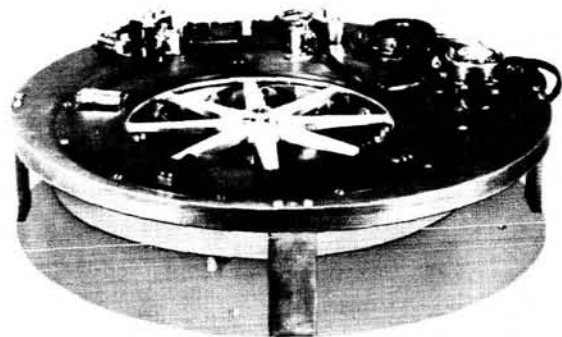


FIGURE 11.—Nimbus tape recorder.

considered the best IR system flown by the experimenters, and still operating as of this writing.

ADVANCED DEVELOPMENT

As a result of the Meteorological satellite programs, the technology in endless loop recorder design has increased significantly. Although flight experience has been limited to endless loop recorders capable of up to 300 feet of tape storage, significant breakthrough have been made in the design and fabrication of large endless loop recorders* capable of containing 1200 feet of $\frac{1}{4}$ and $\frac{1}{2}$ inch wide tapes (Figures 12 and 13).

FIGURE 12.— $\frac{1}{4}$ "—1200 ft advanced tape recorder.

*NASA TN D-2316 "Development of a 1200 Foot Endless-Loop Tape Transport for Satellite Applications"—K. W. Stark.

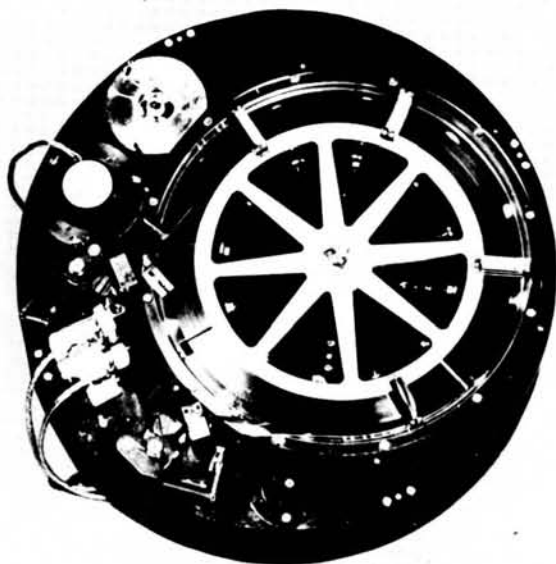


FIGURE 13.— $\frac{1}{2}$ "—1200 ft advanced tape recorder.

Interplanetary capsules and large satellites, where data to be stored can be considerable, require large capacity endless loop tape recorders. Endless loop recorders with storage capacities of 300 feet of tape are in some designs no longer satisfactory. In addition to requiring 1200 feet of tape on an endless loop cartridge, a rigid set of design specifications must be employed to assure long life and reliability. New design considerations, with respect to obtaining a functionally operational recorder, are required to achieve the flutter, low power, long life and low weight limitations. As a recorder's storage requirement is increased, many new design problems must be taken into consideration.

One design parameter that must be investigated is that of tape life. Included in this term "tape life" are the deterioration of magnetic properties, abrasion loss of the lubricant, edge wear and the basic characteristics of the mylar base material and oxide binder. Each one of these considerations is affected by one or more of the following: temperatures, tape velocities, mean pack diameter (which determines the slippage between outer and inner tape diameters) and total recorder operational lifetime required. Because of these tentative trouble areas, it is not just a simple matter of making a larger cartridge to accommodate the larger tape requirement.

Another area to consider in the design is the selection of the proper motor. Again, this is not simply utilizing existing motors in other tape recorders. Power requirements are different and motors must be matched to the recorder to prevent undesirable hunting which produces detrimental flutter and wow.

The method in which capstans are made (e.g., bearing preload techniques, machining tolerances, material and assembly and mounting arrangements) determines to a great extent how performance will be affected such as flutter and wow, skew and amplitude variations.

Of equal importance in the major design considerations is environmental testing. A 1200 foot endless loop recorder is much more sensitive to temperature and vibration testing than a small compact 300 foot endless loop recorder. The large loop tends to act very erratic under random and sinusoidal vibration tests if left in a free condition in a cartridge.

The 1200 foot endless loop tape recorder is a lightweight, low power, multi or single speed system. The overall external dimensions are $13\frac{3}{8}$ inches in diameter, $3\frac{1}{2}$ inches in depth, and weighs 10 pounds.

A maximum mechanical power of 1.29 watts is required to drive the transport at 30 inches per second at 0°C . At $3\frac{3}{4}$ inches per second tape speed, only .083 watts of mechanical power are required at 0°C to operate the transport. These values vary slightly for the $\frac{1}{2}$ " widetape and the type of tape used. Speed reductions are obtained through the use of accurately machined pulleys and seamless polyester film belts. Although the optimum motor has not been obtained as of this writing, wow and flutter measurements have been obtained at 1.14 percent P-P from 0-1000 cps bandwidth.

The item which required the longest development time is the tape cartridge. At the time, the only cartridges in common usage had tape capacities from 200 to 300 feet. To attempt to design a cartridge for 1200 feet of tape, various design considerations must be taken into account such as tape life, as defined earlier in the introduction, transport operational life, performance requirements, and environmental testing. A 9 month life was a design requirement which allowed for approximately 3 months of satellite testing on the

ground and 6 months of operation in orbit. The life requirement and environmental testing was most severe on the tape.

The tape cartridge had to be designed so that it would effectively allow 1200 feet of magnetic tape to function properly. Such parameters as power, flutter, signal to noise ratio, and operation at two speed record and playback had to be taken into consideration in addition to the life requirement. The 9 months of continuous operation necessitated that the cartridge design must impose a minimum wear condition on the tape. This meant that the tape must not suffer deterioration of the oxide, lubricant and base material to the extent where its operation would be detrimental to the system as a whole.

The first approach to a cartridge design was to obtain an optimum mean diameter of the cartridge with minimum pack thickness. It is important to have a narrow pack thickness to minimize tape wear induced by the angular velocity difference between inner and outer tape layers. In addition, the mean tape diameter should be within reasonable limits to prevent an excessively heavy and large recorder design. From a curve of mean tape pack diameter vs pack thickness for various lengths of tape an optimum mean diameter of 8 inches was chosen for a pack thickness of .8 inches for the 1200 feet of tape.

After the basic size of the cartridge was determined, there were two possibilities for a design. One was to expand the TIROS type into a larger size. As a result of conducting some preliminary tests, it was found that flutter was appreciably reduced when the tape was supported on rollers during operation. Thus the second design consisted of utilizing rollers to support the tape in the cartridge.

A large TIROS type cartridge (Figure 14) was built and installed in the breadboard for testing. Performance tests at various speeds were run. Although the cartridge visibly functioned well, there were severe tape pulsations which increased with tape speed. As a result, this type cartridge was eliminated from further consideration for two reasons. First, the pulsations result from a condition of excessive tape friction between each layer and between the tape edges and reel flange. This will cause accelerated tape wear and be detrimental to overall recorder reliability. Secondly,

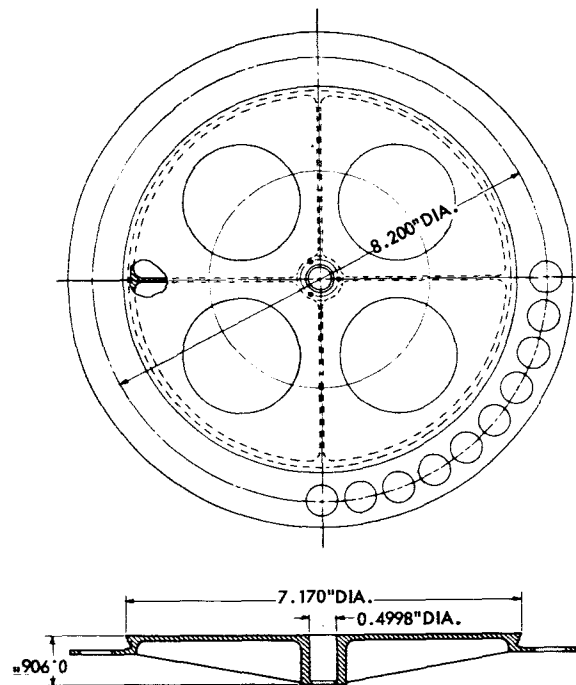


FIGURE 14.—Tape reel and cross-section.

because of the magnitude of the pulsations, filtering could not be easily accomplished with simple devices. This would result in large periodic flutter and wow disturbances.

At this point, the roller-reel tape cartridge was put into the breadboard (Figure 15) and testing proceeded. The main problem area developed

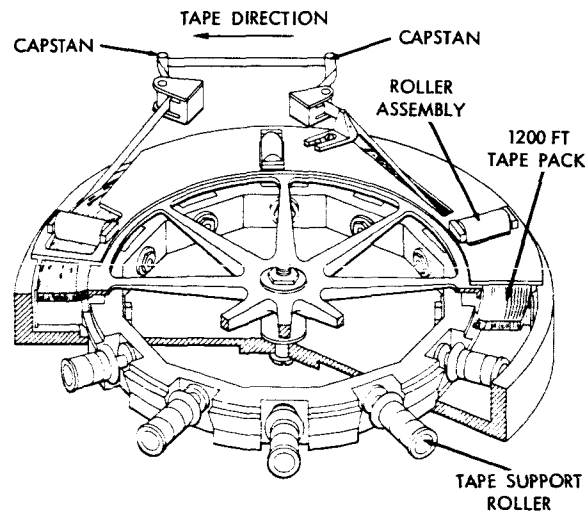


FIGURE 15.—Tape cartridge.

with the tape support rollers. The original rollers were straight cylinders. During testing, it was observed that upon acceleration the tape would form a loop on the return side which would not pull in and which remained at a constant length. This situation resulted because there was no friction between the outer layers of tape and the tape rollers as existed between the outer layers of tape and the flange of the larger TIROS reel. With the TIROS type cartridge, the flange which supports the tape has higher linear velocities at every point than the tape itself; therefore, there is a constant frictional force pulling the tape into the cartridge on the return side. This effect is absent with the rollers. To correct this situation, a series of tests were conducted which resulted in two cylindrical steps being added to the straight roller section. It was found that the RPM of the rollers was determined at the smallest diameter which provides a higher circumferential velocity at the two larger steps. The resultant relative velocity between the roller and tape edge was sufficient to prevent the tape from forming and maintaining a loop on the return side of the cartridge.

With this problem solved, the cartridge functioned very uniformly; and it was decided that further recorder development would be centered around this cartridge.

The cartridge consists of a tape reel, 12 tape support rollers, a roller mounting ring, a tape guide plate, a tapered guide, cover plate rollers, and, of course, 1200 feet of lubricated mylar base magnetic tape.

Although the problem of obtaining a functionally operating endless loop tape cartridge for 1200 feet of tape was solved, there still remained an area which required investigation. This area concerned the magnetic tape itself. The tape which was used for smaller endless loop cartridges on the TIROS and Nimbus recorders was not satisfactory for use on the larger cartridge. Initial testing with this tape revealed deficiencies in the mechanical stability of the tape during temperature testing. Primarily, there were two temperature problems, (a) the oxide binder became soft at 70°V, and (b) the tape had a very high shrinkage rate.

Temperature tests were run on the 1200 foot endless loop cartridge using Minnesota Mining

and Manufacturing's LF 1220 tape at 30 inches per second and temperatures of 0°C, 25°C., and 60°C.

At 0°C, the transport operated if the recorder was first placed in a plastic bag, flushed with dry air and sealed before being brought down in temperature. This was necessary to prevent the cartridge from jamming. This occurrence is attributed to the possibility of ice condensing on the transport and tape.

At 25°C the cartridge performed satisfactorily. However, at 60°C, problems arose where the tape exhibited properties of excessive inter-layer friction. This interlayer interference caused the cartridge to tighten up and eventually jam. A series of tests were conducted to determine whether the oxide, lubricant, or both were causing the trouble. At about 70°C, the sliding friction of oxide against oxide increased; but the sliding friction of lubricant against lubricant remained constant. At 100°C, the same effect was noted. This fact was brought to the attention of the tape manufacturer. The manufacturer conducted his own tests and found that the oxide binder was stable only up to about 80°C.

Although the recorder temperature test is run at 60°C, it is possible that the innerlayer tape angular velocity difference causes a localized heating effect raising the tape surface temperature above 70-80°C.

As a result of tests, the manufacturer provided new samples of tape (LR 1259 and 8943) which were considered identical in magnetic and lubrication properties to the LR 1220 but with a higher temperature binder on the oxide side. The LR 1259 was tested at 60°C in the cartridge and operated for a considerable length of time over and above the LR 1220 at the same temperature indicating the new binder was the solution to this problem.

However, at this point the second problem arose. When using the LR 1220 tape, the cartridge never operated long enough at the high temperature to observe other characteristics.

As the LR 1259 continued to run at 60°C, it was observed that the tape pack was gradually tightening up. Finally, the point was reached where the pack became so tight that the cartridge jammed. This tightening process was attributed to the shrinkage of mylar at high temperatures. Although this characteristic of mylar is well

known, the exact rates of shrinkage were not as well known for various temperatures. Tests were run on LR 1220, LR 1259, 8943, and a tensilized tape sample to determine their shrinkage rates.

The percent shrinkage for LR 1259 at 60°C is .13 percent; however, if localized heating raised the temperature to 80°C, the shrinkage becomes .34 percent. For 1200 feet of tape and .34 percent shrinkage, the tape becomes shorter by 49 inches thus increasing tape tension excessively.

The percent shrinkage for LR 1220 at 60°C is .075 percent and at 80°C it becomes .19 percent. For 1200 feet of tape with .19 percent shrinkage, the tape will lose 27.4 inches or about one-half that of LR 1259. The 8943 tape had a shrinkage rate of .13 percent. Base mylar (.001" thick) had a shrinkage rate of .284 percent.

Because these results showed large shrinkage rates and the variation of rates with tape type,

it was decided to preshrink the LR 1259 tape at 100°C for one hour and then repeat the test in the cartridge at 60°C.

This was done and the tape ran continuously at 60°C without any indication of tightening or jamming; however, during life testing a wear problem became apparent with the lubricant. Again 3M studied the problem and supplied NASA with a new sample which passed the required 9 month accelerated life test. There were not enough samples to determine the recording characteristics although it was specified that the oxide should have the same magnetic properties as the tape it replaced. This new tape has a designation of LR 1353.

The following performance measurements were taken at room temperature with the ¼" and ½" 1200 foot preprototype recorders.

	¼" Preprototype		½" Preprototype
Length of tape (feet).....	1200	1200	1200
Temperature (°C).....	25	25	25
Record Speed (IPS).....	3¾	30	30
Reproduce Speed (ISP).....	30	30	30
Flutter (% P-P) 0-1000 cps...	1.14	.64	.7
Power to Drive.....	.0717 record	.897	.88
Entire System (watts).....	.897 reproduce		
Signal to Noise Ratio (db)....	40		26.6
Tape Type.....	LR 1220	LR 1220	LR 1353

This transport has survived operation at 0°C and 60°C. However, during vibration testing, a problem of tape spew and tightening arose. This problem was solved by snubbing the tape, against its edges, so that the layers of tape could not move during vibrations. In addition the recorder was mounted on vibration isolators.

The following specifications are the vibration levels to which the recorder was tested:

- a. Sinusoidal
 - 10 g's (limit to ¼ inch single amplitude)
 - 5 to 2000 cps
 - 18 minute duration
- b. Random
 - 20 g's RPS
 - 20 to 2000 cps
 - .2 g²/cps spectral density
 - 4 minute duration

A third type endless loop cartridge (Figure 16) utilizing 1200 feet of ¼ inch wide tape was built and demonstrated good performance characteristics with respect to power consumption and smoothness of tape motion. The uniqueness of this cartridge is that the tape can be stored in a recorder approximately ½ the recorder diameters of those previously mentioned. However, this cartridge will have to be incorporated into a transport system to obtain more complete performance and operating specifications.

A complete tabulation of recorder performance for all meteorological satellite endless loop tape recorders is shown in Figure 17.

CONCLUSION

The ½" 1200 foot recorder has shown the greatest promise for solving immediate needs for large

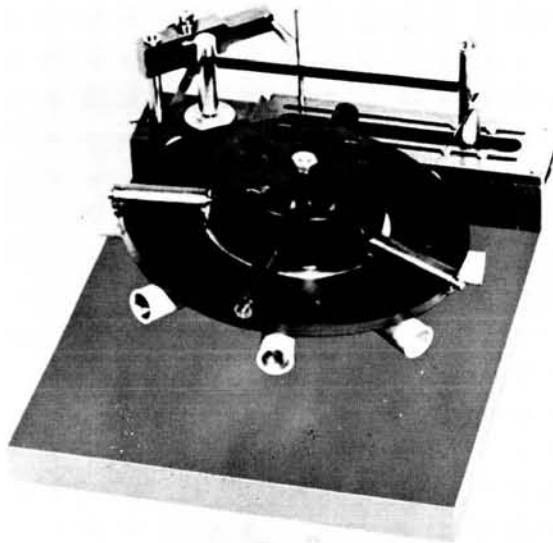


FIGURE 16.— $\frac{1}{4}$ "—1200 ft compact advanced tape recorder.

	VANGUARD II	TIROS	NIMBUS	ADVANCED RECORDER
WEIGHT (POUNDS)	1.5	4.0	8	10
SIZE (INCHES)	5.5 DIA	6 DIA	8X6X6.5	13.375 DIA
HEIGHT (INCHES)	3	2.3	—	3.5
TAPE SPEED (IN./SEC)	RECORD	.4	.4	33
	REPRODUCE	15	12	12
POWER (WATTS)	RECORD	.09	3	1
	REPRODUCE	1	1	4
FLUTTER % (PEAK TO PEAK) FROM 0-1000 CPS	6	2.5	2.5	.7
ANALOG OR DIGITAL	A	A	A&D	A
AMPLITUDE MODULATION (% PEAK TO PEAK)	—	7	7	4.0
PACKING DENSITY (BITS PER INCH)	—	—	250	—
TAPE LENGTH (FEET)	75	200	200	1200
MAXIMUM LIFE OBTAINED	2 WEEKS	17 MOS	13 MOS	—

FIGURE 17.—Performance characteristics of various meteorological satellite endless loop recorders at 25°C.

data storage. On the basis of its past performance tests this recorder was recently selected to be tested in recording and reproducing T.V. pictures utilizing the Nimbus Meteorological Satellite Camera system. Having passed this test successfully it is planned to use large endless loop recorders in future meteorological satellite programs.

DEVELOPMENT OF A BRUSHLESS DC MOTOR FOR SATELLITE APPLICATION*

PHILIP A. STUDER

The Goddard Space Flight Center is sponsoring a development program which has developed a brushless motor with true dc motor characteristics. Photo-optical detectors and transistorized switching duplicate the functions of the conventional commutator without physical contact. The 3-watt prototype motors have a starting torque of 2.7 oz-in and achieve fifty percent efficiency (.67 oz-in) at 3000 rpm, the speed chosen to reduce bearing and gearing problems anticipated in operation in outer space.

The motor performance compares favorably with conventional dc motors and should significantly improve the efficiency of present systems employing ac motors and inverters. An exceptionally fast response time (<.010 sec) and control at the milliwatt power level will be of value in servo applications. High starting torque and system efficiency recommend the motor for consideration whenever a source of mechanical power is required and brush life or arcing have previously ruled the dc motor out of consideration.

INTRODUCTION

The Goddard Space Flight Center is sponsoring a program which has developed a "brushless" motor (Figure 1) of true dc motor characteristics. This motor designed specifically for spacecraft operation, promises to significantly increase the efficiency of power conversion from electrical to mechanical energy. The motor has other assets which further tend to make it a very attractive component for a wide variety of applications, such as servo mechanisms, tape recorder drives, reaction wheels, etc.

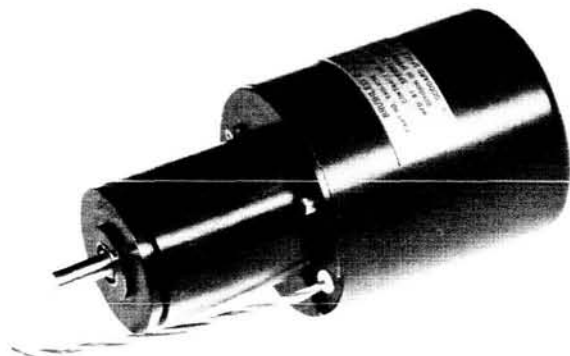


FIGURE 1.—The prototype motor.

HISTORICAL REVIEW

One of the basic mechanical problems is motion. The most effective and versatile type of motion is provided by rotating machinery. Since the readily available and storable form of energy aboard spacecraft is electrical, motors are a key element in all but the most elementary form of spacecraft.

Motors, in general, are chosen according to the type of electrical power available—either ac or dc. Spacecraft power sources are almost universally dc because of the nature of the devices—i.e., solar cells, thermoelectric converters—and because electrical energy is storable in this form. Hence, dc motors should be ideally suited for spacecraft use. But, the barrier to the use of dc motors in space has been the problem of brush life in a vacuum environment.

Another reason for preferring dc motors is their high efficiency. Electrical power is valuable; and, at the present time, energy collection, storage, and the associated conversion, regulation, and appendages represent between 20 and 40 percent of the total weight of satellites now on the drawing boards and under construction. The higher figure is generally related to smaller types and the smaller percentage to larger, oriented devices. These systems have weights ranging between $\frac{3}{4}$ and

*Published as NASA Technical Note D-2108, February 1964.

13/4 pounds per watt and these figures do not include the fact that significant portions of the structural and control devices are required specifically for the paddles, battery support members, etc. Any savings that can be made in reducing power conversion losses represent sizable decreases in the overall payload weight and booster requirements. Conversely, with a given amount of available power, a satellite's capabilities can be increased dramatically if the power needed to perform a given function is reduced.

With these facts in mind, the number one problem heretofore plaguing the use of dc motors in space was attacked in early 1962. This was that brush life of conventional carbon brush commutators was measured in minutes when operated in a vacuum environment. Development of electronic commutators in the industry was rumored, but a preliminary search found none commercially available.

DEVELOPMENTAL PROGRAM

Proposals were solicited from the industry for the development and prototype manufacture of a brushless motor with true dc motor characteristics. A contract was awarded to Sperry Farragut Company whose proposal indicated a firm groundwork had been laid. Basic to the problem were two functions:

1. Detecting the position of the rotor and,
2. Driving switches of sufficient power handling capability from a relatively weak detected signal.

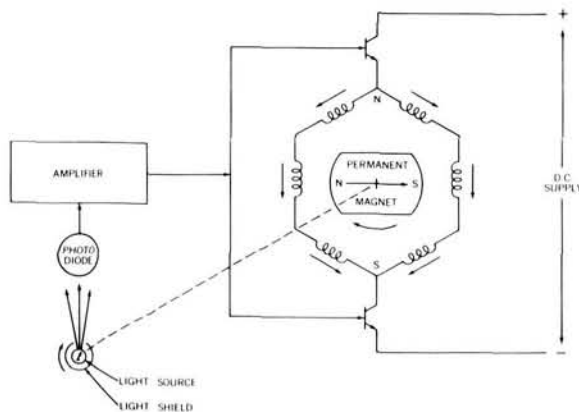


FIGURE 2.—Schematic diagram of the motor.

The contractor's approach had been to use light sensitive photodiodes and a rotating shutter to determine the rotor position, and transistorized amplification and power switching to commutate (or direct) the input power.

The principle of operation of this motor is identical to the conventional dc motor (see Figure 2). A torque is produced when two magnetic fields are established with a relative angle between them. This torque reaches a maximum when the relative angle is ninety degrees. It is the function of the commutator to apply the external power to the motor windings in such a manner as to maintain this angle for all positions of the rotating member.

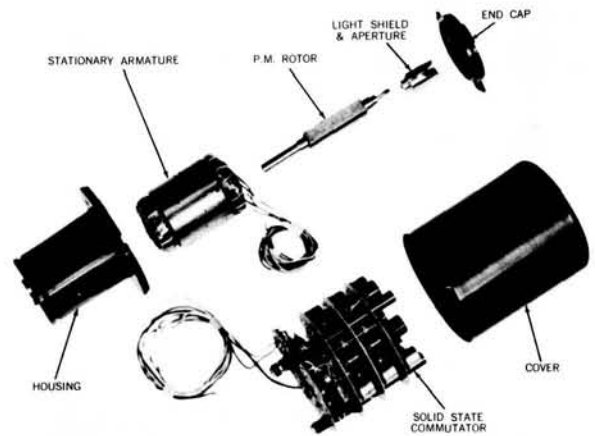


FIGURE 3.—The motor parts.

This motor is inverted in comparison to a conventional dc motor, in that the rotor is a permanent magnet and the stator is wound, as may be noted in Figure 3. Attached to the rotor is a light shield which serves to direct the light beam of the stationary lamp to a particular photo-diode as it rotates. When illuminated, the photo-diode conducts a current which is amplified and directed to the bases of a pair of switching transistors. These apply the power supply voltage to the proper points on the ring type stator winding. There are six identical sets of photo-diodes and transistor switches, each turned on in sequence for 60°/revolution. Thus, this commutation action is identical to that provided by the conventional carbon brush and segmented copper ring, but performed without physical contact of moving parts.

In the course of this development, several alternate methods were investigated as new solid state devices became commercially available (high-impedance thin-film Hall-effect devices, magneto-sensitive resistors, and photosensitive silicon controlled rectifiers). Hall-effect devices have the distinct advantage of being bilateral devices, capable of both positive and negative outputs, thereby reducing by one-half the number of detectors required. They are also solid state devices with no wear-out mechanism inherent in their operation. The disadvantages, shared by the magnetoresistive elements, include the distributed area of sensitivity, the relatively broad slope of the flux change possible, and the weight of the magnetic assembly required. The photosensitive silicon controlled rectifiers are, at first glance, capable of yielding the ultimate in simplicity due to their ability of combining the detecting and switching functions into one unit. Problems associated with their application hinge on the turn-off circuitry and the fact that a light source was still required.

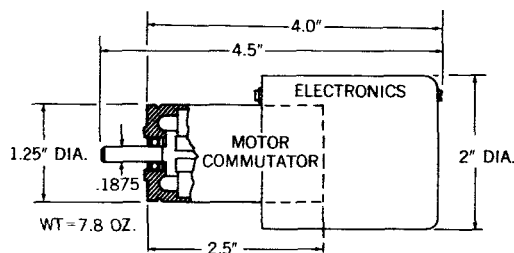
While none of these approaches presented insoluble problems, the experimental evidence continued to favor the photo-optical transistorized commutator. Several developments prompted a decision to continue in this direction; among them were the rather surprisingly good preliminary results of light source environmental tests, and a circuitry of demonstrated capability, light and small enough to satisfy the packaging goals for prototype hardware. Low voltage tungsten filament lamps survived vibration levels of 50 g's and shocks of 130 g's on a hard mount.

In spite of the fact that supplier delays prevented experimenting with some of the newer magnetic alloys; the motor design meets the original goal of 50 percent efficiency, the system's outstanding attribute. Coupled with this is the fact that this design, due to the small rotor dimensions ($\frac{3}{8}$ " diameter), exhibits an exceptional torque to inertia ratio, giving it superior servo motor characteristics.

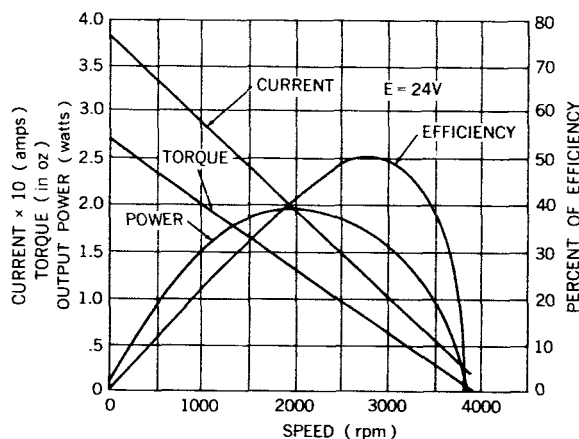
THE PROTOTYPE MOTOR

The specific motor design as developed under the present contract has the following performance specifications (Figure 4):

Power Input 3 watts at 24 v dc
 Torque .67 oz-in at 3000 rpm
 Torque, Stall 2.7 oz-in at 10.2 watts
 No Load Speed 3900 rpm



(a) Motor outline



(b) Performance Characteristics

FIGURE 4.—Performance curves.

Although this motor was designed for a specific set of conditions, the design concepts are versatile and, in fact, higher efficiencies have been obtained at higher operating speeds. Motor size, power, and speed can be scaled up or down with a fair degree of confidence.

The commutator utilizes a miniature tungsten lamp which is derated to give a theoretical filament life of one million hours. It produces an output of three thousandths candlepower while using .3 watt. The detecting device is a silicon photo device with a rise time of 1.5 micro-seconds and a fall time of 15 micro-seconds. Consequently, it is capable of far higher speeds than are presently being employed. The power switches are silicon power transistors rated to 3 amperes

with a leakage current well under 1 milliamp at 150°C.

The motor design employs a minimum diameter Alnico V rotor capable of handling a reasonable air gap clearance. The length of the motor was chosen to minimize the power loss in the "end turns." A high silicon steel was utilized for the stator on the basis of high saturation flux density. The stator is laminated to reduce core losses and the slots skewed to minimize cogging. Ball bearings, the only moving parts in physical contact, are preloaded to reduce runout of the titanium shaft. Heavy polyvinyl acetal insulation was specified for the motor windings.

At the design speed of 3000 rpm, ten percent of the power is used by the lamp and only six percent in the electronics—more than half of which is dissipated in the isolation resistors and not in the active devices. Winding resistance accounts for twenty-two percent and "speed" losses, which include bearings, air friction, and core losses in the iron, the remaining thirteen percent. Mechanical power available at the shaft is one-half the input power.

COMPARISON WITH OTHER SYSTEMS

The gap that this motor fills is best illustrated by looking over a list of typical commercially

Table 1

Typical Characteristics of 9 Types of Motors

Type	Torque at Rated Speed (oz-in)	Efficiency Percent at Rated Speed	Speed (rpm)	Power Input (watts)	Current (volts)	Torque Stall (oz-in)	Diameter (inches)	Length (inches)
GSFC-NASA Brushless D.C.	.67	50	3,000	3	24 dc	2.7	2.0	4.0
Brushless Induction	1.5	36	30,000	100	28 dc	1.5	1.5	
Brushless Induction	.8	22	5,100	14	28 dc	.4	3.0	2.5
A.C. Induction	2.5	55	10,400	36	115 dc	1.0	1.5	2.25
A.C. Hysteresis	.75	17	3,600	12	115 ac	.75	1.2	2.6
Perm. Mag. Servo	.35	5	725	4	115 ac	.73	1.1	1.5
Stepper	.20	1	200	3.8	27 dc (Chopped)	.20		
D.C. Servo	.21	34	13,500	7	28 dc		.8	1.0
D.C. Servo	19.0	44	3,500	96	12 dc	60.0	4.0	2.0
D.C. Motor	.7	53	17,000	23	27 dc	4.6	1.0	1.7
D.C. Motor	1.8	63	15,000	32	26 dc	9.7	1.25	2.5
D.C. Torquer	NA	--	Zero	3	28 dc	5.8	1.9	.5

available devices, just as a design engineer might do when looking for a device to meet a specific satellite requirement (Table 1).

Take a case, for purpose of illustration, that a shaft must turn at 30 rpm with a 40 ± 20 percent

oz-in load on a semicontinuous basis.

The ordinary dc motors with their tempting power efficiencies, must be bypassed due to the problem of brush life. Also bypassed is the stepper motor due to its very low efficiency.

The permanent magnet induction servo merits some consideration due to its low speed characteristic:

$$\frac{725 \text{ rpm}}{30 \text{ rpm}} \times .35 \text{ oz-in} = 8.5 \text{ oz-in (insufficient torque).}$$

The hysteresis motor looks somewhat better:

$$\frac{3600 \text{ rpm}}{30 \text{ rpm}} \times .75 \text{ oz-in} = 90 \text{ oz-in}$$

$$\begin{aligned} (120:1 \text{ reduction}) \times .75 \text{ gear train efficiency.} \\ \text{Output} = 67.5 \text{ oz-in} \end{aligned}$$

The 12 watt ac power requirement is rather large and with 20 percent losses anticipated in the inverter; this would require 15 watts dc battery power.

Induction motors are more efficient:

$$\frac{10,400 \text{ rpm}}{30 \text{ rpm}} \times 2.5 \text{ oz-in} = 870 \text{ oz-in}$$

$$\begin{aligned} (346:1 \text{ reduction}) \times .70 \text{ gear train efficiency,} \\ \text{Output} = 610 \text{ oz-in} \end{aligned}$$

which is an order of magnitude higher than necessary. However, a similar motor should be using one-tenth, the power, or 3.6 watts ac. This would be acceptable even though inverter losses would increase the battery supply requirement to 4.5 watts dc. This is still not a good solution on two counts, bearing and gear life are seriously degraded at these speeds, and starting torque is too low (0.42 oz-in). Now investigating a brushless inductive motor, we see that

$$\frac{5,100 \text{ rpm}}{30 \text{ rpm}} \times .8 \text{ oz-in} = 136 \text{ oz-in;}$$

$$\begin{aligned} (170:1 \text{ reduction}) \times .70 \text{ gear train efficiency.} \\ \text{Output} = 96 \text{ oz-in} \end{aligned}$$

Again 14 watts dc power is high.

Another motor of the same type would cut this to 8.5 watts but only by operating at a speed which is out of the question.

Consider this brushless dc motor:

$$\frac{3000 \text{ rpm}}{30 \text{ rpm}} \times .67 \text{ oz-in} = 67 \text{ oz-in};$$

$$(100:1 \text{ ratio}) \times .75 \text{ gear train efficiency};$$

$$\text{Output} = \frac{67}{50} \text{ oz-in}$$

$$\text{Power} = 3 \text{ watts};$$

$$\text{Starting Torque } 240 \text{ oz-in} \times .75 = 180 \text{ oz-in.}$$

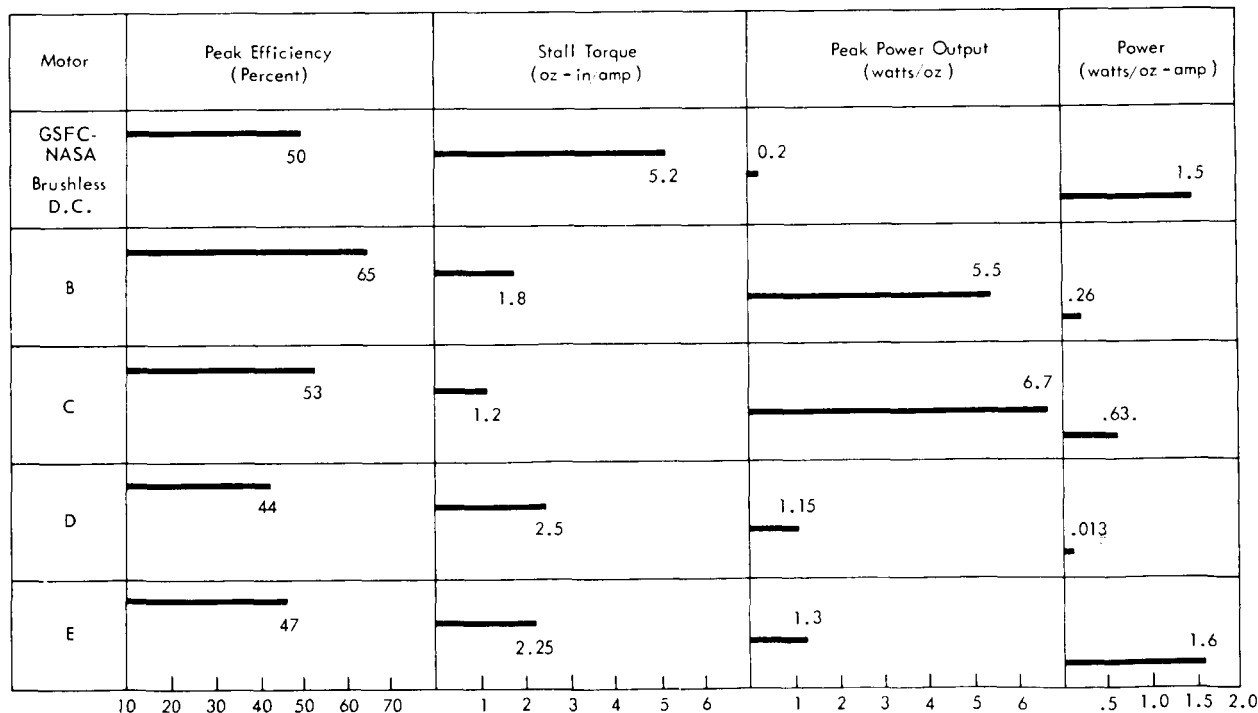
The relatively low speed (3,000 rpm) alleviates the bearing and gearing problems and requires a smaller gear reducer. No additional weight or power loss is involved since it operates directly off the dc battery supply without an inverter.

Another worthwhile comparison is to see where this motor stands with respect to conventional

permanent magnet dc motors. By referring to Table 2 it can be seen that its efficiency is comparable to some good quality motors but 15 percent under the best. Stall torque per ampere exceeds the conventional motor by a factor of two. In the category of peak power output per ounce, a large variation is found, with this motor an order

Table 2

Characteristics of Several dc Motors



of magnitude below a typical motor. This is the result of the effort to design for a low power input capability and the low speed operation desired. But, when this value is divided by the current used, the figures in the next column are obtained, and where weight and power are both limited, as in spacecraft, this motor again appears at the top of the list. It should be remembered that these latter comparisons are made to motors not

suitable for spacecraft operation due to the brush life problem.

Early in the test program the fast response exhibited by the motor was noted. This led to a comparison with typical servo motor characteristics and it was found that it could easily meet SAE and BuOrd specifications for a Size 18 servo even though the motor frame is only a Size 12. They are compared in Table 3.

Table 3

Comparison of this Motor and a BuOrd Servo

Quality	GSFC-NASA	A BuOrd Quality ac Servo
Starting Torque	2.7 oz-in	2.4 oz-in
Diameter	2.0 in	1.75 in
Weight	8.5 oz	12.2 oz
Theoretical Acc. at Stall	70,000 rad/sec ²	39,700 rad/sec ²
Input Watts at Stall	10.2 w	18.2 w
Input Watts at Speed	3.0 w	13.6 w
Servo Time Constant	< .010 sec	.0131 sec

It is felt that this motor concept may well stimulate great interest in the servomechanism field because, in addition to the excellent response time, it is controllable at a milliwatt power level. This is due to the nature of the electronic commutator which has self contained power amplification.

Numerous brushless motor developments have been published recently. Not all of these have the dc motor characteristics of high starting torque and high efficiency. This motor has a linear torque-current relationship with maximum torque developed at stall.

The current and (assuming a constant dc supply voltage) the power used by this type of motor is only what is required by the load. In many cases this is only 25 percent or less of the peak requirement. Consequently, the average power requirement of a system employing a motor of this type may be less than one quarter of what is used by a constant speed motor, and still have a two to one

factor of safety in starting ability under load. In addition, this motor has been specifically designed for low speed operation, easing the problem and gearing life in a vacuum. At higher speeds the motor is more efficient.

Now, not all points of comparison are this favorable. In the use of ac motors some capabilities are obtained which are not normally associated with dc machines. Number one is the inherent synchronous operation, another is the relatively well developed technology in inverter design which over the years has produced a product of known capability and reliability.

RESULTS

To summarize the present status of the brushless dc motor program:

1. Four prototype motors of fine capability specifically designed for low power spacecraft applications have been designed and built;

2. Sufficient knowledge has been developed regarding the design parameters to be able to confidently design motors of similar quality for applications requiring widely different power levels.

Further, groundwork has been laid for the following development efforts:

1. Breadboard circuits have been operated demonstrating the feasibility of considerably reducing the complexity of the electronic circuitry;
2. Design calculations have been made which utilize some of the newer magnetic alloys and special winding techniques which potentially are capable of efficiencies in the range of 60 to 70 percent;
3. The possibility of obtaining synchronous speed control without losing the performance of the dc motor appears feasible; and

4. A two speed motor with two independent windings is being fabricated for a developmental meteorological satellite tape recorder.

CONCLUSIONS

It is felt that this effort is a step forward in spacecraft technology, namely in providing more efficient motive power in space. Further, it is believed that it has opened the door to new areas of worthwhile development, specifically: speed control of dc motors, simplifications yielding increased reliability and economy in the solid state commutation circuitry, and further increases in the efficiency of miniature motors.

Now that the fundamental limitation on the use of dc motors in spacecraft has been overcome, it is reasonable to expect their increasing application. The high starting torque and high system efficiency recommend it for consideration whenever a source of mechanical power is required.

A MEDIUM - DATA - RATE DIGITAL TELEMETRY SYSTEM*

MARJORIE R. TOWNSEND, PAUL M. FEINBERG, AND JOHN G. LESKO, JR.

Earth-orbiting satellites require on-board data storage. A suitable medium-data-rate digital telemetry system with 10^7 -bit storage is described. Since this storage is provided by a tape recorder, a unique technique for wow and flutter compensation has been developed. Hardware applicable to the radiation experiments on the Tiros and Nimbus meteorological satellites has been built and is currently being tested.

INTRODUCTION

In addition to television cameras for taking pictures of the earth's cloud cover, the Tiros series of meteorological satellites carries a radiation experiment.† The Nimbus series will carry a similar experiment. The earth's temperature is produced by self-emitted radiation and reflected solar radiation. Present medium resolution radiometers filter the combined radiations into five spectral regions. Their geometry is such that the signal has a $(\sin x)/x$ response, the first null occurring at 8 cps for Tiros and at 16 cps for Nimbus. The sampling rate required to reproduce this signal should therefore be greater than twice the maximum frequency response. In this system, each of the five radiometer channels is sampled $33\frac{1}{3}$ times per second.

DATA STORAGE

In a circular earth orbit, data storage on board the satellite is a difficult problem. The most efficient storage system is a tape recorder. In the old FM telemetry system, the limiting factor in terms of a signal-to-noise ratio was its wow and flutter. This limit was about 27 db. In the present system, where the data is digitized before it is recorded, it is anticipated that 40 db can be achieved; the limit will then be due to the radiometer itself. Error analyses, not theoretical but practical, which have been run in the laboratory

show that the bit-error probability due to the transmission link will be less than 0.1×10^{-7} . This can be achieved even on a recorder having the equivalent of 2 months' operation in orbit.

A brief description of the digital storage system is in order. The tape recorder can store 10^7 bits of data, adequate for an orbit greater than 100 minutes. It measures 6 by 8 by $6\frac{1}{2}$ in., weighs $10\frac{1}{2}$ lb, and uses less than 2 watts in the record mode and about 12 watts in the playback mode. The two-speed recorder (0.45 ips and 11.7 ips) contains up to 250 ft. of lubricated $\frac{1}{4}$ -in. tape, in a continuous-loop cartridge driven across an 8-track staggered digital record-playback head. Saturation recording is used, the direction of the flux alternating for "ones" and "zeroes", with the bits conservatively packed at 500 per in.

Eight record and eight playback amplifiers can require a considerable volume. In addition, head-lead length should be kept to a minimum to reduce noise pickup. This, then, seemed to be a critical area in which microelectronics could provide a considerable advantage. The result of a contract with CBS Laboratories to develop suitable amplifiers is a package 4 by 1.359 by 1 in. (Figure 1) containing ten record and ten playback amplifiers (two spares) which can be mounted directly on top of the tape recorder, accomplishing the triple purpose of minimizing weight, size, and head-lead length. Each head, which has a nominal impedance of 22 ohms, requires 3.5 ma of current. By deriving this from a constant current source, the power requirement is 10 mw per track during the record cycle.

*Published as *NASA Technical Note D-2315*, June 1964 and as *Goddard Space Flight Center Document X-650-63-194*, September 1963.

†Davis, J. F., Hanel, R. A., et al., "Telemetry Infrared Data from the Tiros Meteorological Satellites," *NASA Technical Note D-1293*, June 1962.



FIGURE 1.—Amplifier package.

SYSTEM DESIGN

The digital electronics package, which weighs about 8 lb, requires 1.5 watts in the record mode and 7.2 watts in the playback mode. This latter figure includes the power used by the transmitter. The total system power is 3.5 watts in record and about 20 watts for the 4 minute playback cycle. The way in which the data and synchronization signals are handled will be considered next. An eight track tape recorder lend itself well to a 7 bit signal recorded in parallel across the tape, the eighth track being used for a clock signal. Seven-bit digital conversion will provide an accuracy of 1 part in 128. The clock track, one of the middle tracks, will be recorded 90 degrees out of phase with the signals on the other seven tracks, so that its edges can be used to trigger the readout in the center of the other data bits.

The record rate chosen was 200 bits per sec on each of the eight tracks in nonreturn-to-zero (NRZ) format, each bit lasting 5 msec. The format (Figure 2) as the signals go on tape is as

follows: Frame sync, a signal composed of all "ones", is recorded laterally across the tape. This is followed by the 7 bit output of the analog-to-digital converter. The analog signal from the first of the five radiometer channels, varying between 0 and -6.4 volts, is converted with an accuracy of 1 part in 128. Similarly, the remaining four radiometer channels are converted sequentially to digital form and the resulting bits become words 3 through 6.

The seventh word will contain timing data for Nimbus or a subcommutated channel for Tiros. Coherent with the sampling rate, the Minitrack time code is generated in the Nimbus clock and presented to this system in NRZ format. This information is updated every second, each of its 100 sequential bits lasting 10 msec. Of these 100 bits, only 36 represent timing data. Hundreds of days, tens of days, units of days, tens of hours, units of hours, ten of minutes, units of minutes, tens of seconds, and units of seconds are each indicated by 4 bits. The remaining 64 bits are for synchronization and station identification data. During the 60 msec frame period, 6 of the Minitrack bits will be stored in a shift register. At the proper time for recording the timing word, a "zero" is added, and this plus the 6 Minitrack bits are recorded on tape. Obviously, the 6 bits do not maintain any specific pattern, but they can be unscrambled in a suitable demodulator.

In the Tiros system, the seventh word is reserved for housekeeping data and a low-data-rate experiment. With a data input channel in this slot every other frame, a piggyback experiment can be sampled 8 times per sec. Alternating with this is subframe sync, the temperature of the radiometer baseplate which is required for accurate data reduction, the temperature and pressure

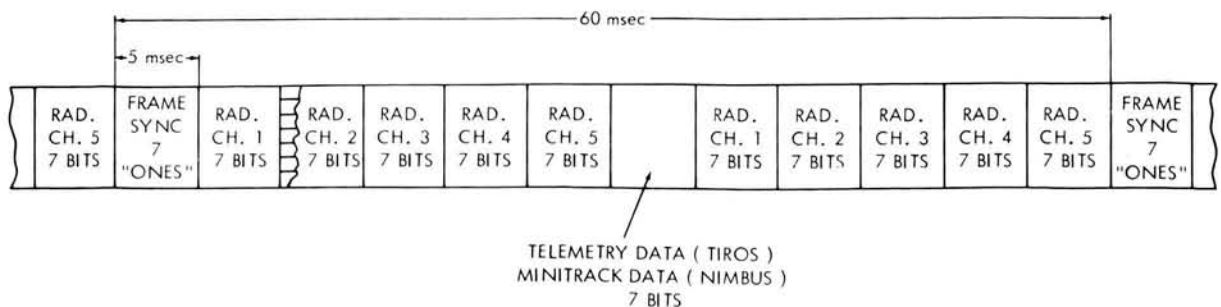


FIGURE 2.—Nimbus and Tiros digital main frame telemetry format.

inside the digital tape recorder, and a signal which correlates the radiation data with the TV pictures of the earth's cloud cover.

Words 8 through 12 are the next samples of the five radiometer channels in the same order as words 2 through 6. Thus, each radiometer channel is sampled once every 30 msec. This constitutes the frame, which is repeated over and over for the length of the orbit. The storage capability of the recorder is about 2 hours. Logic will prevent all "ones" from occurring at any time other than frame sync; i.e., 7 data "ones" (127) will never occur.

The Nimbus system block diagram in Figure 3 will help to clarify the system description. The

frame format gating network selects 7 "ones" for frame sync, 7 binary bits in parallel from the analog-to-digital converter whose input is determined by the analog gating network, or the "zero" and 6 bits from the Minitrack storage shift register. The selected 7 pulses, each 5 msec long, either "ones" or "zeros," are applied to seven record amplifiers and recorded on the tape. On the clock track is recorded a 100-cps square wave, equivalent to alternating "ones" and "zeros" at the 200 bit per sec rate. The change of state of these pulses is delayed by 2.5 msec (half a pulse-width) in relation to that on the other seven tracks. The purpose of this, as described before, is to use the edges of the clock pulses to sample the

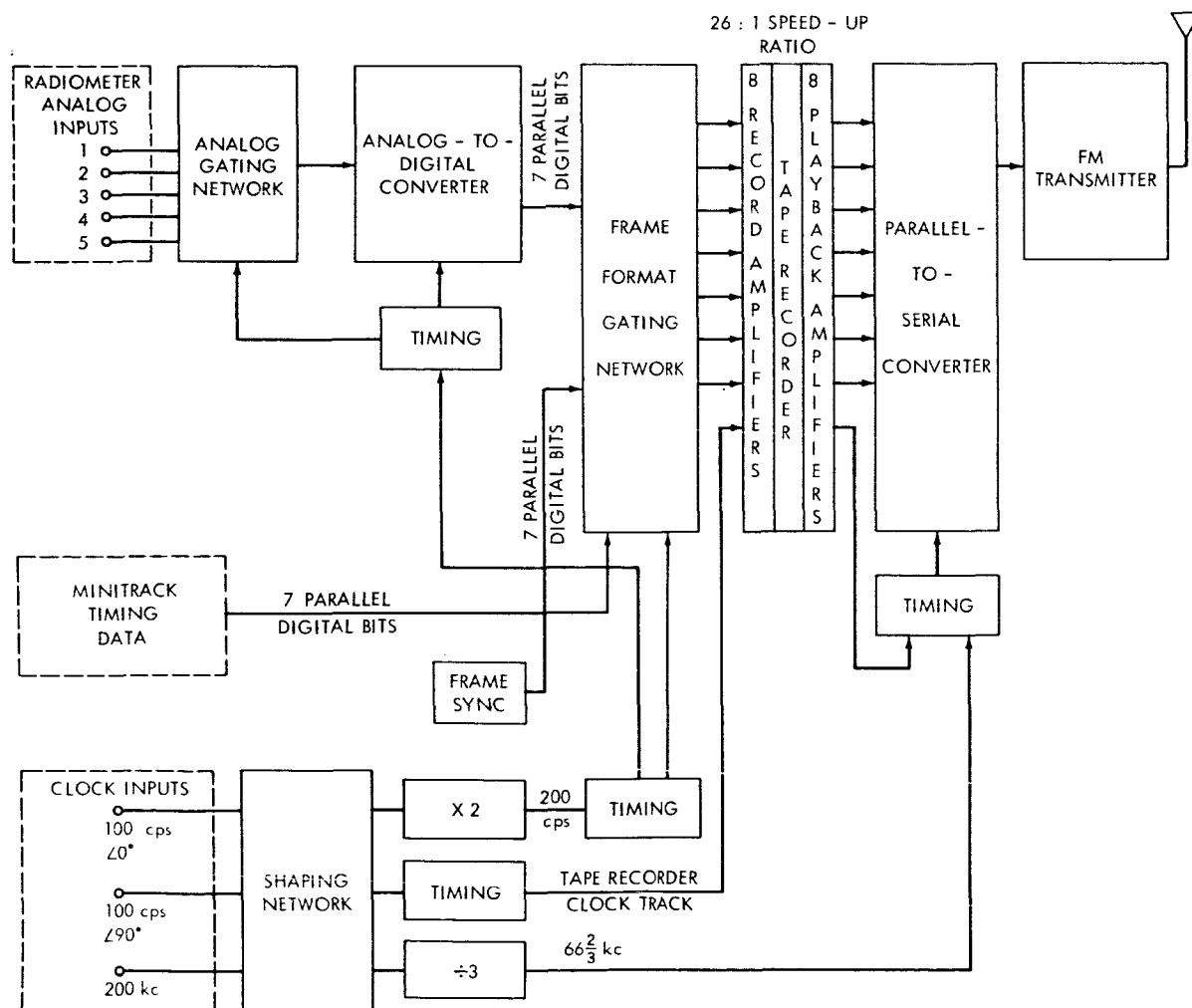


FIGURE 3.—Nimbus medium resolution infrared-pulse code modulation (MRIR-PCM) system.

format of the frame is completed. A "one" for word sync is added to the beginning of each word in the parallel-to-serial converter as the data is strobed out of the storage flip-flops. A stable frequency from the Nimbus clock ($66\frac{2}{3}$ kc), or from a crystal-controlled oscillator for Tiros, is used for strobing these flip-flops so that the transmitted bit rate will be very stable. If no new data is in the storage flip-flops after one strobe is completed, the transmitter will transmit "zeros" until new data is available. This will happen because the time required to strobe and transmit the seven tracks plus the added "one" will be less than the $5/26$ msec, 192μ sec, between words on the tape. Logic will prevent reading out of the tape while the output register is shifting. This system is designed to guarantee at least one "zero" to prevent the possible occurrence of eight "ones" in a row at a time other than frame sync, and will average about four (4.82) "zeroes" or twelve (12.82) bits per word. This guarantees a minimum bit rate of 5000 bits per sec. Thus, the transmitted serial train of NRZ pulses contains words of varying length (9 to 15 bits per word) to compensate for the changes in speed of the tape recorder and any variations in the recording rate. The cost of keeping the transmitted bit rate stable is the need for a 50 percent wider video bandwidth. However, the advantage of the stable transmitted frequency and much less complex spacecraft system make the trade-off worthwhile.

Table 1

Transmission Data

Datum	Value
Range at horizon	2270 st. mi.
Orbital altitude	600 st. mi.
Path attenuation at horizon	147 db
Path attenuation at zenith	135 db
Satellite antenna gain	-3 db
Ground antenna (85 ft dish) gain	29 db
Antenna temperature	290° K
Cabling losses	-2 db
Transmitter frequency	136 Mc
Transmitter power	1.75 watts
Receiver bandwidth	100 kc
Receiver noise figure	5 db
Signal-to-noise ratio at horizon	28 db
Signal-to-noise ratio at zenith	40 db

The serial train of pulses thus created frequency-shift keys a 1.75 watt FM transmitter. Because of the system design, the frequency can vary only between 5000 bits and $66\frac{2}{3}$ kilobits per sec. The deviation selected was ± 25 kc, providing a minimum modulation index of 0.8. Table 1 gives the parameters for a 600 mi. orbit. The results show a signal-to-noise ratio into the bit synchronizer of 28 db at the horizon, with 40 db when the satellite is at its zenith point. The use of the 60 ft dish for Tiros reduces these figures by 3 db if the same frequency is used, but the safety factor is adequate.

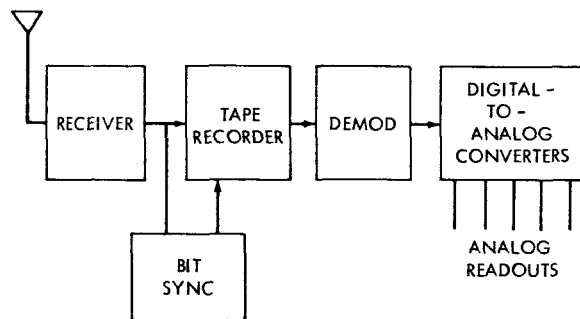


FIGURE 5.—Data acquisition station diagram.

GROUND DATA PROCESSING SYSTEM

After demodulation by the receivers (Figure 5) the digital signal, still in serial form, is recorded at the ground station on an instrumentation tape recorder. Simultaneously it is presented to a bit synchronizer whose output is recorded on another track of the same instrumentation tape recorder. A demodulator and digital-to-analog converters are used at the data acquisition station to judge the overall quality of the satellite data. Full-scale data processing must be handled by a computer because of the copious quantities of data.

Figure 6 illustrates the operation of the ground-station complex. After presentation to the bit synchronizer, the serial train of bits is converted to parallel format in an 8 bit shift register where a search is made for the unique frame sync word. When this is detected, a second search begins for the word sync bit which is used to direct the parallel output of the input shift register to the proper digital-to-analog converter flip-flop storage network. Eight bits are counted out for each word, if the "zero" bits which follow it are ignored, and

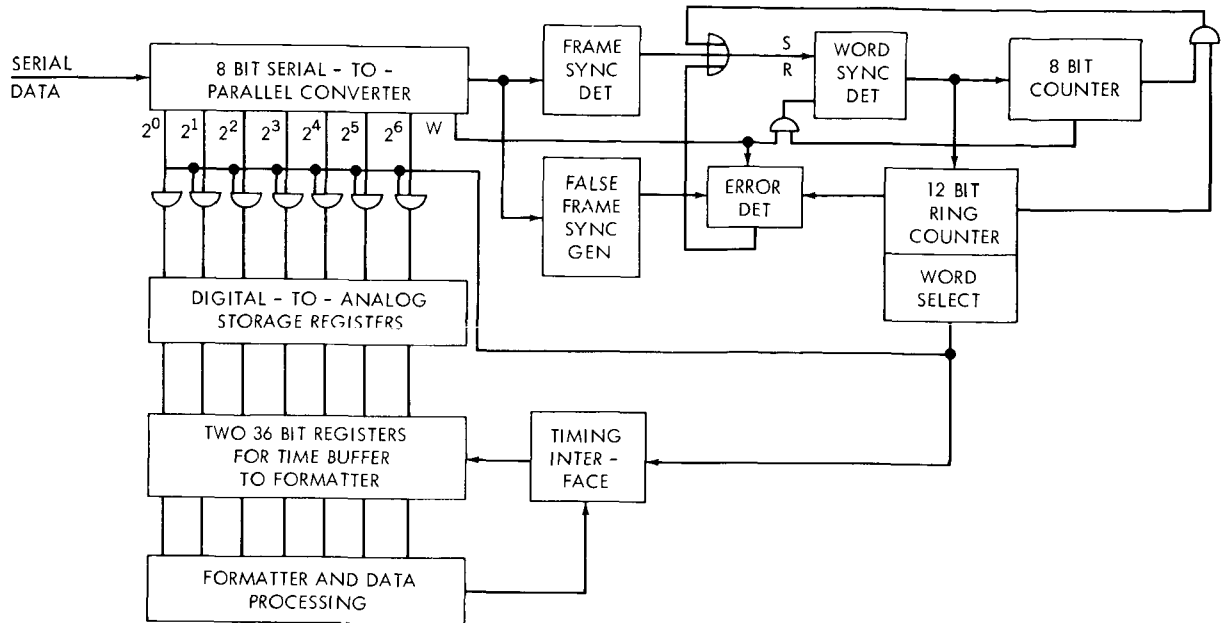


FIGURE 6.—Simplified ground station diagram.

twelve words are counted out between frames. This procedure continues through the whole twelve word frame, after which the search for frame sync begins again.

If frame sync should fail to appear within a selected time tolerance, a false frame sync is generated in the Tiros ground stations, since the loss of data words could mean a loss in time correlation and thus an error in determining the correct orbital geometry. To provide a surer selection of true frame sync, three frame syncs must be detected, without error, before a frame sync error is locked out. Three errors in one frame will put the system back into the search mode again, looking for true frame sync. An error is defined as finding that the ninth bit after word sync is a "one" instead of a "zero", or that frame sync does not occur within the proper time interval.

An output signal will occur after 64 false frame syncs to flag the occurrence of the splice or serious losses of data. Almost that many can be expected each orbit at the tape splice. Sixty-four frames are equivalent to 3.84 sec of data, a small percentage from a 90 minute orbit, but important in terms of maintaining time.

The individual words can be directed to digital-to-analog converters, so that the signal levels for

each channel may be analyzed by hand, or they may be directed to a formatter where they are combined into standard 36 bit IBM words and transferred to a digital tape which can be handled automatically by a computer. The format will be the same as that now used for the Tiros radiation experiment data. However, the housekeeping data which appear on the subcommutated channel will appear periodically as a special word on the tape and will be decommutated by the computer. With this data the calibration of the radiometer data can be adjusted automatically to conform to the changing satellite temperature. The error-output signal described above will generate a special code word to tell the computer about the occurrence of unacceptable data.

For Nimbus, absolute time in the form of the Minitrack time code will appear on the computer tape as a single word, which is updated every real-time second. There will be about 33 data words between each timing word, which will be flagged to the computer by a special code word.

The normal five channel data will be handled by shifting each 7 bit data word in parallel into a 36 bit shift register. After the register has five words or 35 bits, a "one" is added in the 36th bit and the IBM-formatted word is shifted out in groups of 6 bits. While this shifting operation is

occurring, the data words will be stored in a second identical 36 bit shift register which alternates with the first in accepting data and shifting them into a buffer. From the buffer the data are formatted onto the digital tape.

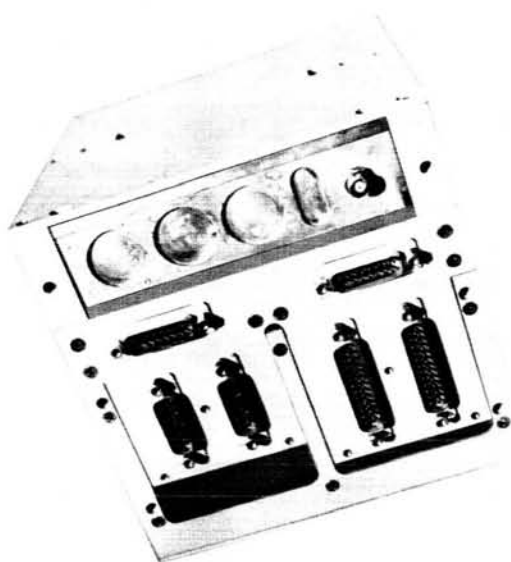


FIGURE 7.—Telemetry system.

HARDWARE

The hardware (Figure 7) for this telemetry system has been designed and built with the Series 51 integrated circuits developed by Texas Instruments. Flip-flops and different types of gates are fabricated in tiny individual modules from which the logic circuitry may be constructed. Using this type of circuitry should represent a significant advance in the reliability achieved in such a system. The record and playback amplifiers previously described and the logic for the analog-to-digital converter are being built in microelectronic form. Analog portions of the analog-to-digital converter and the analog gates are constructed from standard discrete components.

The converter (Figure 8) encodes using the successive approximation technique; it completes



FIGURE 8.—Analog-to-digital converter.

the encoding process in 320 μ sec. The total conversion error over a temperature range of 0° to 60°C is ± 10 mv. The total accuracy is ± 25 mv, half the least significant bit of 50 mv. The converter is designed so that a simple change would make it an 8 bit converter; its inherent accuracy provides that capability. The power required to operate the converter is 170 mw; of this, only 12 mw is required for the micropower logic circuitry, the remaining power being used for the conventional solid-state circuits.

CONCLUSION

This digital method of treating the data from the medium resolution radiation experiments planned for future Tiros and Nimbus satellites is expected to result in a fivefold improvement in accuracy over the FM analog system previously used. The integrated circuits are versatile, reliable, and ideally suited for present and future aerospace requirements.

III. SPACECRAFT AND SUBSYSTEMS

DIRECT MICROWAVE TO MICROWAVE TRANSPONDERS FOR COMMUNICATION SATELLITES*

WALTER K. ALLEN

INTRODUCTION

The primary function of a communication satellite involves the reception, amplification, and retransmission of information-bearing radio frequency signals between various ground terminals. The satellite equipment for performing this function is termed a "transponder". The transponder must also convert the signal to a different frequency at the output in order to prevent feedback interaction between the low level received power and the higher level transmitted power.

Present frequency band allocations for communications satellites require that the received, or uplink, frequency fall in the 600 Mc region and the retransmitted, or downlink, frequency fall in the 4000 Mc area.

TYPES OF TRANSPONDERS

The techniques used for obtaining frequency conversion and amplification in the transponder can be divided into two types: (1) the conventional or indirect method, where the received RF signal is down-converted to an intermediate frequency (IF) for amplification then up-converted to a new RF frequency for further amplification and re-transmission; and (2) the direct microwave to microwave method, where amplification and frequency conversion take place at microwave frequencies, with no IF present. The conventional method is, in general, complex, requiring many frequency sensitive and active devices, which lower the overall system reliability. In addition, the need for an intermediate frequency sets a limitation on the maximum information bandwidth available. The direct RF to RF system requires fewer frequency sensitive components and is not as limited in bandwidth

since microwave amplifying devices such as traveling wave tubes are available with octave bandwidths. The direct system also is not affected by the modulation scheme present since no intelligence detection is required.

DIRECT MICROWAVE TO MICROWAVE TRANSPONDER SYSTEMS

Crystal Converter Re-entrant TWT Amplifier

The first type of direct microwave to microwave transponder to be discussed utilizes a traveling wave tube (TWT) for amplification and a non-linear impedance device as the frequency converter. The TWT is operated in a re-entrant, or reflex, mode to obtain the necessary gain for the transponder. Figure 1 shows a simplified block diagram of the re-entrant traveling wave tube amplifier.

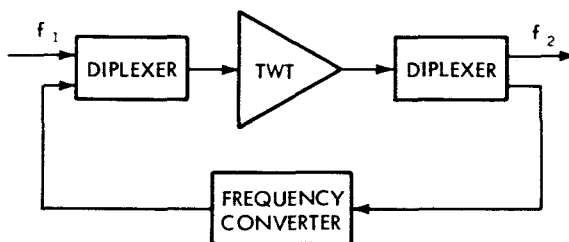


FIGURE 1.—Re-entrant TWT amplifier.

The input signal at a frequency f_1 passes through the first diplexer and is amplified by the TWT. It then is directed by the second diplexer to the frequency converter where it is converted to frequency f_2 . This signal again passes through the first diplexer and it re-enters the TWT where it is again amplified. The signal is then directed by the second diplexer to the output port. In this system the gain of the TWT is utilized twice, once at a frequency f_1 and again at a frequency f_2 .

*Published as *Goddard Space Flight Center Document X-625-64-276*, September, 1964.

Several methods of obtaining frequency conversion by non-linear impedances have been investigated. A tunnel diode or a varactor can be utilized and the gain of the converter made greater than unity. Matching the device over a usable bandwidth is a problem and, to date, sufficient stability for use in a re-entrant loop has not been achieved.

A non-linear resistance in the form of a crystal diode has given satisfactory results for frequency conversion.

Figure 2 shows a crystal converter re-entrant system which utilizes a 2220 Mc local oscillator and a crystal diode to down-convert the input

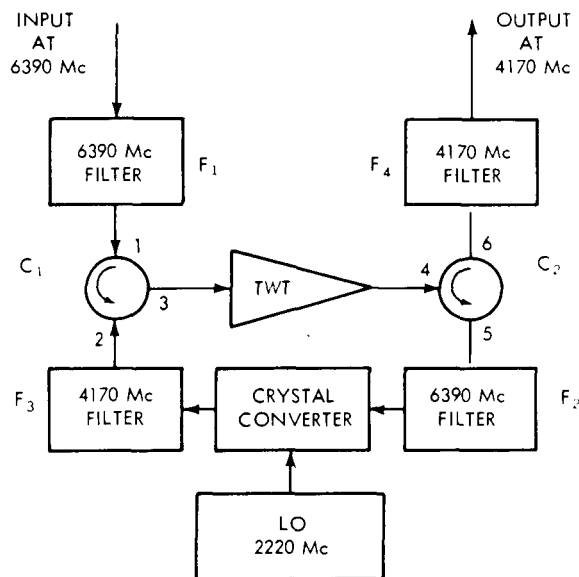


FIGURE 2.—Crystal mixer re-entrant amplifier.

signal at 6390 Mc to an output signal at 4170 Mc. The input signal at 6390 Mc passes through the filter F_1 and enters port 1 of the ferrite circulator C_1 . It is reflected at port 2 from F_3 and enters the TWT through port 3. After amplification in the TWT C_2 directs the signal to port 5 and it enters the crystal down-converter through F_2 . After passing through F_3 , it is directed by C_1 through the TWT at port 3. The signal then enters C_2 at port 4, is reflected at port 5, and is finally passed through F_4 from port 6.

Present day ground station transmitting and receiving capabilities require the satellite to receive signals near the -80 dbm level and retransmit at a $+33$ dbm level, with information band-

widths up to 25 Mc, depending on the mode of operation. Future communication satellite requirements, such as the use of multiple access ground stations and wideband video transmission, indicate that bandwidths of 200 Mc or greater will be required.

For a 200 Mc system with an input level of -80 dbm, the necessary gain to provide a 2 watt ($+33$ dbm) transmitted signal cannot be provided by the re-entrant loop alone, and a receiver noise figure of 5 db or better is required. Therefore, in addition to the re-entrant amplifier, an output power stage is required to realize the necessary gain, and a low noise pre-amplifier stage is required to realize the necessary sensitivity. Figure 3 shows the re-entrant transponder with the power levels, gain, and noise figure of each stage, for an input signal of -80 dbm.

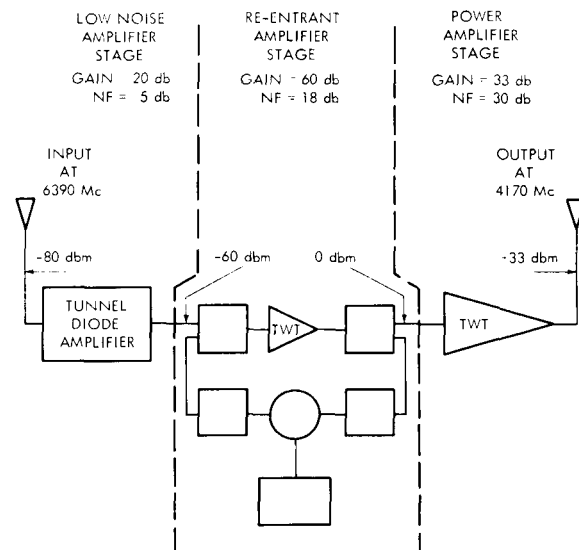


FIGURE 3.—Power levels for a re-entrant transponder with a 200 mc bandwidth.

The noise figure values were derived assuming a single channel at a carrier power level of -80 dbm with a baseband of 10 Mc. The receiver carrier to noise ratio, C/N , for receiver noise power density of -114 dbm/Mc is:

$$\begin{aligned} N_p \text{ (receiver noise power)} &= -114 \text{ dbm/Mc} \\ NF \text{ (noise figure)} &= \frac{5}{-109 \text{ dbm/Mc}} \\ B \text{ (noise bandwidth)} &= \frac{10 \text{ db (10 Mc)}}{-99 \text{ dbm}} \\ N \text{ (receiver noise)} &= -99 \text{ dbm} \\ C/N \text{ (carrier/noise)} &= 19 \text{ db} \end{aligned}$$

For modulation schemes with thresholds of 10 db a margin of 9 db can be realized. Figure 4 shows the receiver carrier to noise ratio for noise bandwidths of 5 to 100 Mc, with receiver noise figures from 3 to 15 db.

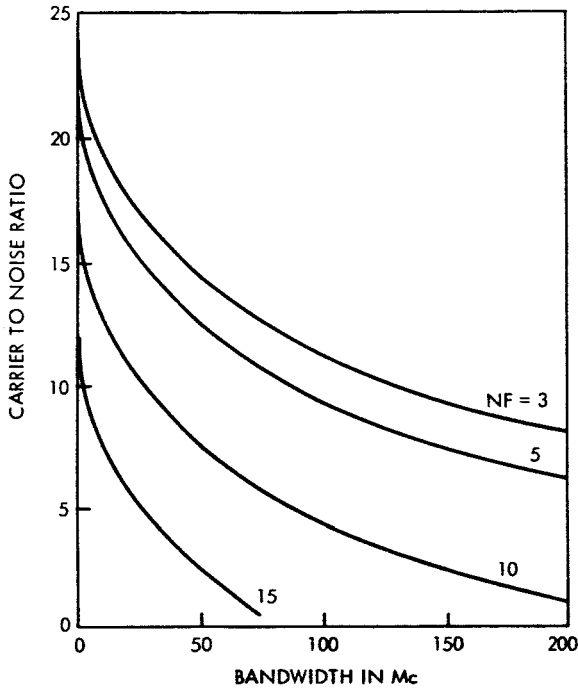


FIGURE 4.—Carrier to noise ratio vs bandwidth for receiver noise figures of 3 to 15 db.

To avoid intermodulation and harmonic distortion products, the TWT in the re-entrant loop must be operated well below its saturated output power. Figure 5 shows the power levels present in the re-entrant system for an input of -60 dbm and a TWT gain of +40 db for both the 6390 Mc signal and the 4170 Mc signal.

The TWT must have sufficient dynamic range to accept both the 6390 Mc signal (-62 dbm) and the 4170 Mc signal (-38 dbm), and the gain suppression at these levels must be held to a minimum. The re-entrant loop filters must have stop band suppressions greater than the gain of the TWT to prevent oscillation of spurious components. The choice of local oscillator frequency to give either sum or difference conversion can be a design problem if harmonics of the local oscillator fall in the information bandwidth of the transponder.

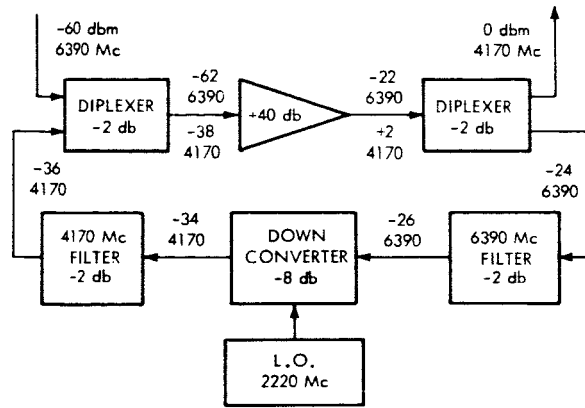


FIGURE 5.—Signal levels in the re-entrant amplifier.

Serrodyne TWT Re-entrant Amplifier

Both amplification and frequency translation can be accomplished in the TWT structure by serrodyne modulation of the TWT helix or beam potential. If the modulating signal is a perfect sawtooth of frequency Δf and the amplitude is adjusted to 2π radians of phase deviation, transit time modulation of the input signal at a frequency f_0 will result in sidebands displaced Δf apart. All of the sidebands and the carrier are suppressed with the exception of one of the first sidebands. This results in a frequency "shift" of the carrier power from f_0 to either $f_0 \pm \Delta f$. The slope of the sawtooth determines which sideband (+ or -) remains. The degree of suppression of the carrier and unwanted sidebands are affected by the flyback time and linearity of the sawtooth, as well as certain helix characteristics. A re-entrant serrodyne TWT amplifier is shown in Figure 6,

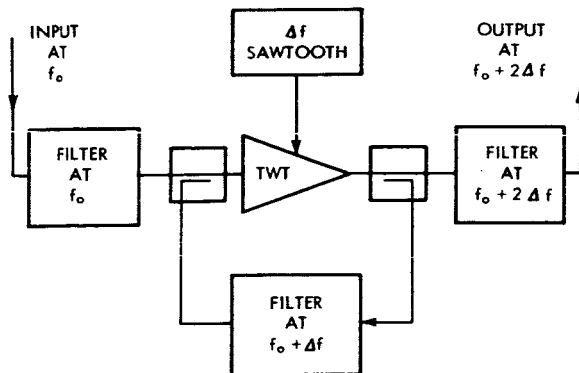


FIGURE 6.—A re-entrant serrodyne TWT amplifier.

for an input at frequency f_0 and an output at frequency $f_0+2\Delta f$. This system requires only one filter in the re-entrant loop. The resultant gain of the loop is the TWT gain at frequency $f_0+\Delta f$ plus the TWT gain at frequency $f_0+2\Delta f$, less the filter and coupler losses.

Various techniques for generating a sawtooth wave and impressing it onto the TWT structure have been investigated. Methods that directly apply the sawtooth waveform are limited by capacitive buildup at the helix, and by the difficulty in producing a linear waveform with a short flyback time at the frequencies required. To date, sawtooth waveforms at sufficient levels and with adequate linearity and flyback times are not available above several megacycles. In addition to conventional techniques, sawtooth waveforms have been generated utilizing harmonically related sinusoidal waves derived from the Fourier series representation of the sawtooth wave. Adequate sideband and carrier suppression have been obtained using the first three harmonics. The use of drift tubes to impress the synthesizing sine waves onto the helix structure has been investigated. This technique reduces the AM to PM distortion of the signal power and minimizes the interaction between the synthesizing signals.

CONCLUSIONS

Future communication satellite applications will require transponder systems with increased bandwidth and signal handling capabilities. The direct microwave to microwave transponder offers these characteristics along with reliability and simplicity of operation.

Investigations of the crystal converter re-entrant TWT amplifier system indicate that frequency conversions above 2000 Mc, with information bandwidths of 200 Mc or more, are attainable in a direct microwave to microwave transponder.

Other techniques for realizing direct conversion, such as a re-entrant serrodyne TWT loop, or an active frequency converter utilizing a tunnel diode or varactor are not, to date, usable in this type transponder; i.e., over large bandwidths. Extensive effort is required to either develop suitable serrodyne generators at the high frequencies or eliminate the problems associated with gain stability and wideband loading and matching of tunnel diodes and varactors.

ACKNOWLEDGMENT

The author wishes to express his appreciation to Louis J. Ippolito for his invaluable assistance in the preparation of this paper.

ELECTRONICALLY STEERABLE ANTENNAS FOR COMMUNICATION SATELLITES*

D. L. BACKUS

This paper presents several electronically steerable antenna techniques which, for frequencies from 1 to 35 gc, show promise for application to communications satellites. It is indicated that such antennas can be realistically designed to provide maximum antenna gain for given weight, maximum bandwidth, and the capability of illuminating only desired geographical areas. Moreover, by incorporating self-tracking capability into the antenna, spacecraft stabilization requirements can be relaxed even to the extent that high gain inertialess antenna systems with omnidirectional response are possible in the case of unstabilized vehicles.

Selected beam forming techniques are presented and it is shown that many steerable antenna designs are derived from these basic techniques. An examination of the circuitry required to implement the design is utilized to prove that the antenna is no longer solely an input or output to a given system but becomes now an inseparable part of the system. Basic problem areas of beam forming, self tracking, and scanning are outlined. Moreover, evidence is given to indicate that lightweight, low loss stripline componentry is presently available to construct real systems.

Several promising techniques, namely the retrodirective and transdirective, are applied to earth-space-earth and space-space communication links. System parameters are listed which show that increased performance and flexibility can be attained.

Present activities underway in space depend in large measure on the communications that can be initiated and maintained with and between the participating vehicles. This requirement for sensory contact is ever present in the planning of anticipated future missions. Substantial effort is being made to establish space communication links from point to point on the earth's surface by means of artificial satellites. Courier, Echo, Telstar, and Syncom are classic examples. The intense interest of the military in satellite data transmission is only an additional example of the recognized promise for such communication links.

Generally, developments and exploration in space require point to point communications where the space vehicle may constitute either a point of origin of transmission or simply a relay station, depending on the particular mission. Figure 1 shows typical data links involving earth, spacecraft and aircraft.

Motion is to be expected of either or both terminals of the link with respect to an arbitrary frame of reference. The volume through which

an antenna system must operate is determined by the orientation of the two ends of a communication link. Usually this orientation is described in terms of the satellite stabilization listed in Figure 2.

When the link is satellite-to-satellite, then the stabilization characteristic must be defined between satellites. However, except in special cases, this system is considered to be unstabilized from the antenna standpoint since control of the orientation between satellites is technically difficult and also expensive in terms of fuel consumption. From the general orbit categories mentioned above six have been chosen to typify communication satellite systems. These are listed in Figure 3.

New environmental demands must be met including those imposed by the ambient vacuum, by temperature extremes, and by micrometeorite bombardment as examples. Restrictions of weight and power consumption become severe, and reliability assumes extreme importance since time and access for repair or service are nonexistent. A stringent weight restriction on equipment is inherent in all space missions. This restriction

*Published as *Goddard Space Flight Center Document X-625-64-208*, October 1964.

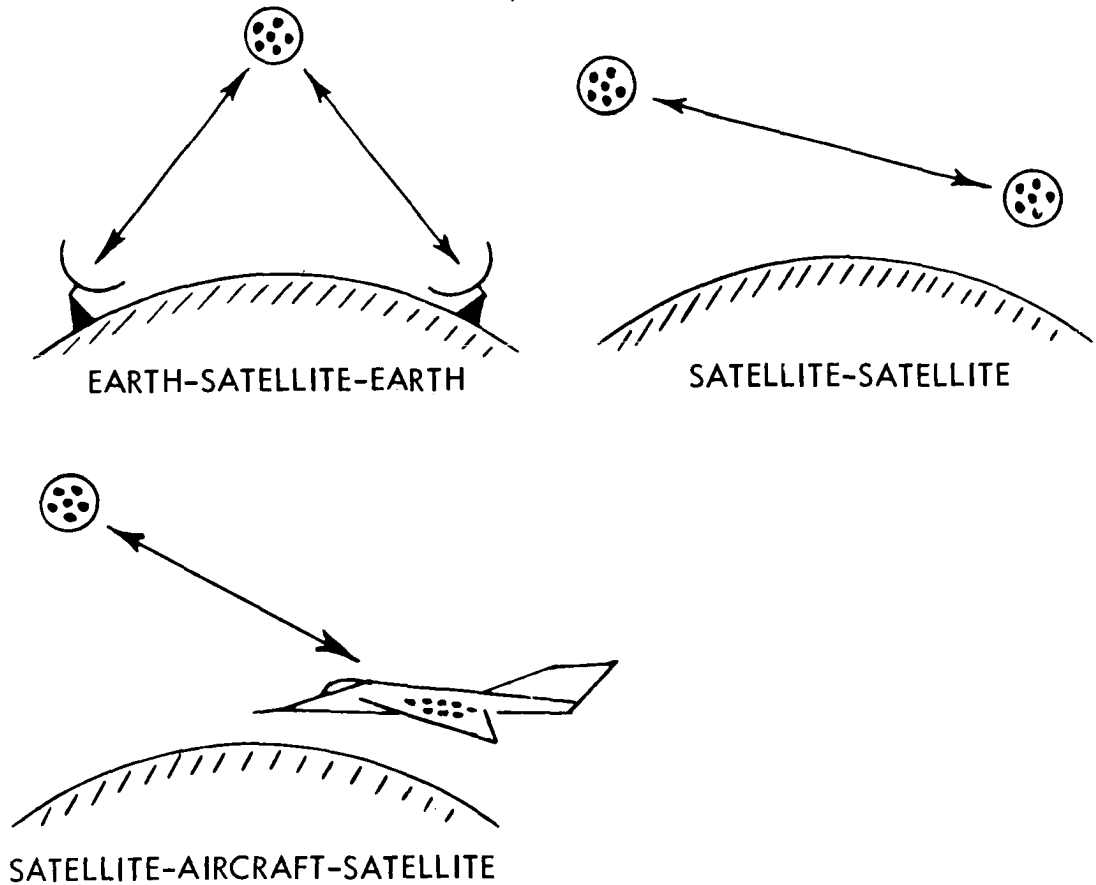


FIGURE 1.—Satellite Communication links.

- (1) Fully stabilized — control is obtained over three axes (e.g. Nimbus, OAO)
- (2) Spin stabilized in plane of orbit — the satellite spins rapidly about one axis — a maximum of 150 revolutions per minute. (e.g. Telstar, Relay)
- (3) Spin stabilized perpendicular to orbiting plane — normally the satellite spin axis lies in the orbiting plane and reorientation is necessary. (e.g. Syncom)
- (4) Gravity Stabilized — the satellite has one axis coincident with a line through the center of the earth. Slow oscillations or librations do occur, however. (e.g. ATS)
- (5) Unstabilized — the satellite may take any orientation. The change in orientation, called tumbling, is usually rather slow.

FIGURE 2.—Degrees of stabilization.

- | |
|--|
| (1) <u>Synchronous orbit</u> – 22,300 miles – spin stabilized |
| (2) <u>Synchronous orbit</u> – 22,300 miles – gravity stabilized |
| (3) <u>Synchronous orbit</u> – 22,300 miles – unstabilized |
| (4) <u>Six thousand mile orbit</u> – spin stabilized |
| (5) <u>Six thousand mile orbit</u> – gravity stabilized |
| (6) <u>Medium altitude orbit</u> – unstabilized |

FIGURE 3.—Satellite orbits.

may be translated rather directly into limitations on available power and on the size of communication equipment. Another factor of extreme significance is that the transmission of signals to and from space vehicles involves communication over great distances. The ranges involved are long even in the instance of an earth orbiting satellite, although such a situation involves short distances for what would be considered a space mission. These two considerations, the importance of weight and the long distances involved, point to the desirability and necessity of achieving as much directivity and efficiency in the antenna structure as the mission situation will permit.

In the past, communication satellite designers have used antennas which are either omnidirectional or quasi-isotropic on spin stabilized spacecraft. Other designs have called for directive horn or reflector antennas placed on spacecraft of varying degrees of stabilization and in certain cases employing electromechanical means to control the beam pattern as dictated by point-to-point spacial variation of the data link. Although generally reliable, and in many applications an optimum design based on system tradeoffs, such antenna/vehicle combinations suffer from a number of limitations (see Figure 4). Reduction in system communication capacity is actually an inherent result of their use. Fig. 5 diagrammatically displays the typical scattering of radio frequency energy throughout space by a transmitting omnidirectional antenna on a communication satellite.

Only the energy impinging on the surface of the ground antenna is usefully recovered. The remaining energy can become a significant con-

tributor to the increasing problem of R.F. interference on the ground. In receive operation the omnidirectional antenna becomes an efficient collector of galactic noise and man-made interference from a large number of geometrically displaced sources.

Certain factors complicate the practical problem. As directivity is increased, the necessary degree of pointing accuracy required also increases. The antenna designer encounters certain limitations. One of these is the capability of the vehicle to control its mechanical orientation. Another is the accuracy with which the orientation may be determined at any given time. One question to what extent it is advisable in a given set of circumstances to attempt to control the vehicle attitude. The answer is of course a practical compromise of those techniques available for antenna searching and tracking, and the degree of vehicle attitude control required.

Inertialess electronically steerable antenna systems are particularly applicable to spacecraft mission problems. Such systems which permit a large number of beams to be independently steered can be made light in weight, low in loss and small in size so that gain and bandwidth can be increased to give better system performance without R.F. interference with existing facilities. Antenna beams can be placed selectively to cover desired geographical areas. A consideration of the antenna techniques must be incorporated in the overall system design for the mission in order to determine essential trade-offs in overall system weight, accuracy, and reliability.

Before presenting two promising antenna systems for space vehicle application, a brief discus-

- (1) Low gain figure
 - (a) results in mediocre system performance
 - (b) for high signal-to-noise ratios, bandwidth is limited
 - (c) require expensive high quality ground terminals (30 foot dish or larger).
 - (d) Space-space operation requires a high power transmitter on the spacecraft.
- (2) Poor directivity
 - (a) antenna does not place the major portion of its effective radiated power at the terminal point.
 - (b) omnidirectional and "earth angle" antenna radiation can cause R.F. interference.
 - (c) antenna exposes spacecraft receiver to noise and man made interference from spacially separated sources.
- (3) Mechanical Inertia
 - (a) electromechanically scanned antennas require spacecraft torque compensators, are subject to bearing and motor failure, in certain applications require rotary joints, and are limited in the scan angle and number of functions they perform.
- (4) Limited Functional Performance
 - (a) electromagnetic access is not selectively controlled by the antenna.
 - (b) the multiple access beam forming capability inherent in inertialess antennas is generally not utilized.
 - (c) the system performance of small aperture ground antennas can be below CCIR standards for T.V. bandwidths.

FIGURE 4.—Communication satellite antenna limitations.

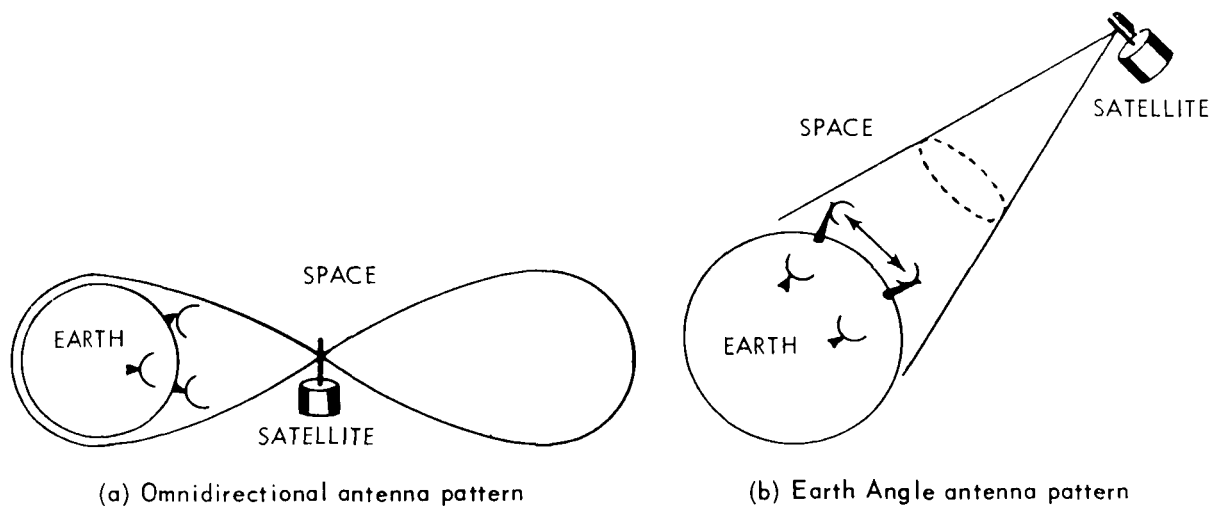


FIGURE 5.—Indiscriminant radiation of R. F. Energy and susceptibility to noise.

sion of the special properties of electronically steered antennas is in order. The assembly and functional operation of a steerable array can begin by referring to Figure 6 which depicts an n -element linear array. The vector normal to the array is defined as the boresight direction. On receive operation energy arrives at the array from some angle θ with respect to boresight. Let the distance separating the wavefront from each antenna element be $0, P, 2P, 3P \dots (n-1)P$. This will result in a corresponding time delay $0, T, 2T, 3T \dots (n-1)T$. Equations are given in Figure 6 which relate the differential time delay to the differential phase shift between adjacent ele-

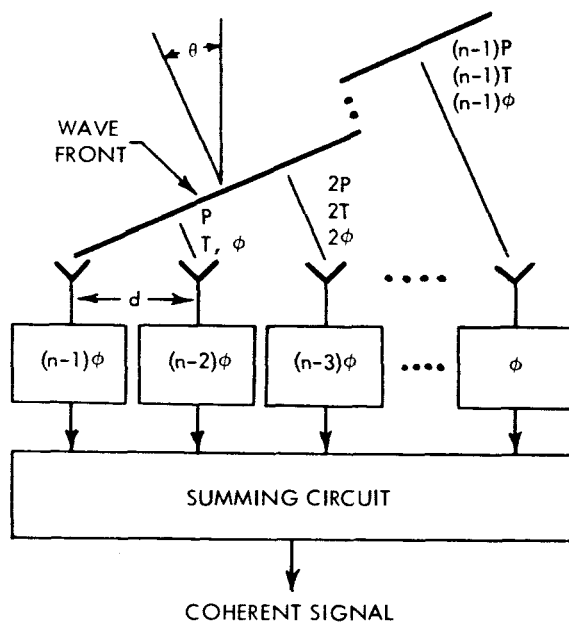
ments of the array. The energy is added incoherently since the phase in any channel differs from the others. The addition of a variable phase shifter behind each element allows the channel outputs to be coherently summed by forcing equivalence of the total argument of the phase function at each channel output. Phase shifts of $0, \phi, 2\phi, 3\phi \dots (n-1)\phi$ would be required. The response of the array or beam pattern is indicated by the magnitude of the sum pattern and will be less for angles of arrival different from θ . The output of the summing bus is more commonly called the beam output; the process of summation is referred to as beam forming.

In conventional arrays, element spacing is approximately one-half wavelength. Spacings less than 0.4λ give mechanical as well as mutual coupling problems, while spacings larger than 0.7λ produce grating lobes in the beam pattern. Under the condition of 0.5λ spacing, the array beamwidth is less than that obtained with a reflector antenna of the same dimensions. This is in agreement with the higher aperture efficiency of the array. As is experienced in conventional antennas, control of side lobe level can be achieved by proper weighting of the illumination function. The signals in each channel are weighted in amplitude prior to summation. This allows the side lobe level of the antenna beam pattern to be set at practical levels.

An infinite number of orientations of the phase front in Figure 6 are possible. If the phase shifter setting $\phi, 2\phi \dots (n-1)\phi$ are varied proportionately, the differential phase shift across the aperture must also change in order to achieve coherent summation. In effect, the angle of arrival θ of R.F. energy which is added coherently by this antenna has been altered. This is generally referred to as beam steering.

The point of importance to be made here is that considerable variation in beam pattern shape can be achieved by exercising proper control of phase and amplitude at each element of the array. Simultaneous beam formation by the antenna in a typical space-earth and space-space link is shown in Figure 7.

Considerable effort has been expended in the area of beam forming and steering by certain investigators. Although much of the earlier work has been applied to the development of



$$T = \frac{P}{C} \quad P = d \sin \theta$$

$$\phi = \frac{2\pi d}{\lambda} \sin \theta$$

where θ = the beam angle
 ϕ = differential phase shift
 T = differential time delay
 P = differential path length
 d = element spacing
 n = number of elements
 c = speed of light

FIGURE 6.—Special properties of an electrically steered array.

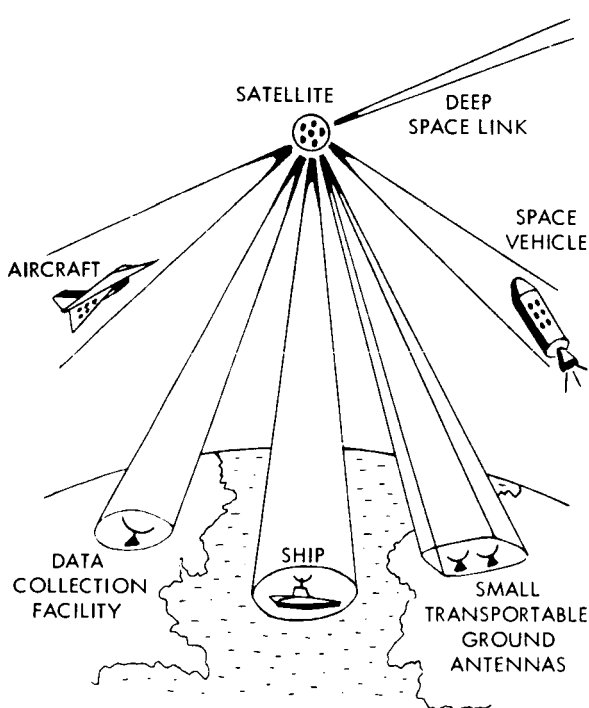


FIGURE 7.—Formation of simultaneous steerable beams by the spacecraft antenna in space-earth and space-space data links.

large ground antennas, in comparison, little (few exceptions)* has been done to further the development of electronically steerable antennas for communication satellites. A large number of techniques for beam forming and steering are available, but only a few are suitable for space application. Two promising techniques, namely the Retrodirective and the Transdirective appear to be practical in the sense that components required for their demonstration are presently available in the frequency range of interest (1 to 10 gc).

The Retrodirective principle in its simplest form is illustrated in Figure 8. A CW signal received by the n^{th} element of an array can be represented by $\sin(\omega t + \theta n)$, where ω is the angular frequency and θn is the phase of the incoming signal at the n^{th} element, relative to some common reference. The signal is subtracted from (mixed with) a common local oscillator, $\sin 2\omega t$, to obtain a transmit signal $\sin(\omega t - \theta n)$ which is the same as the received signal

*See reference.

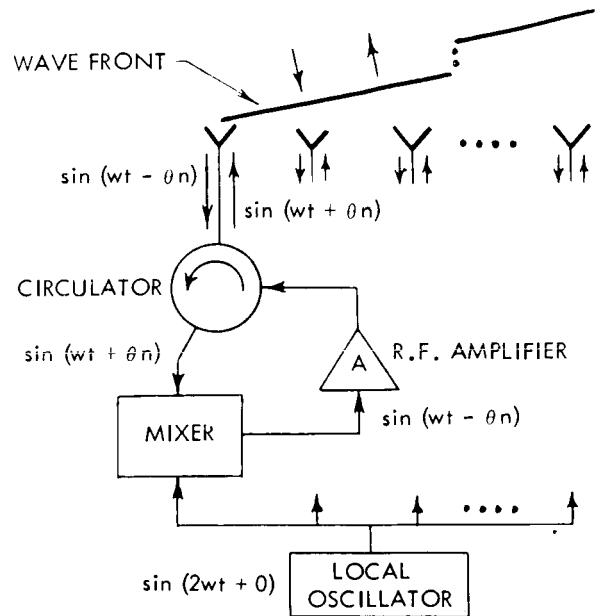


FIGURE 8.—Simplified retrodirective principle.

except that it has an inverted phase angle. The transmit signal is amplified and transmitted out the same antenna element. A circulator is used to separate the transmit and receive signals. The phase angle is similarly inverted at every element; this is the condition for transmitted energy to form a beam in the same direction as the received signal. Such an array works as an active device to reflect a signal back towards its source. No frequency shift between the uplink and the downlink is accomplished and no signal output is available at the satellite.

A practical variation of the retrodirective principle is given in Figure 9. This design is called a Retro-Redirective Antenna since it has the capability to receive energy from one direction at frequency f_1 and re-transmit it in another arbitrary direction at frequency f_2 . In addition, the array may be operated in the simple retrodirective mode similar to that of Figure 8 but with the advantage of frequency separation between the transmit and receive signals. In its present form the more directive array is not used to receive the wideband signal (normally in earth-space communications the uplink is the strong link). In this system a pilot signal at frequency f_p arrives at the array from station B, and is mixed to accomplish phase inversion. A wideband

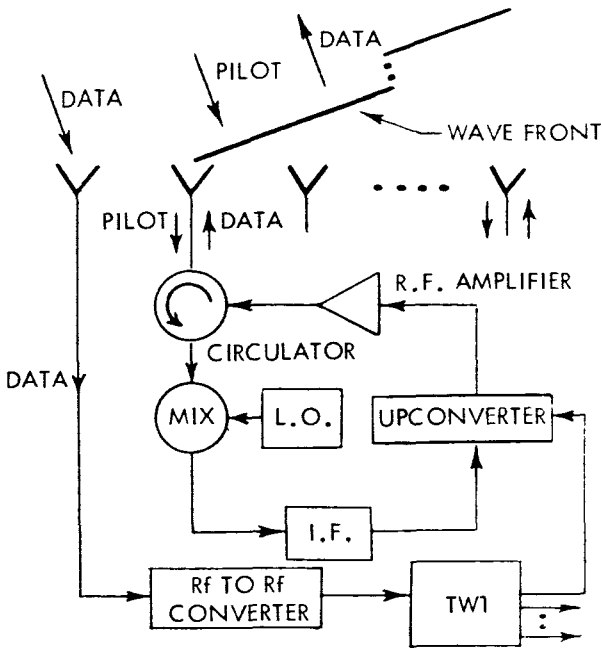


FIGURE 9.—A retro-redirective antenna system.

signal at frequency f_1 from station A is received by a horn antenna of earth angle coverage, is frequency shifted to f_2 in an RF to RF converter, and amplified in a traveling wave tube. The inverted phase information is applied to the wideband signal at f_2 in a upconverter, is amplified in a tunnel diode amplifier, and radiated from the array with a high gain figure toward station B.

The system block diagram for a Transdirective Antenna System is indicated by Figure 10. This

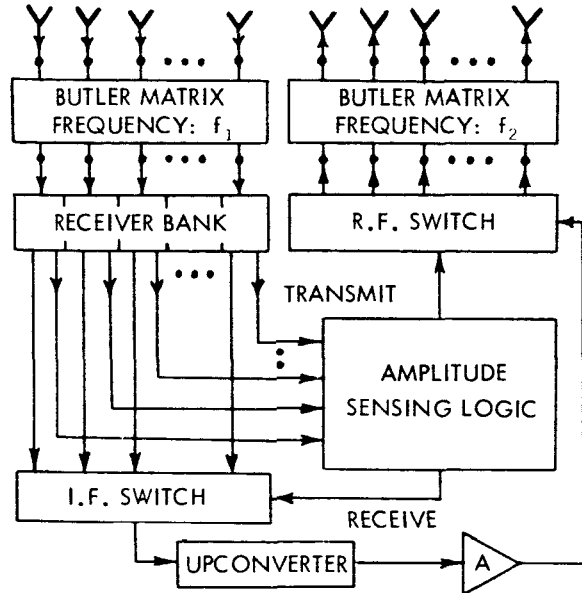


FIGURE 10.—A transdirective antenna system.

technique makes use of the properties of a Butler matrix array to form multiple beams throughout a 2π steradian volume at two separated frequencies. Similar arrays are utilized at each frequency; the elements of each spaced at approximately one-half of a wavelength. The N by M Butler Array will then form approximately $N \times M = n$ discrete beams which are spacially superimposed at the frequencies f_1 and f_2 . Therefore, corresponding terminals of the matrices actually represent two identical beams displaced only in frequency. Self tracking is accomplished

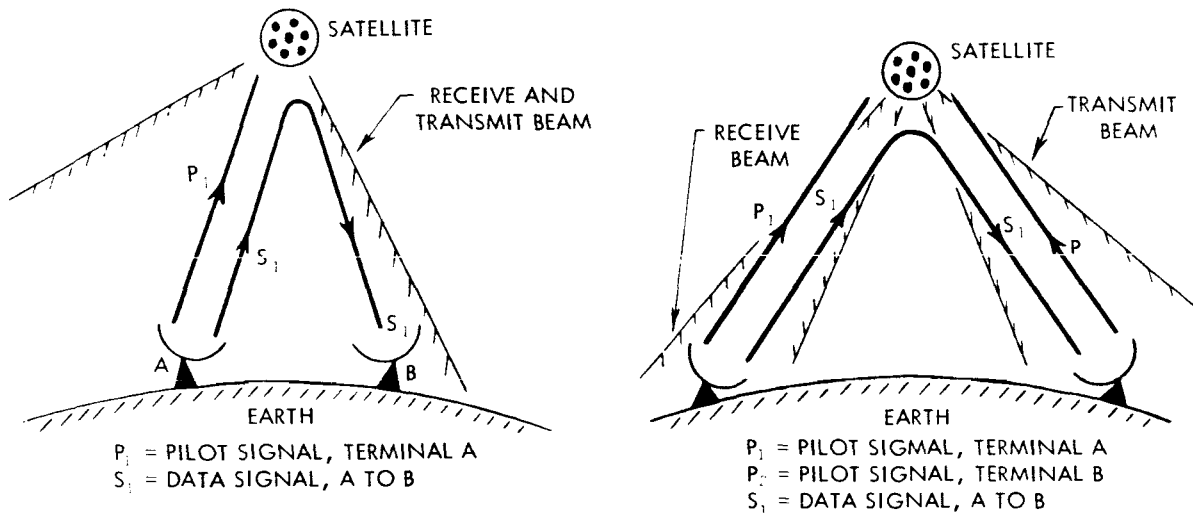


FIGURE 11.—(a) Retrodirective antenna (monostatic). (b) Re-directive antenna (bistatic).

Link Frequency gc.	earth-space 6	space-earth 4
Xmtr Power, dbm	70 (10 KW)	33 (2 W)
Xmt Antenna Gain (15' dia. dish), db	46.5	3
Space Loss, db	-186.5	-183
Rcv. Ant. Gain, db	3	43
Rcvr. Noise Density dbm/mc	-104.5 (10 dbNF)	-116 (2 dbNF)
Rcvr. Noise Bandwidth (10 mc), db mc	10.0	10.0
Rcvr. Signal Power, db	-67.0	-104
Rcvr. Noise Power, dbm	-94.5	-106
Predetection Carrier to Noise, db	27.5	2

A. Dipole "earth angle" Spacecraft Antenna

Link Frequency gc.	earth-space 6	space-earth 4
Xmtr Power, dbm	70 (10 KW)	33 (2 W)
Xmt Antenna Gain (15' dia. dish), db	46.5	30
Space Loss, db	-186.5	-183
Rcv. Ant. Gain, db	30	43
Rcvr. Noise Density dbm/mc	-104.5 (10 dbNF)	-116 (2 dbNF)
Rcvr. Noise Bandwidth (10 mc), db mc	10.0	10.0
Rcvr. Signal Power, db	-40.0	-77
Rcvr. Noise Power, dbm	-94.5	-106
Predetection Carrier to Noise, db	54.5	29

B. Electronically Steerable Spacecraft Antenna (3 foot by 3 foot aperture)

FIGURE 12.—Predicted performance of a 5000 mile orbit communication satellite system utilizing a 15 ft diameter ground antenna.

by sensing the amplitude of pilot signals transmitted from each ground terminal. An I.F. switch is used to select the receive beam. Although amplitude variation is expected due to the

4 db crossover between adjacent beams, cross coupling circuitry should reduce this excursion.

The basic techniques presented can be extended to allow multi-station operations in a multiple

access mode. Additional work is required in this area in order to facilitate a practical design. Typical earth-space-earth Retro-directive and Re-directive links for communication satellite application are given in Figure 11.

CONCLUSION

System calculations are provided in Figures 12 and 13 which show that increased system performance can be attained in the data link through the use of high gain inertialess antennas. Wide-

Link Frequency, gc	earth-space 6	space-earth 4
Xmtr Power, dbm	70 (10 KW)	33 (2 W)
Xmt Ant. Gain (15' dia. dish) db	46.5	5
Space Loss, db	-201	-197.3
Rcv. Ant. Gain, db.	5	43
Rcvr. Noise Density, dbm/mc	-104.5 (10 dbNF)	-116 (2 dbNF)
Rcvr. Noise Bandwidth (10 mc), db mc	10.0	10.0
Rcvr Signal Power, dbm	-79.5	-116.3
Rcvr Noise Power, dbm	-94.5	-106
Predetection Carrier to Noise, db	15.0	-10.3

A. Dipole "earth angle" Spacecraft Antenna

Link Frequency, gc	earth-space 6	space-earth 4
Xmtr Power, dbm	70 (10 KW)	33 (2 W)
Xmtr Antenna Gain (15' dia. dish), db	46.5	30
Space Loss, db	-201	-197.3
Rcv Ant. Gain, db	30	43
Rcvr Noise Density, dbm/mc	-104.5 (10 dbNF)	-116 (2 dbNF)
Rcvr Noise Bandwidth (10 mc), db mc	10.0	10.0
Rcvr Signal Power, dbm	-54.5	-91.3
Rcvr Noise Power, dbm	-94.5	-106
Predetection Carrier to Noise, db	40.0	14.7

B. Electronically Steerable Spacecraft Antenna (3 foot by 3 foot aperture)

FIGURE 13.—Predicted performance of a 22,300 mile synchronous communications satellite system utilizing a 15 ft diameter ground antenna.

band communication between small inexpensive ground terminals is possible while stabilization requirements on the spacecraft are actually relaxed.

A number of high gain independently steered beams can be formed. Beamwidths can be synthesized to cover only desired earth or space areas, reducing R.F. interference with existing facilities. Power densities resulting from spacecraft radiation will be high at the terminal points, but such is not the case elsewhere. To achieve the system performance predicted by Figures 12 and 13, the recommended earth power density limits of the CCIR might have to be exceeded; however, favorable location of the ground terminals should preclude any problem of interferences. Beamwidths of 3 or 4 degrees appear appropriate for most applications. Alternately, fan beams can be used as the mission requires particularly where narrow ground areas are to be illuminated. The availability of lightweight, low loss stripline componentry used in the antenna element circuitry is a requirement if electronically steerable antenna systems similar to ones discussed are to become spacecraft hardware. Extension of the ideas presented to the millimeter wavelength region of the electromagnetic spectrum could result in the development of high performance, high resolution, self tracking antenna systems with larger information bandwidths than are presently available with existing systems. It should be noted that while the term "antenna"

is generally used to describe conventional designs, the antenna techniques presented require that the spacecraft antenna be an inseparable part of the receiving/transmitting system. This salient difference demands that the more accurate term "antenna system" be used.

BIBLIOGRAPHY

1. "Final Report on Electronically Steerable Antenna Systems," Sanders Associates, Inc.; June 1960 to June 1962; Signal Corps Contract No. DA-B6-039SC-85198.
2. "Spherical Retrodirective Array," E. M. Rutz-Philipp Federal Systems Division of IBM; Bethesda, Maryland.
3. "Satellite Data Transmission System, Final Report," ASD-TDR-63-759; May 1963, Sylvania Electronic Systems.
4. "A Self-Steering Array Repeater," C. C. Cutler, et al; BSTJ; September 1963.
5. "Beam Forming Matrix Simplifies Design of Electronically Scanned Antennas," J. Butler and R. Lowe; Electronic Design; April 1961.
6. "Directive Array for a Spinning Vehicle," H. A. Rosen; Proceedings of the National Telemetry Conference, May 1962.
7. "The Advanced Syncom Communications Antenna System, A Directional Antenna for a Spin Stabilized Satellite," H. R. Erhardt, G. Gerson, D. C. Mead; National Space Electronics Symposium, September 1963.
8. "Study of Spacecraft Antenna Systems, Final Report," Hughes Aircraft Co.; August 1964; NASA/GSFC Contract NAS 5-3545.
9. "Final Acts of the EARC," Geneva 1963, ITU.

ORBITAL PLACEMENT AND CONTROL OF THE SYNCOM II SPACECRAFT*

JAMES L. BAKER

PREFACE

This paper is intended to present some detailed data on the performance of control system elements in addition to the principles and sequence of operation of the orbit placement and control system for the SYNCOM II spacecraft.

ACKNOWLEDGMENT

I wish to express my appreciation for the guidance in preparation of this paper contributed by Mr. Alton E. Jones, Chief, Communications Satellite Projects Branch, and Mr. R. J. Darcey, SYNCOM Project Manager of the Goddard Space Flight Center of NASA. The members of the Space Systems Division of the Hughes Aircraft Company, who were responsible for developing the SYNCOM for NASA, are to be commended for their outstanding implementation of this project.

I. INTRODUCTION

Successful NASA ventures in low altitude spacecraft systems carrying scientific and technological payloads were augmented substantially by the achievement of a synchronous orbiting satellite. The spin stabilized SYNCOM II, with its gas reaction motors, provides a simplified approach to stability and control in orbit which has been demonstrated since its launch on July 26, 1963.

II. OBJECTIVES

The basic objectives of the SYNCOM II mission were as follows:

(a) To gain early experience with a communication satellite in synchronous orbit.

(b) To develop the capability of launching satellites into a 24-hour orbit, using existing launch vehicles with additional apogee-boost propulsion capability.

(c) To determine the communications system's performance parameters, in a design configuration capable of relaying with high-quality two voice-modulated carriers (either FM or PCM/PM) between widely separated surface terminals.

(d) To obtain scientific data on space propagation.

(e) To test the life of communications-satellite components at the 24-hour orbit altitude.

(f) To obtain data from the quantitative effects of earth's "triaxiality" (asymmetry around its spin axis) on the synchronous orbit.

(g) To develop a spinning communications spacecraft with gas reaction control systems and demonstrate its stability and maneuverability in the synchronous orbit.

This paper deals only with aspects of the synchronous orbit placement and control of the SYNCOM II spacecraft. Objectives related to communications experiments were met with a high degree of success and have been reported elsewhere.

As a minimum, the control system should have the capability to (1) remove initial launch and apogee motor injection errors; (2) move the spacecraft to a westward (or eastward) position (3) change its spin axis relative to the orbital plane, (4) accommodate corrections to the orbit perturbation caused by the earth's "triaxiality," and (5) remove secular and periodic drift. The spacecraft control and communication system should be designed to operate a minimum of one year in orbit.

The choice of spacecraft design parameters, the ground tracking equipment and communications system is dependent largely on the characteristics of the synchronous orbit. It is of interest to the communicator to consider these characteristics of the orbit from the point of view of the sub-satellite trace.

Behavior of the spacecraft orbital motion, as manifested by this sub-satellite trace, may result

*This work was presented at the 15th International Astronautical Congress, Warsaw, Poland, September 1964.

from design, hardware limitations, excessive errors, or combinations of these factors. Errors may be reduced or eliminated if capability is present within the spacecraft and ground control system. If the capability exists one may correct the orbit subsequent to launch and injection by changing either inclination, eccentricity, period or a combination of these.

Classical elements of the orbital plane geometry are provided in Figure 1.

Figures 2(a) and 2(b) show the results of the synchronous mission where a periodic drift occurs east and west of the nodal longitude due to eccentricity, and north and south due to inclination of the orbital plane to the equatorial plane. Figures 2(c) and (d) result from argument of perigee ($\omega = 0$ or 180°) for small value of inclination and eccentricity. A closely related condition is shown in Figure 2(e). In addition to the periodic drift,

ORBITAL PLANE RELATIVE GEOMETRY

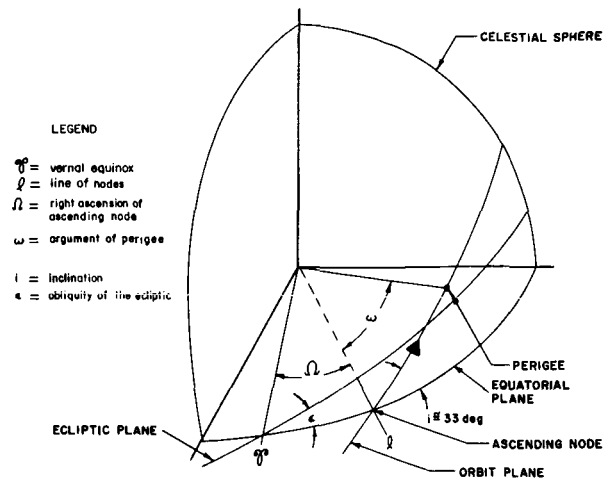


FIGURE 1.—Orbital plane relative geometry.

TRACES OF SUB-SATELLITE POINTS

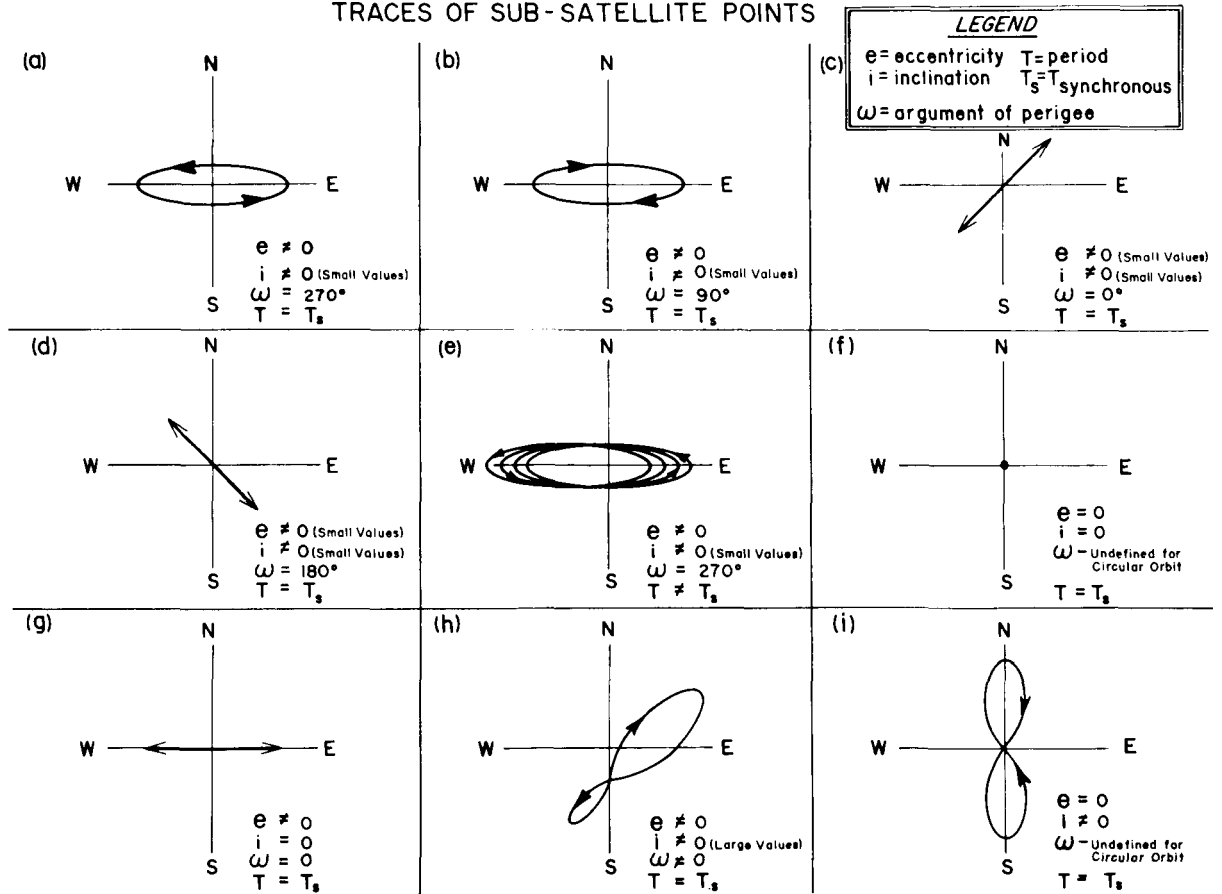


FIGURE 2.—Traces of sub-satellite points.

a secular drift is induced when the orbital period is greater (westward drift) or less (eastward drift) than synchronous.

Figure 2(f) shows the sub-satellite point of a synchronous orbit spacecraft which is placed in a circular equatorial orbit with no inclination to the equatorial plane. This condition is popular from the standpoint of eliminating the Doppler shift and reducing complexity of ground station antennas. It requires, however, that either the booster be launched from a point on the equator, or a plane change effected by the booster and/or spacecraft apogee motor. Eccentricity imposed on the orbit has the effect of producing an oscillation east and west along the equator with a 24-hour period, as indicated in Figure 2(g).

A skewed trace exists when the period is synchronous, but eccentricity, inclination and argument of perigee are not equal to zero as shown in Figure 2(h). A symmetrical trace, Figure 2(i) is present about the nodal longitude and the equator when the synchronous orbit possesses an inclination to the equatorial plane and eccentricity is eliminated.

This latter orbit, characterized by the "figure eight" sub-satellite trace, was selected for SYNCOM II, because of performance capability of the available DELTA vehicle. This, obviously, resulted in payload restrictions in weight and the minimum inclination between the equatorial plane and the orbital plane.

A target nodal longitude of 55° W was selected beforehand to provide for communications visibility between the United States and Africa and to satisfy the phase of the experiment designed to demonstrate the capability of the spacecraft to maneuver to and remain at a precise location.

III. TECHNIQUE OF ORBITAL PLACEMENT AND CONTROL

The design of the SYNCOM II flight mission is predicated on the use of a classic technique, the Hohmann transfer ellipse, to place the maximum payload at the synchronous altitude in a circular orbit.

The planar injection geometry for a typical 24 hour mission without drift is shown in Figure 3. The spacecraft is mounted on top of the launch vehicle in what appears to be an inverted position. The nozzle of the apogee "kick" (injec-

A. SYNCHRONOUS ORBIT
 $r_o = 22,752.3$ n. mi
 $V_o = V_a + \Delta V_a = 10,087.5$ fps

B. TRANSFER ORBIT
 $r_o = 3560$ n. mi
 $V_c = 33,546.0$ fps
 $V_a = 5248.8$ fps

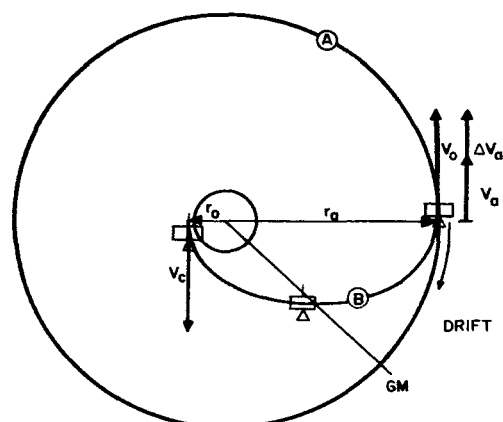


FIGURE 3.—Planar injection geometry.

tion) motor, fixed within the spacecraft, faces the cone end of the shroud. This position is germane to the design of the injection phase.

At launch the three stage Thor-Delta booster delivers a thrust sufficient to provide a perigee tangential velocity $V_c = 33,546.0$ feet per second (fps) at a radius r_o of 3560 nautical miles. Prior to attaining this velocity at perigee the solid propellant third stage, to which the spacecraft is attached, is spun up to 150 rpm by tangentially mounted solid propellant motors fastened to the periphery of the spin table on the forward end of the second stage. When spin-up is complete the spinning third stage, with spacecraft, is ignited and provides the velocity increment to achieve V_c whereupon the spacecraft is separated from the third stage. The spin stabilized craft "coasts" toward apogee along its transfer orbit path, B, with its spin axis fixed in inertial space. At apogee, the altitude r_a is equal to the synchronous altitude and the velocity is 5248.8 fps. The apogee motor nozzle is aligned, due to gyroscopic stability, so that its thrust axis is tangent to the transfer orbit at apogee. The velocity ΔV_a , imparted by the motor, added to the transfer orbit velocity \bar{V}_a at apogee is 4,838.7 fps which yields the required synchronous orbit velocity \bar{V}_o , of 10,087.5 fps. If an eastward drift is desired, the energy of the transfer orbit should be programmed to be less than that for synchronism, assuming nominal performance of the apogee

motor. The spacecraft will drift westward if the energy in the transfer orbit is biased toward greater energy than that required for injection into a synchronous orbit.

Following injection into a circular orbit by the apogee "kick" motor, the spacecraft drift, orientation and orbit eccentricity are controlled by the on-board gas-reaction jets. Initial injection errors are removed by the axial jet whose thrust axis is parallel to the spin axis of the spacecraft. The jet is initiated by command at the required point(s) in the orbit and functions continuously over many revolutions to produce the desired synchronism (or drifts) and circularity. Spacecraft orientation maneuvers for purposes of (1) jet thrust axis alignment (2) improved solar array—sun angle (ϕ), and (3) communications antenna pattern coverage are accommodated by the same axial jet; however, it is pulsed in synchronism with the spin rate for a fixed sector (approximately 60 deg.) of each revolution.

The drift of the spacecraft after its precession to the required orientation is controlled by pulsing a radial jet, whose thrust axis passes through the center of gravity of the spacecraft and is normal to its spin axis.

Velocity corrections to the spacecraft for nominal conditions in order to obtain synchronism with the earth's rotation and circularize the orbit are shown in Figures 4 through 9 for various initial orbital conditions. Although corrections

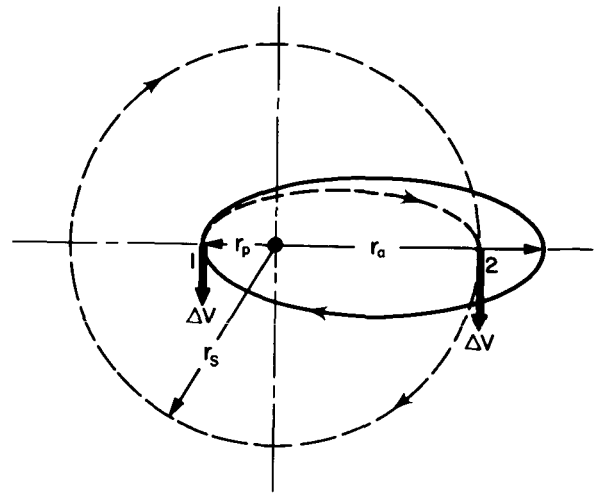


FIGURE 5.—Correction sequence for $r_p < r_s < r_a$

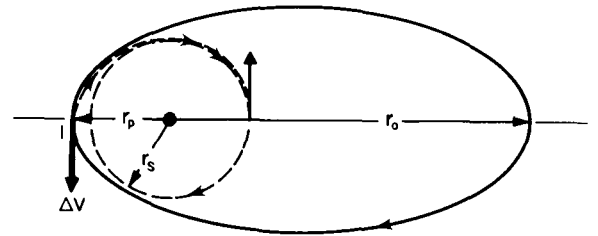


FIGURE 6.—Correction sequence for $r_s < r_p < r_a$

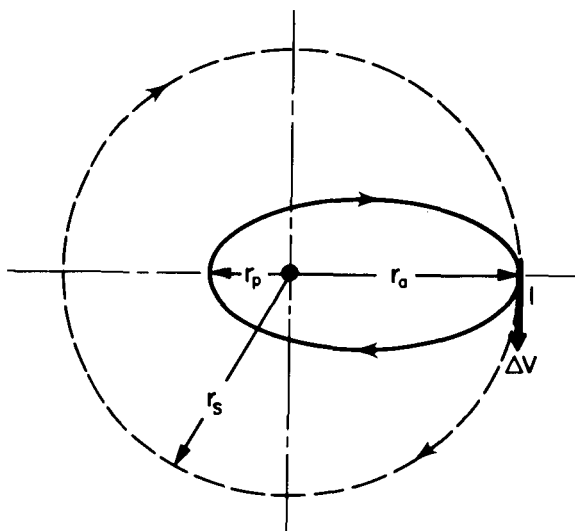


FIGURE 4.—Correction sequence for $r_p < r_a = r_s$

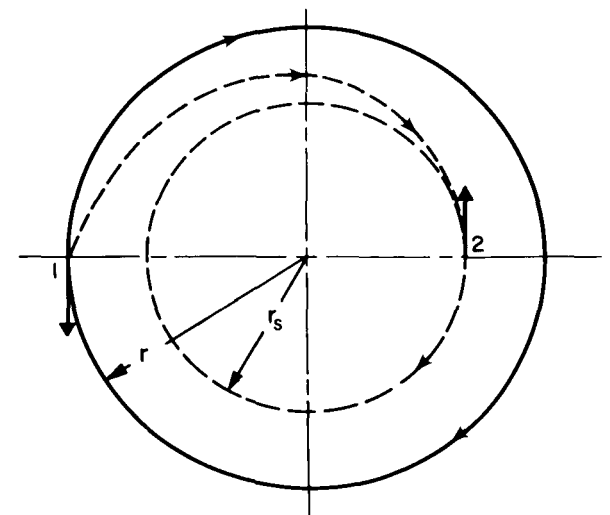


FIGURE 7.—Correction sequence for circular orbit. $r > r_s$

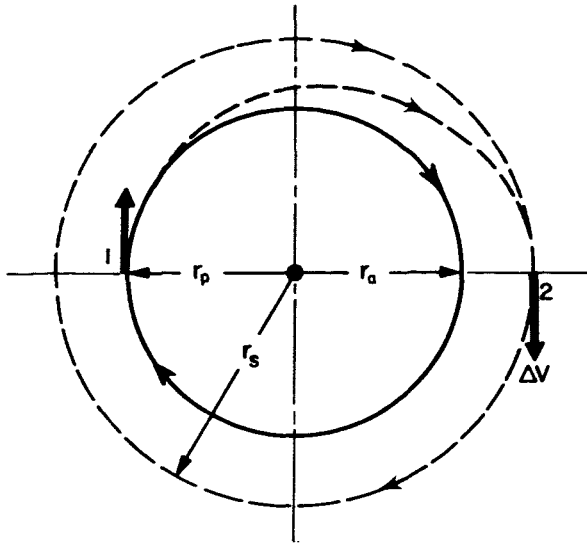


FIGURE 8.—Correction sequence for $r_p < r_s < r_a$.

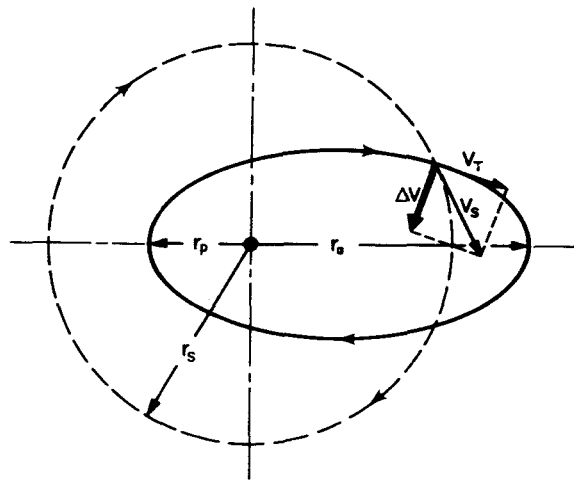


FIGURE 9.—Combined correction to provide a synchronous and circular orbit.

to these orbits by thrusting the spacecraft jets may be achieved by more than one approach, the intent is to illustrate the usual technique employed for efficient maneuvers. The dotted circle in each figure represents the circular synchronous orbit, and the velocity increment (ΔV) arrows represent the direction only of the tangential thrust vector relative to the orbital motion. In most cases, the overall maneuver is most efficient if the first correction is made at perigee.

Figure 4 illustrates the eccentric orbit where apogee occurs at the synchronous radius (r_s) with

perigee (r_p) less than r_s . An appropriate velocity increment (ΔV) added in parallel to the orbit velocity vector at apogee (r_a) will circularize and synchronize the orbit.

Figure 5 is representative of an initial orbit where $r_p < r_s < r_a$. Orbital energy is removed at perigee by phasing the thrust of the jet in the reverse direction to the orbit velocity vector such that the resultant apogee is equal to r_s . A second thrust at the new apogee adds the required energy to obtain a synchronous, circular orbit.

Where $r_p > r_s$, as shown in Figure 6, orbital energy is removed by phasing the radial jet average thrust opposite to the orbital velocity vector. The required amount of thrust puts the new perigee at r_s . A thrust opposite to the orbital velocity vector is again applied at the new perigee, 2, to remove the required energy from the orbit for synchronism and circularity.

The method for achieving synchronism and circularity in Figure 7 is the same basic technique as applied in Figure 6, and is merely the correction of a special case of the higher energy orbit.

Figure 8 depicts the orbit whose apogee and perigee is less than r_s , and, although an ellipse is shown, the basic sequence of events is applicable for correcting a circular orbit with $r < r_s$.

A velocity correction to change the period should be planned for the time in orbit which also has the combined effect of reducing the eccentricity as well as changing the drift rate.

Corrections requiring very small velocity increments may be combined into one maneuver to change both the period and the eccentricity as illustrated in Figure 9. If a large change in either one or both is required the inefficiency of the compound maneuver may be prohibitive. SYNCOM II was corrected in this manner and this maneuver suffered the loss of only 1 fps out of the 3 fps required.

As stated earlier, no attempt was made to remove inclination due to the large amount of fuel required. The synchronous equatorial orbit, however, requires this correction maneuver following injection into the synchronous orbit and nulling the north and south (N-S) drift due to the sun's and moon's gravitational fields. The relative geometry of inclination removal is shown in Figure 10 to aid in explaining the use of the axial jet in the continuous mode. In this maneuver

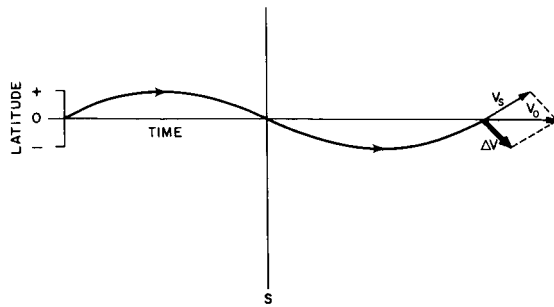


FIGURE 10.—Maneuver for inclination removal.

the spacecraft must be precessed relative to the inclined orbit's velocity vector, \bar{V}_s , to align the thrust of its axial jet coincident with the required $\Delta\bar{V}$. When the spacecraft intersects the equatorial plane, the axial jet is operated in the continuous mode for the amount of thrust required to produce a resultant velocity, \bar{V}_0 , lying in the

equatorial plane with a magnitude equal to the synchronous velocity.

Reorientation can be accomplished at any point in the orbit; however, the initiation of continuous jet function for inclination removal must occur at the ascending or descending node.

Future spacecraft missions may require that inclination corrections be made following injection into the equatorial synchronous orbit. In addition requirements may provide for correcting inclination perturbations caused by the gravitational force of the sun and moon. Their combined effects result in a latitude change of $\pm .85$ degrees and the estimated average velocity increment for one year's station keeping is 150 fps.

The technique of orbital placement and control of SYNCOM II requires a high degree of reliability and wide range of control. Comparison of the velocity increments imparted by the maneuvers listed in Table 1 shows this variation, and

TABLE 1.—Spacecraft Maneuvers Completed as of August 15, 1964

Maneuver	Fuel Jet Used	Jet on Time (Sec)	No. of Pulses	Fuel Weight Used (ΔW) Predicted (lbs)	ΔW Actual (lbs)	Incremental Velocity ΔV Predicted (Ft/Sec)	ΔV Actual (Ft/Sec)
Apogee Motor Burn 7/26/63	-----	20.20	-----	61.3	61.28	4696.1	4712.1
1st Vel. Corr. 7/27/63	H ₂ O ₂ Axial	140	Continuous	1.80	1.80	107.5	110.47
1st Reorientation 7/31/63	H ₂ O ₂ Axial	143	350	.40	.40	15.5	29.35
2nd Vel. Corr. 8/11/63	H ₂ O ₂ Lateral	393	894	1.03	.91	44.06	37.28
3rd Vel. Corr. 8/12/63	N ₂ Lateral	510	1152	.495	.405	13.75	13.15
4th Vel. Corr. 8/15/63	N ₂ Lateral	924	2023	.471	.502	12.72	14.70
5th Vel. Corr. 8/16/63	N ₂ Lateral	120	264	.049	*	1.33	-----
2nd Reorientation 8/18/63	N ₂ Axial	95	195	.060	.06	<1.0	<1.0
6th Vel. Corr. 11/28/63	H ₂ O ₂ Lateral	31.5	65	.060	.08	2.79	4.39
7th Vel. Corr. 3/18/64	H ₂ O ₂ Lateral	150	308	.280	.257	12.52	10.83
8th Vel. Corr. 4/25/64	H ₂ O ₂ Lateral	70	140	.140	.140	4.90	4.79
9th Vel. Corr. 7/3/64	H ₂ O ₂ Lateral	49	98	.090	.106	3.02	3.50

*Too small to detect pressure change with sufficient accuracy.

reliability is limited only by the amount of fuel available.

IV. APOGEE INJECTION MOTOR & ORBIT CONTROL SYSTEM

Two basic reaction systems aboard the spacecraft provided the implementation of the orbital placement technique and corrections to the orbit and attitude of the craft. These were the solid propellant apogee motor and the gas reaction motors, respectively.

The SYNCOM II apogee injection motor was developed for the Goddard Space Flight Center by the Jet Propulsion Laboratory of the California Institute of Technology. Motor weight was approximately 70 lbs. with estimated burn time of 19.7 seconds providing a velocity increment of 4700 fps. Figure 11 shows the overall configuration of the motor, and its location within the spacecraft is indicated in Figure 12.

The system design of the gas reaction control equipment, as mentioned previously, is based on the concept of the spin stabilized craft and provides control for correcting injection errors, longitude drift, eccentricity, spacecraft attitude and station keeping. The system utilizes two independent units; each one having a lateral (radial)

jet motor and axial jet motor as indicated in Figures 13, 14 and 15.

The axial motor is located at the apogee motor end of the solar array and its thrust vector is parallel to the spin axis of the spacecraft. By ground control, it may be operated in a pulsed mode or a continuous mode. The continuous mode of operation, shown in Figure 13, provides a force parallel to the spacecraft axis, and it is used in this mode to remove gross injection errors. The pulsed mode of operation is designed to provide the precession torque necessary to change the orientation of the spacecraft. In this mode the jet, Figure 14, is made to function over an arc of 60 deg. (or 90 deg.) of each revolution and in synchronism with the spacecraft spin rate. Pulsing in this manner results in an average torque vector to the spinning vehicle in accordance with Newton's second law of motion such that

$$\bar{T} = \frac{d}{dt} \bar{H} = \bar{\omega} \times \bar{H}, \text{ with } |\bar{H}| \text{ constant}$$

where \bar{T} = applied torque vector
 \bar{H} = angular momentum vector
 $\bar{\omega}$ = precession rate vector

Note: All vectors are mutually perpendicular in this technique of generating torque.

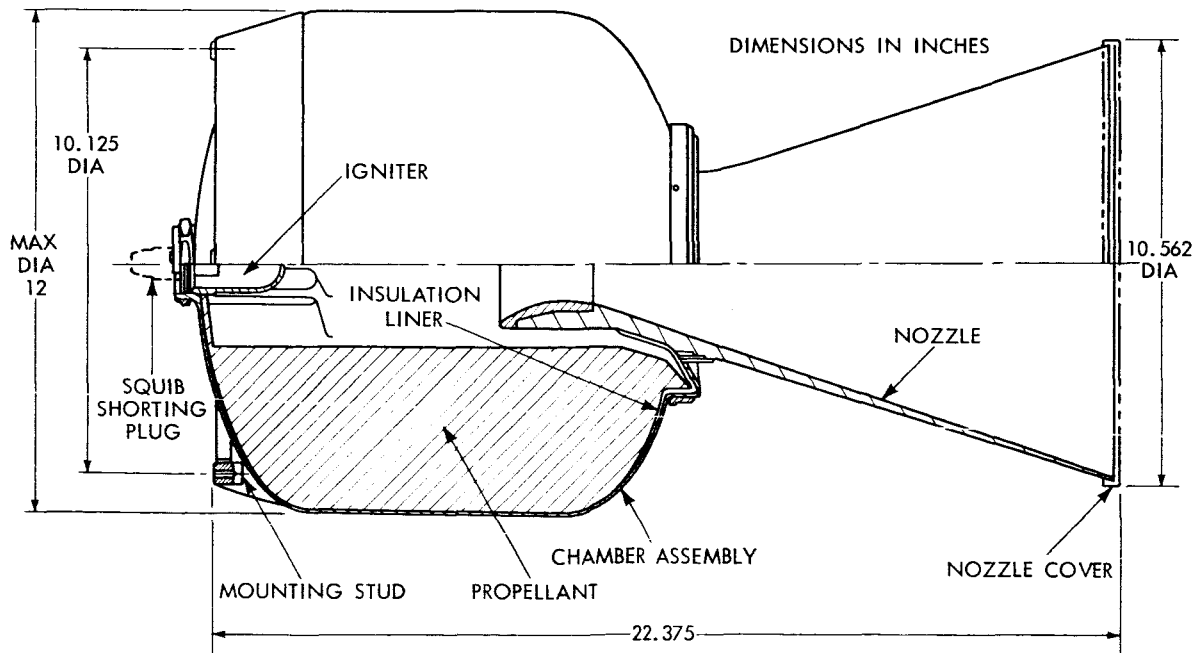


FIGURE 11.—Syncom II apogee injection motor.

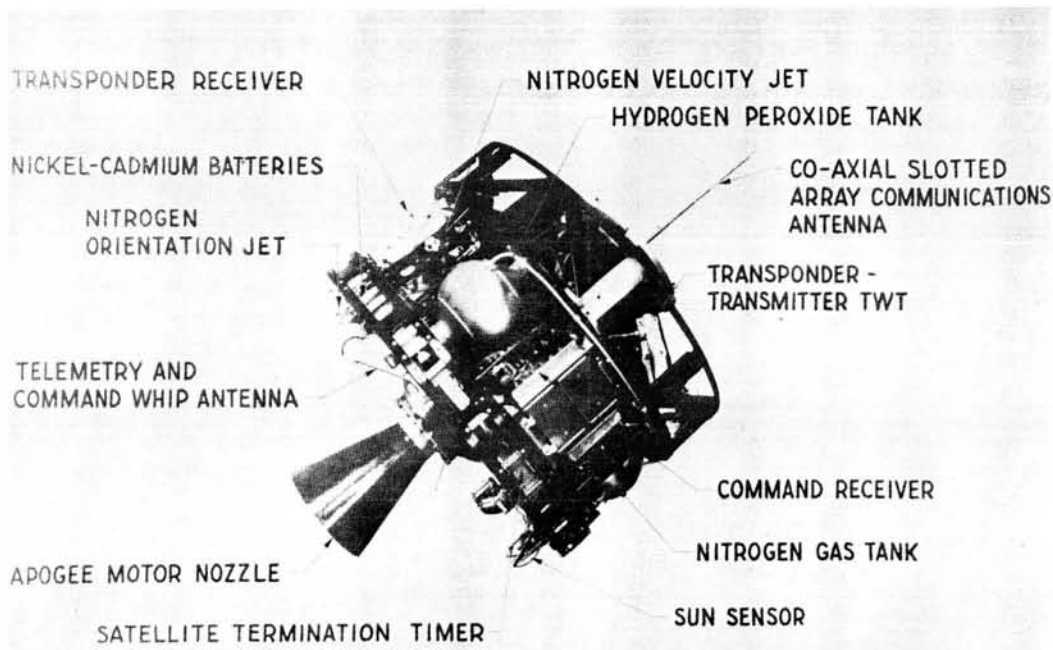


FIGURE 12.—Syncom

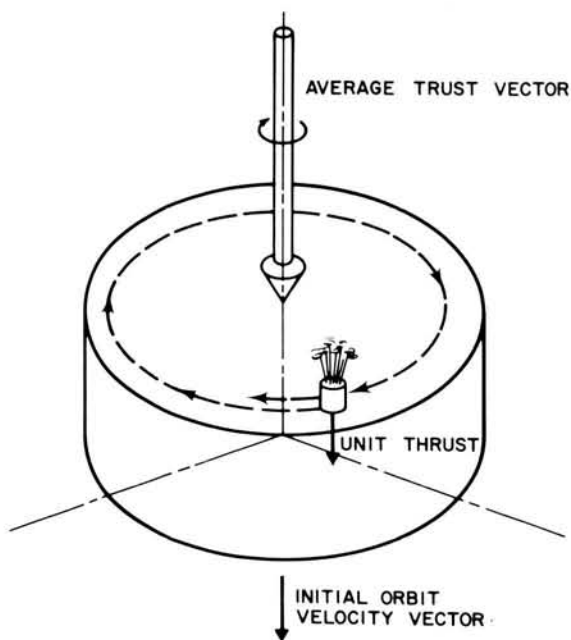


FIGURE 13.—Initial velocity correction.

The direction of spacecraft precession is accomplished by appropriate time phasing of jet pulsing in each revolution, and the magnitude of

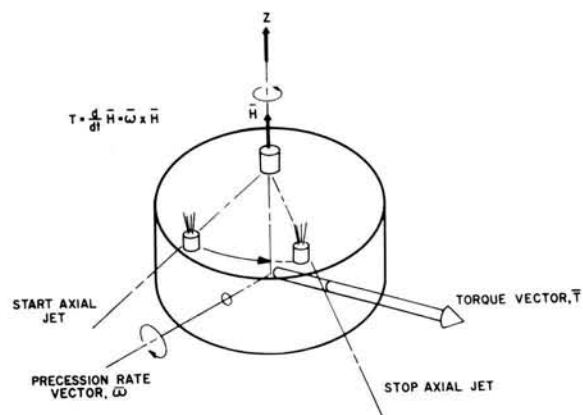


FIGURE 14.—Orientation control.

precession is dependent on pulse train duration. It is desirable to offset the axial jet motor as much as possible from the axis in order to increase torque.

The lateral (or radial) motor is located in the plane of the center of gravity, and its thrust center line passes through the axis of the spacecraft as shown in Figure 15. It operates only in the pulsed mode and is the motor used to make velocity corrections normal to the spacecraft spin axis. This jet is likewise controlled to operate over a

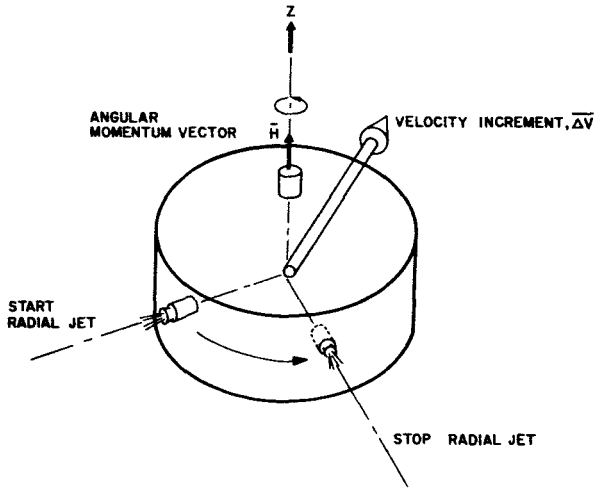


FIGURE 15.—Velocity control.

60 deg. arc (or 90 deg.) of each revolution and varied in phase during spin with relation to some reference such as the spacecraft-sun line.

One control unit of the system uses high pressure nitrogen. It was intended as the vernier control for attitude correction and velocity changes. The nitrogen unit could be used also for continuous thrusting with equal facility.

The second unit of the system was selected to be fueled with 90 percent pure hydrogen peroxide. Previous experience proved that the design of a hydrogen peroxide was feasible. Its main function was intended for continuous thrusting of the axial motor and as a back-up for the nitrogen for operation in the pulsed mode. The characteristic lag in rise and decay of its thrust was considered initially to be a problem.

The spherical tanks for the H_2O_2 unit are connected with fuel outlets at their outboard surfaces so that the advantage of centrifugal force is utilized and a bladder is not required. The in-board tankage surfaces are connected with a pressurization equalization line for the nitrogen which is used to expell the H_2O_2 . Figure 16 shows the general arrangement of the fuel supply and jet motors.

The initial design was configured to correct a 3σ ($\sigma=1$ standard deviation) error of both the launch vehicle and the apogee motor plus east and west (E-W) station keeping. As development proceeded and weight and space considerations for other areas increased, the space and weight

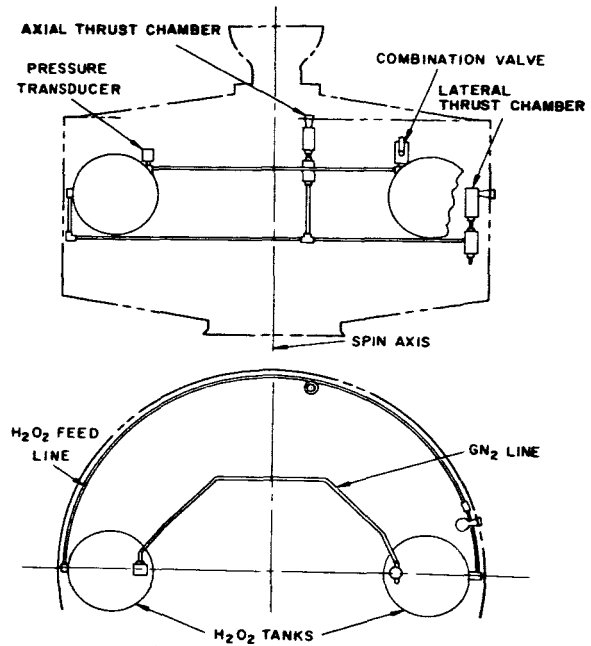


FIGURE 16.—Hydrogen peroxide control unit.

available for the two reaction control systems provided velocity increments of 297 fps for the H_2O_2 system plus 48 fps from the cold gas (N_2) system for a total capability of 345 fps. Estimated velocity incremental values for SYNCOM II are provided in Table 2.

TABLE 2.—Estimated Velocity Requirements for the Syncom II Reaction Control System

Maneuver	Velocity Increment (ΔV)
1. Correction of booster error (1)...	110 fps
2. Correction of apogee motor error (1)	24 fps
3. Control of drift.....	9.3 fps/deg/day
4. Reorientation of spacecraft spin axis	20.0 fps
5. Circularizing orbit (for small values of e)	$e \times 5000$ fps
6. E-W station keeping.....	7.0 fps/year

Estimated system performance of the H_2O_2 and the N_2 and hardware characteristics of the units are shown in Tables 3 & 4, respectively.

TABLE 3.—Estimated System Performance of the Hydrogen Peroxide and Nitrogen Units

	N ₂	H ₂ O ₂
Thrust (vacuum)		
Initial at 70 deg. F.....	0.8 lb _f	3.1 lb _f
Final at 70 deg. F.....	0.0 lb _f	1.2 lb _f
Specific Impulse.....	60.0 lb _f -sec	157 lb _f -sec
	lb _m	lb _m
Total Impulse.....	110 lb _f -sec	754 lb _f -sec
Thrust Coefficient.....	1.46	1.63

TABLE 4.—H₂O₂ and N₂ Design Characteristics

	N ₂	H ₂ O ₂
Fuel Weight, lbs....	1.82.....	4.9
N ₂ Pressurization, psia, lbs., (at 70 deg. F)	2500.....	200
Material.....	Tanks-titanium Tubes-stainless steel	High-purity 1060 Aluminum
Valve-Solenoid (actuated)	Stainless Steel (Class 321) with gold plated ball seal	1060 Aluminum with KEI-F valve poppet
Expansion ratio....	145/1.....	30/1

Stability Control

The spacecraft system design provided for the effect of nutation by including a nutation damper. Tumbling and severe despinning are prevented by taking special care to maintain a stable ratio of roll to pitch moment of inertia (greater than 1) and alignment of the jet motors.

The nutation damper is a linear fluid damper consisting of a fiberglass cylinder partly filled with mercury and mounted near the periphery of the craft with the axis of the cylinder parallel to the axis of the spacecraft. It is possible to design a similar damper, and with the careful application of dimension and amount of mercury, it will resonate at the desired angular frequency. However, tests of the scale model on a pendulum indicated that resonance is unnecessary to achieve satisfactory performance.

The ratio of roll to pitch moments of inertia ($\frac{I_R}{I_P}$) of the spacecraft was kept higher than 1.25 to 1, because if $\frac{I_R}{I_P}=1$, the spacecraft would

eventually start to nutate followed by tumbling. Dynamic balancing was accomplished to a tolerance such that the unbalance was 10 ounce-inches².

Both the H₂O₂ and N₂ jet nozzles were aligned in the spacecraft, which was mounted on an air bearing, so that the angular offset was no greater than 0.2 degrees.

Spin Speed and Attitude Determination

The use of the solar sensor as the primary means of determining spin speed, orientation and reference pulse for jet actuation has proven to be a simple, but most effective method. The solar sensor employed consists of a dual arrangement of silicon cells connected in parallel. The cells are mounted behind an appropriate mask which provides a narrow slit opening over the cells. The sun will energize the cells when it lies in the plane of the slit which also passes through the spacecraft spin axis. The slit provides a 3 db beamwidth of 0.8 ± 0.2 degrees and the resulting solar pulses of one millisecond (ms) duration are telemetered to the ground equipment for extraction and processing.

Figure 17 shows the arrangement of the solar sensors on the craft. A redundant pair of sensors (4A and 4B) is used for spin speed determination and, with the use of telemetry, the spin data from them is automatically used as an input to the ground synchronous controller. These are called psi (ψ) sensors. Spacecraft spin axis angle, ϕ , relative to the sun is effected through the placement of a redundant pair of ψ_2 sensors (5A and 5B) at angle of 35 deg. with the ψ sensors (4A and 4B). Figure 18 shows the orientation angle and sun observation geometry. The orientation of the spacecraft is derived from the difference in arrival time of the telemetered sun pulses. Where the arrival time is the same for both pulses, it is evident that the sun lies in the plane of both the ψ and ψ_2 slits simultaneously. In this case the sun line is normal to the spacecraft spin axis. Obviously then the greater the excursion from normal, the greater will be the time difference between pulses. For ease of determining the

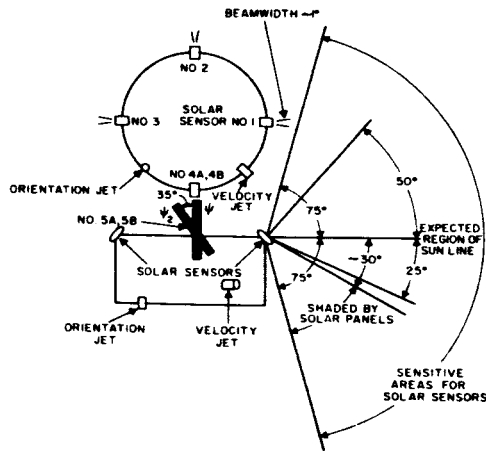
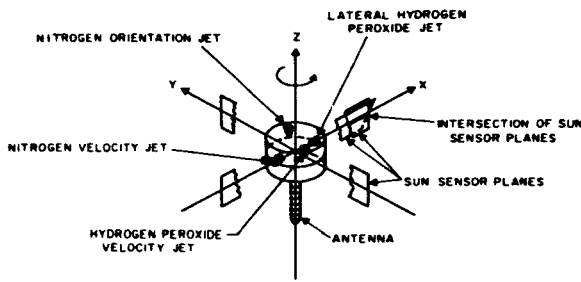


FIGURE 17.—Solar sensor and pulsed-jet system geometry.

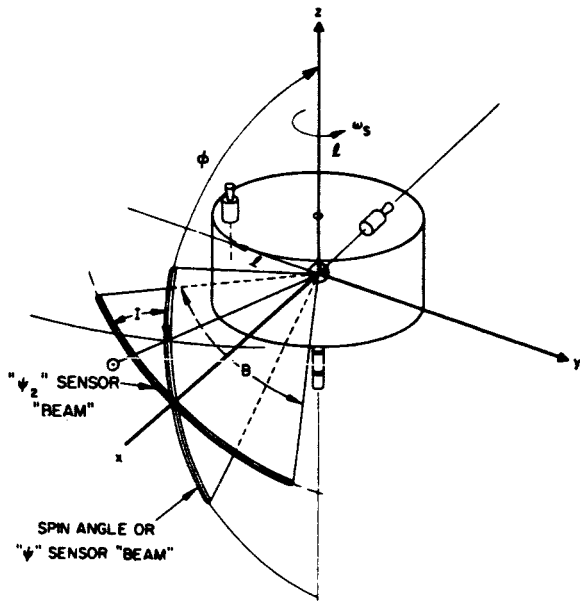


FIGURE 18.—Geometry of the attitude sensing and control system.

direction of tilt, the ψ and ψ_2 pulses have different polarities and amplitudes.

Sensors 1, 2 and 3 (mounted with their beam planes parallel to the spin axis) may be used in conjunction with 4A or 4B in a back-up mode for pulsing the axial or radial engines should the telemetry system fail to operate. This is referred to as the quadrant mode of operation since these four sensors are located 90 degrees apart and bear a known reference to each of the four jets. In operation, both the jet (axial or radial) and the quadrant are selected from the ground. A multi-vibrator connected to the desired sensor will actuate the jet upon each excitation of the sensor, and the spacecraft will precess (axial jet) or accelerate (radial jet) in a known direction to the sun-line.

Where solar sensors must be used exclusively, observations of attitude angle at least two weeks apart will establish the angle, ϕ , shown in Figure 18 with sufficient accuracy.

The relative positions of spacecraft and sun, in the case of SYNCOM II, permit accurate measurement of the right ascension of the spin axis using solar sensors. However, the more accurate measurement of declination is achieved through the determination of the antenna polarization angle.

Determination of polarization is a necessary adjunct to the sensor data in obtaining angle of the spin axis in the plane orthogonal to the plane containing the spin axis and the sun line. Using the spacecraft RF transmitted signal, whose electric vector is parallel to the spacecraft spin axis, the varying of the received signal level by the rotation of the ground receiving feed horn will indicate the relative angle of the incoming vector within 1 degree.

Ground Synchronous Control

The ground synchronous controller provides the desired initiation and duration of jet pulse command to the spacecraft to execute either precession or velocity maneuvers. The major sub-system in this equipment is the drum controller which is shown in Figure 19. This unit simulates the motion of the spinning craft, the solar sensors and a selected control jet using a rotating drum and adjustable contacts to fix the relationships between sensors, jets and spin angle of the spacecraft.

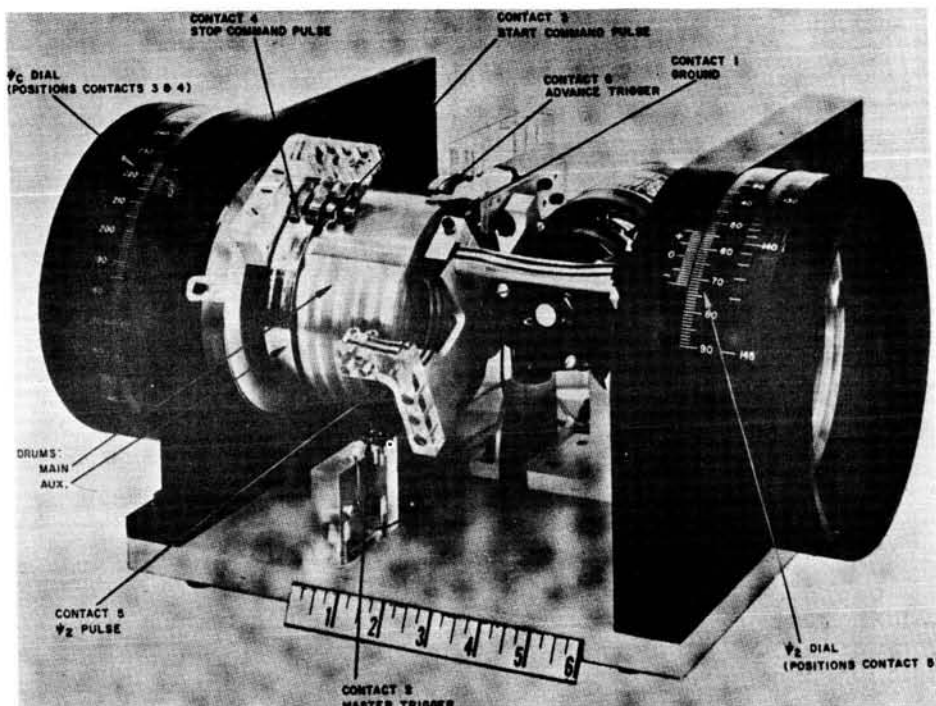


FIGURE 19.—Synchronous controller drum unit.

The drum is made to rotate at the same speed as the spacecraft using the pulses received from the solar sensor. The capability to adjust contacts on this device accommodates the requirement to locate precisely the initiation and cut-off points of the desired jet relative to the spacecraft-sun line during each revolution. The ground contact, number 1, provides electrical ground for the main and auxiliary drums. The master trigger (2) provides the drum synchronizing pulse on contact with the main drum. Jet command initiation occurs when No. 3 contacts with the main drum followed by the cut-off command when No. 4 contacts with the auxiliary drum. The ψ_2 pulse used with the ψ_2 pulses telemetered from the spacecraft occurs when contact No. 5 makes contacts with the main drum.

As a first measure, delays in signal transmission to and from the spacecraft and through the controls are calibrated out of the system. This is accomplished by transmitting a command pulse (not a jet actuation command) to the spacecraft and telemetering it back so that it can be compared on an oscilloscope with the received solar sensor, ψ , pulse. Rotation of the command dial

so that these pulses are coincident provides a direct measure of the delay. Delays in valve opening and closing are well established and are added to the propagation delay.

Assuming that delays are known and comparing the received ψ and ψ_2 pulses, the adjustment of ψ_2 contact dial, so that the pulses are coincident, yields the angle ϕ between the spacecraft spin axis and sun directly from the scale.

Actuation of the jet at the desired angle relative to the spacecraft-sun line is accomplished by adjusting the command dial to a value computed from orbital geometry and the parameters of the control system.

For precession or velocity corrections the required number of pulses is determined and translated into the required time for the maneuver, considering the spin speed extent of pulse duration, i.e. 60° or 90° of arc. The synchronous controller will phase the pulses for the length of time required by the closure of a manual switch.

V. ORBITAL MEASUREMENTS

Determination of Syncom orbits has been facilitated by the use of the range and range rate

system which was developed by the NASA—Godard Space Flight Center, Greenbelt, Maryland. The system is a two-way coherent communication technique utilizing the existing communications ground station power amplifier, antenna, ancillary electronic ground equipment and the frequency translation communications transponder in the spacecraft.

Since a direct relationship exists between the distance and time of wave propagation, advantage is taken of this to derive the range from the observed delay in signal transmission from the ground station to the spacecraft and return. Several side tones, equally separated, from 8 cps to 100 kc phase modulate the carrier transmitted from the ground. The spectrum is received by the spacecraft, translated in frequency, amplified and retransmitted. The retransmitted carrier and the spacecraft local oscillator (used as the beacon) are received on the ground and processed. The processed tone time delays correspond to increments of range.

The range rate determination is based on the Doppler principle, and is derived from the combination of the spacecraft beacon from the local oscillator, the transmitted carrier and the retransmitted carrier. The output from a Doppler extractor provides a two-way Doppler shift measured at that instant of time. These data, processed with antenna angle and time, determine the orbit. Although the orbit can be determined from one station, a more accurate location is obtained from two or more stations suitably displaced from each other.

Table 5 provides the half wave lengths of frequencies used in the determination of range.

Figure 20 shows the range and range rate system block diagram. Range tones of 100 kc, 20 kc, 4 kc, 800 cps, 160 cps, 32 cps and 8 cps phase modulate a carrier to produce sideband pairs. The phase delay of the 100 kc signal to and from the spacecraft provides the range measurement. Counting of a 10 mc repetition rate by the time interval unit is initiated by a reference data rate of 1 cps which is synchronized to the WWV time standard. Phase delay of the 10 kc tone corresponds to 1 count of the 10 mc or ± 15 meters resolution. Phase deviation of the 100 kc signal is maintained to less than ± 3.6 degrees of phase in order to provide the range accuracy. This is equivalent to knowing the positive going slope of the zero crossing of the wave within one percent.

The lower frequency tones are used to resolve the ambiguity, as the range increases, of the next higher tone to the whole cycle. Therefore, individual phase measurements of these tones are required to within ± 36 degrees.

The spacecraft transmits the translated spectrum (range tones) of the up-link carrier and the local oscillator beacon frequency. The first mixing operations are accomplished by the ground receivers which phase lock to the local oscillator and the translated up-link carrier. The range measurement is determined by detecting the range tone from the translated up-link carrier.

These two frequencies provide knowledge of the following: (1) the spacecraft local oscillator, (2) the one-way Doppler shift of the up-link carrier frequency, (3) the one-way Doppler shift of the up-link carrier minus the down-link carrier, (4) the one-way Doppler shift of the spacecraft oscillator drift, (6) the Doppler frequency.

The phased locked loop outputs of both receivers and the ground transmitted signal are supplied to the Doppler extractor which removes the effects of the satellite local oscillator drifts and produces the Doppler signal. Doppler shift is impressed on a 1 mc bias frequency to eliminate the ambiguity for negative Doppler values. This in turn is fed to a digital unit which counts the time interval at the rate of 10 mc. The bias frequency is removed to yield the Doppler frequency.

Measurements are made over a 10 minute period usually, and the data format is produced

TABLE 5

Half Wavelength $\left(\frac{\lambda = 299,792.5 + .4\text{km/sec}}{f}\right)$		
Frequency	km	Statute Miles
100 kc.....	1.498963	0.931
20 kc.....	7.498963	4.657
4 kc.....	37.47406	23.285
800 cps.....	187.3703	116.42
160 cps.....	936.8516	582.13
32 cps.....	4,684.258	2,910.6
16 cps.....	9,368.516	5,821
10 cps.....	14,989.63	9,314
8 cps.....	18,736.90	11,642.5

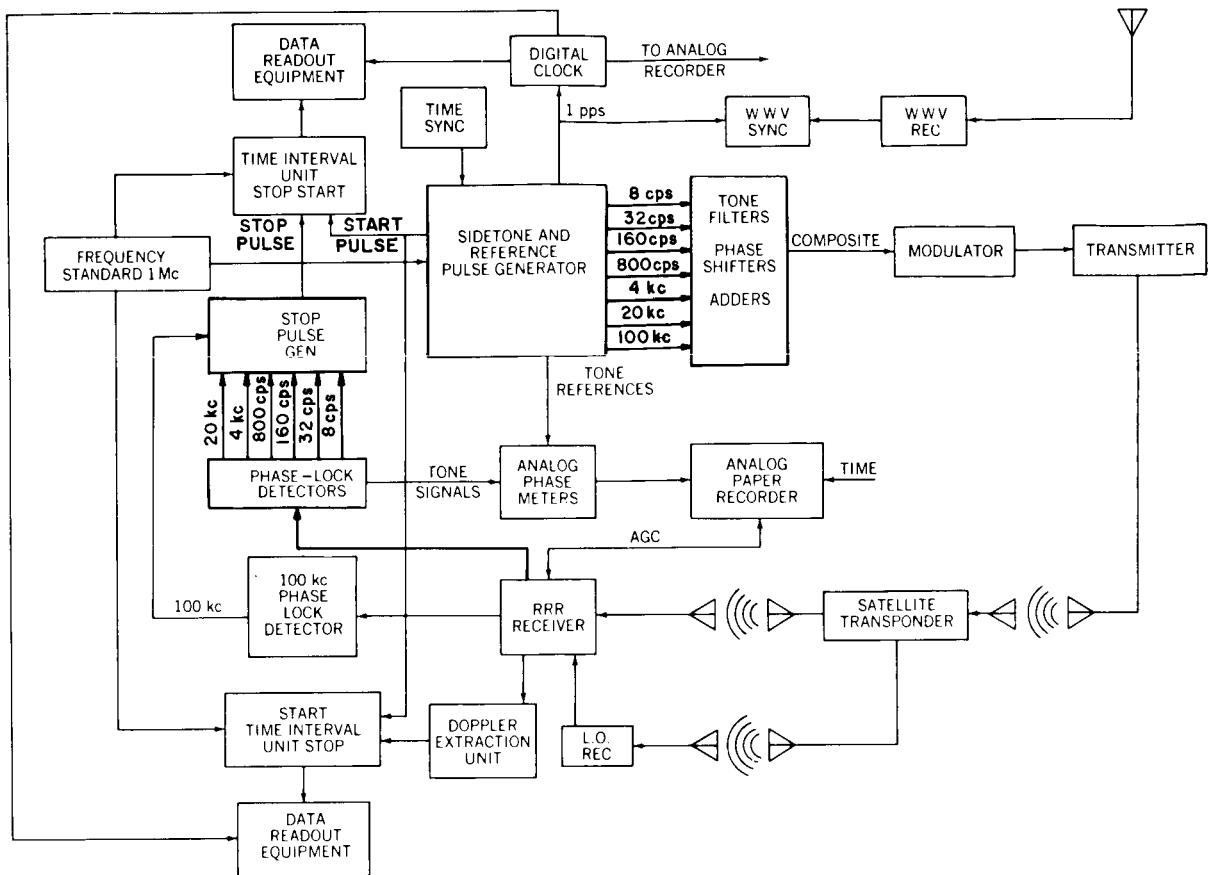


FIGURE 20.—Range and range rate system diagram.

by the system in punched paper tape for teletype relay to the GSFC. The data are processed first by the CDC-160 Computer which

1. Computes range in meters from the time of propagation or phase delay of the ranging tones to the spacecraft and return.
2. Computes range rate in m/sec from the time interval required to count "N" cycles of a standard frequency \pm the Doppler frequency of the carrier.
3. Reproduces the 30 foot communications antenna angle readout which the system measures to 0.1 degree.
4. Makes the necessary adjustments to time for the propagation time of the ranging tones to and from the spacecraft, and for the difference in time between the sampling of the time register and the range and range rate registers.

These data are transferred to cards or magnetic tape for orbit determination by the IBM 7090 or

7094 Computer. The standard output is an ephemeris tape providing position and velocity vectors corresponding to a time interval. The data are smoothed and calculations are made of the orbital elements and standard deviation.

Range and range rate data were taken by the communications ship at Lagos, Nigeria, 0.5 hours after lift-off while the spacecraft was in the transfer orbit. Four and one-half hours later the orbit was obtained within an accuracy of ± 20 meters. The standard deviation for SYNCOM II is ± 15 meters in range and 0.02 meter per second in range rate. Use of this system has resulted in establishing orbits with eccentricities less than 0.0001.

VI. SPACECRAFT CONTROL SYSTEM PERFORMANCE

A launch "window" was selected in order to maintain continuous sunlight while in the transfer orbit to maximize power output and maintain

apogee motor temperature within the desired range. The available power permitted range and range rate measurements during the transfer orbit and still assured sufficient battery power to provide for the high current needed to ignite the apogee motor and, afterwards, supply the axial jet solenoids current to obtain the desired drift orbit.

SYNCOM II, weighing 146.92 pounds, was placed in a Hohmann transfer ellipse by a THOR-DELTA launch vehicle on July 26, 1963. The resulting transfer orbit was within two-thirds of the standard deviation expected and inclined 33.16 degrees with the equator. Eccentricity was 0.732718. Perigee velocity was 33,569.97 fps and at apogee 4,712.1 fps.

The apogee injection motor was initiated by the on-board timer 5 hours, 25 minutes and 57 seconds after lift-off and burned approximately

20 seconds with a resultant spacecraft weight of 85.64 lbs. The apogee injection motor performance was 0.3 percent above nominal and injection occurred at latitude 21.83° E.

Although the apogee motor's performance improved the situation, the inertial velocity remained approximately 68 fps (0.6σ) below that required for obtaining synchronism. As a result of the lower altitude the spacecraft drifted eastward at a rate of 7.03° per day. Orbit eccentricity was 0.02844.

For convenience, Figure 21 shows the nominal trace of the spacecraft sub-satellite point.

Orbital Maneuvers

Approximately 24 hours after injection, the axial H_2O_2 jet was operated continuously for 140 seconds in order to change the drift from east to west. As a result of this maneuver, the drift was

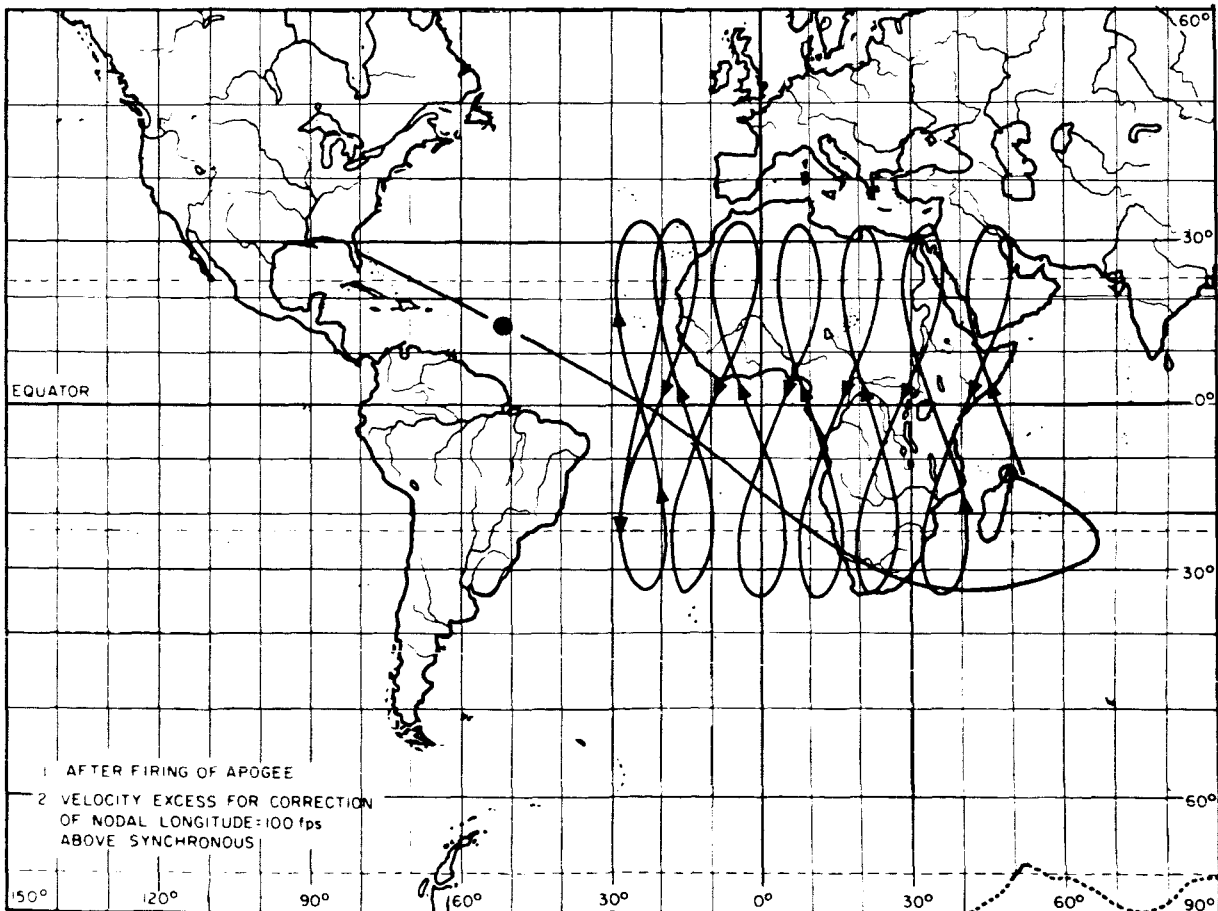


FIGURE 21.—Motion of subsatellite point.

westward at a rate of 4.53° per day. The following table compares the predictions with the actual results.

TABLE 6

July 27, 1963 19^h 10^m 00^s Z to 19^h 12^m 20^s Z, Nodal Longitude 37.4° E Hydrogen Peroxide Axial Jet (First Velocity Correction)

	ΔV fps	Drift Deg/day	e	ΔW Pounds	I_{sp} Seconds
Predicted	107.5	4.21	0.01329	1.80	157
Actual	109.8	4.53	0.01305	1.80	160

The second maneuver was designed to alleviate a problem of low temperature and power caused by the shallow angle of incidence of the sun line with the solar cell array. Two series of pulses of the H_2O_2 axial motor precessed the spacecraft spin axis so that the resulting sun line angle of incidence was approximately 10 degrees from the normal to solar array.

The following table (7) provides comparisons of the predicted parameters with the actual values for these precession maneuvers.

The third correction consisted of a velocity change using the H_2O_2 radial jet to reduce the westward drift rate prior to reaching the desired

TABLE 7

July 31, 1963 20^h 40^m 07^s Z, Nodal Longitude 20.6° E Hydrogen Peroxide Axial Jet (First Attitude Correction)

First Pulse Train						
	Pulsing Time, Sec.	No. of Pulses	Precession Angle-Deg.	Precession Phase-Deg.	Initial Sun Angle Degrees	Final ϕ_F Sun Angle Degrees
Predicted	56	135	30	114.3	64.0	76.4
Actual	56	135	30	123	64.3	80
Second Pulse Train						
Predicted	87	211	54	113.2	79.5-80.5	101.2
Actual	87	215	53	112.5	79.5-80.5	100.1

TABLE 8

August 11, 1963 06^h 00^m 00^s to 06^h 14^m 13^s Z Nodal Longitude 50.0° W Hydrogen Peroxide Lateral Jet (Second Velocity Correction)

	Pulsing Time Seconds	ΔV fps	Resulting Drift, °/orbit	ΔW lbs.	I_{sp} Seconds	Total No. of pulses
Predicted	395	44.06	2.0	1.05	111	938/906
Actual	395	*39.5 37.5	2.67	0.91	106	894

*Based on Pressure Drop for Actual ΔW .

55° W longitude. Table 8 provides the predicted and actual values for this maneuver.

In order to demonstrate the accuracy of the reaction controls, it was decided that another attempt should be made to modify the rate to achieve the experiment objective of a nodal crossing at 55° W. The drift of 2.67 degrees per orbit was the wrong rate to obtain the target nodal

longitude precisely at 55 degrees W. This velocity correction would be made with the nitrogen jet for the first time. The N₂ pressure had dropped from 2500 psia to 2100 psia by this time due to leakage. Following the orbit determination the next day, the lateral nitrogen motor was used for the maneuver. Results of this correction are provided in Table 9.

TABLE 9

August 12, 1963 05^h 45^m Z to 06^h 14^m 19^s Z Nodal Longitude 51.5° W Nitrogen Lateral Jet
(Third Velocity Correction)

	Pulsing Time Seconds	ΔV fps	Drift, Deg/orbit	ΔW lbs	I_{sp} Seconds	i Degrees
Predicted.....	510	13.75	1.189	0.495	(74 × 0.9549) = 70.66	33.12
Actual.....	510	12.8	1.31	0.485	67 to 72	33.12

There are minor discrepancies between calculations for ΔV and ΔW . These errors were comparable to effects which would be caused by calibration errors. The correction factor 0.9549 applied to the specific impulse is a geometrical correction arising from thrusting over a 60 degree arc of each revolution.

The resulting drift rate was still too high, but no attempt was made to reduce it prior to reaching 55 degrees W longitude. A subsequent plan was formulated in order to obtain sufficient range and range rate data, determine the orbit exactly and then make a second velocity correction aimed to cause a slight eastward drift. The purpose of

a 0.05 degree per orbit eastward bias was to compensate for the estimated force effect of triaxiality which would overcome the eastward drift and cause the spacecraft to stop and, gradually, drift westward again. At the time this determination was made, other calculations indicated that velocity corrections should be made approximately at the end of each 40-day period amounting to a 12 fps per year correction. The results of the third velocity correction (second Nitrogen operation) are provided in Table 10.

A fourth velocity correction was needed to reduce the eastward drift, and the results of this correction are provided in Table 11.

TABLE 10

August 15, 1963 05^h 20^m 00^s Z to 06^h 06^m 22^s Z Nodal Longitude 55.35° W Nitrogen Lateral Jet
(Fourth Velocity Correction)

	Pulsing Time Seconds	ΔV fps	Resulting Drift Deg/orbit	ΔW lbs	I_{sp} Seconds	i Degrees
Predicted.....	924	12.72	-0.05	0.471	71.5 × 0.9549 = 68.18	33.13
Actual.....	924.	13.97	-0.17	0.502	70.6	33.13

TABLE 11

August 16, 1963 15^h 44^m to 15^h 46^m Z Nodal Longitude 55.1° W Nitrogen Lateral Jet (Fifth Velocity Correction)

	Pulsing Time Seconds	ΔV fps	Resulting Drift Deg/orbit	ΔW lbs	I_{sp} Seconds	i Degrees
Predicted.....	120	1.33	-0.025	0.0488	$72.0 \times 9549 = 68.7$	33.15
Actual.....	120	-----	-0.085	-----	-----	33.15

The nodal longitude achieved was 0.03° greater than the target of 55° W.

Due to uncertainty of measuring small increments of tankage pressure the data for ΔW , ΔV , and I_{sp} are not shown. At this time the magnitude of the velocity correction relative to the spacecraft inertial velocity was 1 to 10,000. The drift eastward of 0.085 deg. per day was still too high, so another attempt to reduce this rate was in order. Care was exercised in selecting the time of day for making the proposed correction so that the maximum drift correction could be made with the least expenditure of nitrogen. In

addition to this maneuver, an 8.5 degree attitude change was necessary to improve the antenna beam angle. Communications between the ship at Lagos, Nigeria, and Lakehurst, New Jersey, were degraded during part of each orbit as indicated by a drop in the received signal strength at the Lakehurst station. The two corrections were combined into one maneuver using the axial jet, and conditions were excellent for using the quadrant mode of jet initiation for the first time. Effects of the combined maneuver on spacecraft drift are provided in Table 12.

TABLE 12

August 18, 1963 01^h 30^m to 01^h 31^m 35^s Z Nodal Longitude 55.03° W Nitrogen Axial Jet (Second attitude correction using quadrant mode of operation)

	Pulsing Time Seconds	ΔV fps	Resulting Drift Deg/orbit	ΔW lbs.	I_{sp} Seconds	Attitude Change	i Degrees
Predicted.....	95	1.58	-0.070	0.0581	$68 \times .9003 = 61.22$	8.5	-----
Actual.....	95	-----	-0.037	-----	-----	8.5	33.05

In this case, the quadrant mode of operation pulses the jet over a 90° arc instead of 60 degrees, therefore, the geometric correction factor is 0.9003.

The results of the attitude correction indicated that a velocity correction was induced; however, the areas of uncertainty prevented confident determination of ΔV , ΔW and I_{sp} . The attitude change occurred as predicted, and the attitude of the spacecraft after reorientation maneuver was within 1° of the desired position.

The resulting eastward drift declined gradually due to the natural orbital perturbations, primarily the effect of triaxiality, and approximately September 9, 1963, the eastward drift stopped and a few days later was observed in a westward drift orbit. At this time the eccentricity was .00015.

The next maneuver was designed to circularize and synchronize the orbit at 59 degrees W longitude. Table 13 provides the results of this correction.

TABLE 13

November 28, 1963 00^h 24^m 00^s to 00^h 55^m 00.5^s Z Nodal Longitude 59.1° W (Eccentricity: 0.00025). Hydrogen peroxide Lateral Jet (Sixth velocity correction combined with an eccentricity correction)

	Pulse Time Seconds	ΔV fps	Drift Deg/orbit	Eccentricity	ΔW	I_{sp} Seconds	i Degrees
Predicted.....	31.5	2.79	0	0.0	.06	111	-----
Actual.....	31.5	2.52	0.027	0.00005	.08	89	32.92

In order to use one maneuver to obtain both an eccentricity and a velocity correction, the average velocity increment was applied at 65.4 deg. to the existing velocity vector. The radial component = 2.54 fps and the tangential component = 1.055 fps.

In anticipation of the next Syncom launch it was desirable to shift the spacecraft to a position which would eliminate communication interference with

the new spacecraft by having the largest difference in the two spacecraft look angles from a ground station antenna. To meet this the next maneuver imparted a westward drift rate such that, on the anticipated launch date of 5 May, 1964, the spacecraft would arrive at 140 deg. W longitude. Table 14 provides the results of this velocity correction.

TABLE 14

March 18, 1964 02^h 31^m 30^s to 02^h 34^m 00^s Z Nodal Longitude 67.0° W Hydrogen Peroxide Lateral Jet (Seventh velocity correction)

	Pulse Time Seconds	Δv fps	Resulting Drift Deg/orbit	ΔW lbs.	I_{sp} Seconds	i Degrees
Predicted.....	150	12.52	1.52	.28	95	-----
Actual.....	150	10.50	1.307	.28	94	32.68

The anticipated launch of May 5 was changed to August; however, it was deemed desirable to reduce the westward drift rate of the spacecraft

to fix arrival at 180 deg. W longitude on July 15, 1964. The results of this maneuver are provided in Table 15.

TABLE 15

April 25, 1964 02^h to 02^h 01^m 10^s Z Nodal Longitude 116.9° W Hydrogen Peroxide Lateral Jet (Eighth Velocity Correction)

	Pulse Time Seconds	ΔV fps	Resulting Drift Deg/orbit	ΔW lbs.	I_{sp} Seconds	i Degrees
Predicted.....	70	4.90	.785	.14	92	-----
Actual.....	70	4.79	.832	.14	88	32.60

The spacecraft arrived at 143.6 deg. West longitude on May 28, 1964, and since the spacecraft was drifting faster than desired as a result of the last correction, it would arrive at 180 deg. W longitude prior to the control ship's arrival on

station. Therefore, a correction to reduce drift was planned for July 3 which would provide arrival at 161 deg. East longitude on September 15, 1964, for direct support of the SYNCOM C launch. The results are provided in Table 16.

TABLE 16

July 3, 1964 02^h 00^m 00^s to 02^h 00^m 49^s Z Nodal Longitude 166° W Hydrogen Peroxide Lateral Jet
(Ninth Velocity Correction)

	Pulse Time Seconds	ΔV fps	Resulting Drift Deg/orbit	ΔW lbs.	I_{sp} Seconds	i Degrees
Predicted.....	49	3.02	.43	.09	82	-----
Actual.....	49	2.5	.49	1.06	103	32.54

The pressure vs time histories for the hydrogen peroxide and nitrogen tankage are provided in Figures 22 and 23, respectively.

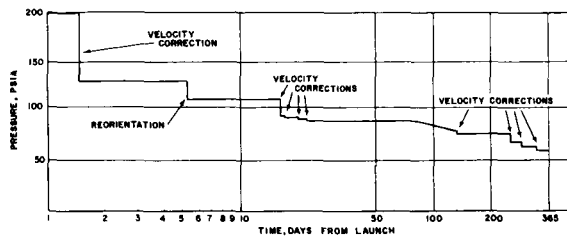


FIGURE 22.—Hydrogen peroxide unit pressures.

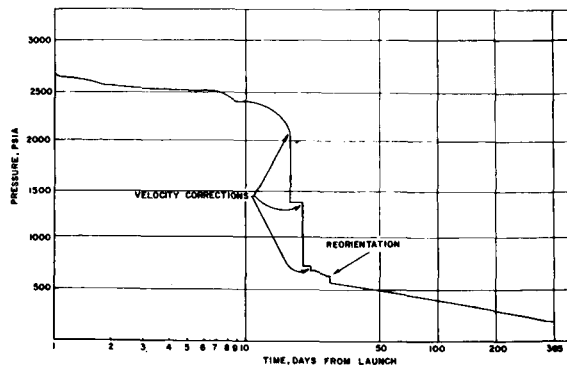


FIGURE 23.—Nitrogen unit pressures.

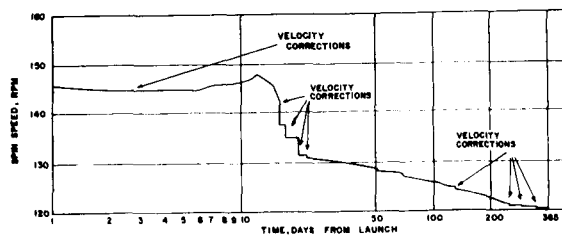


FIGURE 24.—Spacecraft spin speed.

Spacecraft Spin Speed

The spin speed history, shown in Figure 24, has provided some interesting phenomena. At launch the spin rate was 146 rpm and increased 0.2 rpm due to apogee motor firing. After the fifth day, the spacecraft spin rate increased and reached 148 rpm on the tenth day. Associated with this increase was a sharp decrease in the N₂ pressure. A nitrogen leak could cause this change, but it cannot be ascertained definitely that this induced the speed increase. Changes in spin rate resulting from maneuvers have been examined in detail and the effects of mass transfer of the fuel on reduction of speed have been verified.

Station keeping will be provided by the H₂O₂ unit by a lateral jet pulsing correction every two months. This has the effect of also reducing the spin speed approximately 0.1 rpm for each correction.

Drift Observations

The determination of the earth's "triaxiality" force and longitude orientation is an important factor in the design of station keeping control systems for spacecraft where several years of operability in orbit are required. The advantages to be gained from the system designer's point of view and operationally are the ability to predict the spacecraft drift at synchronous altitudes and determine exact fuel requirements for station keeping a specified number of years at the desired longitude.

C. Wagner's preliminary report on the determination of the J_{22} and λ_{22} coefficients in the earth's gravity potential is based on limited acceleration data at two longitudes. These observations were obtained at 59.15 deg. W and 54.76

deg. W with accelerations of -1.32×10^{-3} deg/day² and -1.27×10^{-3} deg/day², respectively. The calculated curve established by these points yields $J_{22} = -(1.70 \pm .05) \times 10$ and $\lambda_{22} = 19 \pm 6$ deg. west. It is determined from the J_{22} value that the difference between the major and minor equatorial radii of the "triaxial ellipsoid" is 212 ± 6 ft.

VII. CONCLUSIONS

Launch Vehicle

The performance of the Delta launch vehicle was excellent. Its third stage placed the spacecraft in the transfer orbit within one standard deviation (1σ) of the nominal altitude and orbital energy required.

Spacecraft Subsystems

(a) *APOGEE MOTOR*—The apogee motor functioned perfectly. Its performance was 0.3 percent above its nominal velocity increment, and was within one standard deviation.

(b) *APOGEE MOTOR IGNITION TIMER*—The timer functioned within the prescribed limits. Its time of function was set for 320 minutes and it fired the motor in 325 minutes and 57 seconds which was well within the ± 3 percent tolerance.

(c) *SUN SENSORS*—These sensors performed in accord with the design specifications.

(d) *HYDROGEN PEROXIDE UNIT*—The performance of this unit surpassed its intended use. In addition to providing coarse correction capability, the unit performed equally as well as the nitrogen unit in making vernier corrections.

Hydrogen peroxide did not decompose when stored at the synchronous altitude as the pressure remained unchanged between maneuvers.

Specific impulse predictions of this unit were verified as 157 seconds in the continuous mode and 111 seconds in the pulsed mode.

Future Syncom type craft will substitute another H_2O_2 unit in lieu of the N_2 . Its performance resulted in the selection of redundant of H_2O_2 units in Syncom C, EARLY BIRD and spin-stabilized series of Advanced Technological Satellites.

(e) *NITROGEN UNIT*—The specific impulse of nitrogen was verified in the pulsed mode operation as 60 seconds.

Problems with leakage and low specific impulse categorized the use of nitrogen as unattractive for primary orbit control for this spacecraft.

(f) *SPIN & JET ALIGNMENT*—Mass transfer effects due to fuel caused the major changes in spin rate and their calculated effects have been verified. Although jet nozzles were accurately aligned on air bearing tables the slight tolerance on their alignment was sufficient to contribute a minor influence on the spin rate. Intentional misalignment of nozzles in the roll plane appears advantageous in producing a bias for compensating the despin associated with conservation of momentum during fuel transfer.

The spin rate will decay on the order of 0.1 rpm per month, due to magnetic damping, now that the nitrogen is expended and the hydrogen peroxide is nearly depleted.

Spacecraft Attitude

The orientation of the spacecraft spin axis in relation to the sun line will change annually by $\pm 22^\circ$ but will remain fixed relative to the earth's spin axis.

Range and Range Rate

This new system, developed by the Goddard Space Flight Center, performed better than expected. Range measurements to the spacecraft were estimated to be within the 15 meters of the actual distance from the measuring station. Range rate measurements resulted in 0.02 meters per second accuracy as compared with the design goal of .05 meters per second. This equipment aided in the fine control of the spacecraft such that an eccentricity of .00005 was obtained on November 28, 1963.

GROUND CONTROL EQUIPMENT—The ground synchronous controller performed as its design intended.

DRIFT OBSERVATIONS—Long term data obtained from SYNCOM II drift should have significant geodetic value compared with lower orbit data.

In conclusion, (a) the technique of orbital placement employing the transfer ellipse and apogee kick motor has been successfully demonstrated. (b) The on-board reaction control system, in conjunction with its complementary

ground equipment, has demonstrated the ease of maneuverability of the spin stabilized spacecraft.

(c) The design goal of one year's operation in orbit has been achieved.

REFERENCES

- "Syncom System Summary," June 1962—NO—SSO—2252R Contract NAS-5-1560, by Hughes Aircraft Company.
- "Syncom II Performance During First 60 Days in Orbit," Nov. 1963—NO. SSO 3103R by Hughes Aircraft Company.
- "Dynamic Analysis and Design of the Synchronous Communication Satellite" by D. D. Williams NO. RDL/TM 649, May 1960.
- "Advanced Syncom Launch and Orbital Analysis"—October 1963, NO. SSD-31078R—Contract NAS5-2797 by Hughes Aircraft Company.
- "Multiple Mission Advanced Syncom"—Jan. 31, 1964—NO. 64 H-0244/A3498—SSO 31265P by Hughes Aircraft Company.
- "Determination of the Triaxiality of the Earth from Observations of the Drift of the SYNCOM II Satellite" by C. A. Wagner, Goddard Space Flight Center, Greenbelt, Maryland—April, 1964.
- "The Range & Range Rate System and Data Analysis for SYNCOM I (1934A) by H. W. Shaffer, W. D. Kahn, W. J. Bodin, Jr., G. C. Kronmiller, P. D. Engels, and E. J. Habib, Goddard Space Flight Center Greenbelt, Maryland.
- "Space Research Vehicle (DELTA) Model Specification DSV-3C," August, 1963 DS-2400 by Missile and Space Systems Division, Douglas Aircraft Company.

APPLICATION OF INTEGRATED CIRCUITS TO TELEMETRY SYSTEMS*

J. MICHAEL BALDERSTON

In June 1962, a continuing study of integrated circuits as applied to scientific satellite telemetry systems was initiated at Goddard Space Flight Center. This paper describes some results of investigations of standard commercially available Texas Instruments Series 51 Solid Circuits. The possibility of altering circuit functions by the addition of external components was verified. Three telemetry encoder subsystems were designed using the solid circuits. Conclusions and recommendations for future studies are presented.

INTRODUCTION

In recent years the continuing search for electronic circuitry of smaller size and increased reliability has brought about the field known as molecular electronics, or integrated circuitry. In one phase of this field, individual components are replaced by areas within a solid state material; these areas are generally produced by means of a planar diffusion technique that is a recent refinement of transistor technology. Interconnections are made by means of deposited leads. Thus, a single block of semiconductor material may replace a complete circuit such as a flip-flop, an IF stage, or an amplifier. Because of tolerance problems and various other difficulties, most commercially available integrated circuits have been digital (flip-flops, gates, inverters, etc.); however, several companies are investigating audio and RF circuit applications.

The particular integrated circuits used in the evaluation discussed here were the Texas Instruments (TI), Inc., "Solid Circuits." When the investigations were started (June 1962), the only circuits available in quantity were these and the Fairchild semiconductor "Micrologic" circuits; the TI circuits were chosen on the basis of power dissipation (2 mw versus 15 mw per circuit) and package configuration (the TI package lends itself more easily to welded, cordwood construction).

The circuits investigated were: a flip-flop (Type 510 or 511) which, by different external connec-

tions, could be used as a counter-divider, a shift register, or a set-reset stage; a six-input NOR gate (Type 512 or 513); a circuit containing two three-input NOR gates (Type 514); and an Exclusive-Or circuit (Type 515).

TIMING CIRCUITS

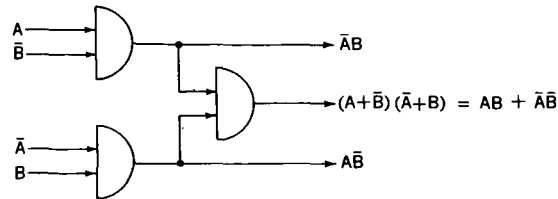
Astable Multivibrator

In any telemetry system involving time-division multiplexing, a timing source is required. For purposes of miniaturization and for compatibility with other parts of a telemetry system realized with integrated circuits, the astable multivibrator shown in Figure 1 was built. It employs one Type 515 Solid Circuit, two capacitors (C), and two resistors (R). Note that by providing the appropriate signals to the unused gate inputs, the astable multivibrator could be synchronized by an external reference, or gated to run or stop. Selection of R and C gives different values of frequency. It was found that C had to be greater than $0.002\mu\text{f}$ for proper operation; R could be varied from 5 K to open. For $R = 120\text{ K}$ and $C = 0.01\mu\text{f}$, the output frequency was approximately 800 cps at room temperature and with a 3 volt supply. The temperature stability was poor, about 0.1 percent per $^{\circ}\text{C}$; this was attributed to the fact that the internal base resistors were of semi-conductor material.

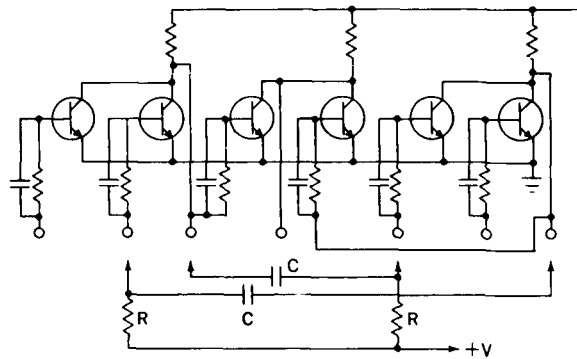
Monostable Multivibrator

In any logic system, it occasionally becomes necessary to generate a fixed delay period. This

*Published as *NASA Technical Note D-1905*, July 1964. This work has also been reported at the NASA Microelectronics symposium, Langley Research Center, July 17, 1963.



Logic diagram of Type 515 Exclusive-Or circuit.



Circuit diagram of an Exclusive-Or circuit and additional components.

FIGURE 1.—Astable multivibrator.

is sometimes done by gating of the system clock; however, it is often simpler to use a monostable multivibrator, which generates a fixed duration pulse each time it is triggered. The monostable

multivibrator of Figure 2 was built with one type 510 Solid Circuit, two resistors, two capacitors, and one diode. At room temperature, and with a 3 volt supply, the following values of output pulse duration resulted from varying C_T :

$t = 40 \mu\text{sec.}$	$C_T = 0.01 \mu\text{f}$
$310 \mu\text{sec.}$	$0.1 \mu\text{f}$
3.0 msec.	$1.0 \mu\text{f}$

The trigger sensitivity was 2 volts.

Elimination of External Components

In order to make production economically feasible TI, Fairchild, and now others have concentrated on a standard line of "building blocks." At the other extreme, Motorola is concentrating on custom circuits. A happy medium, now being introduced by TI and General Electric, is a standard substrate used as a "parts bin," which allows a variety of circuit configurations merely by the changing of internal and external interconnections. The astable and monostable circuits discussed above, for instance, could be reduced to the point where the only externa

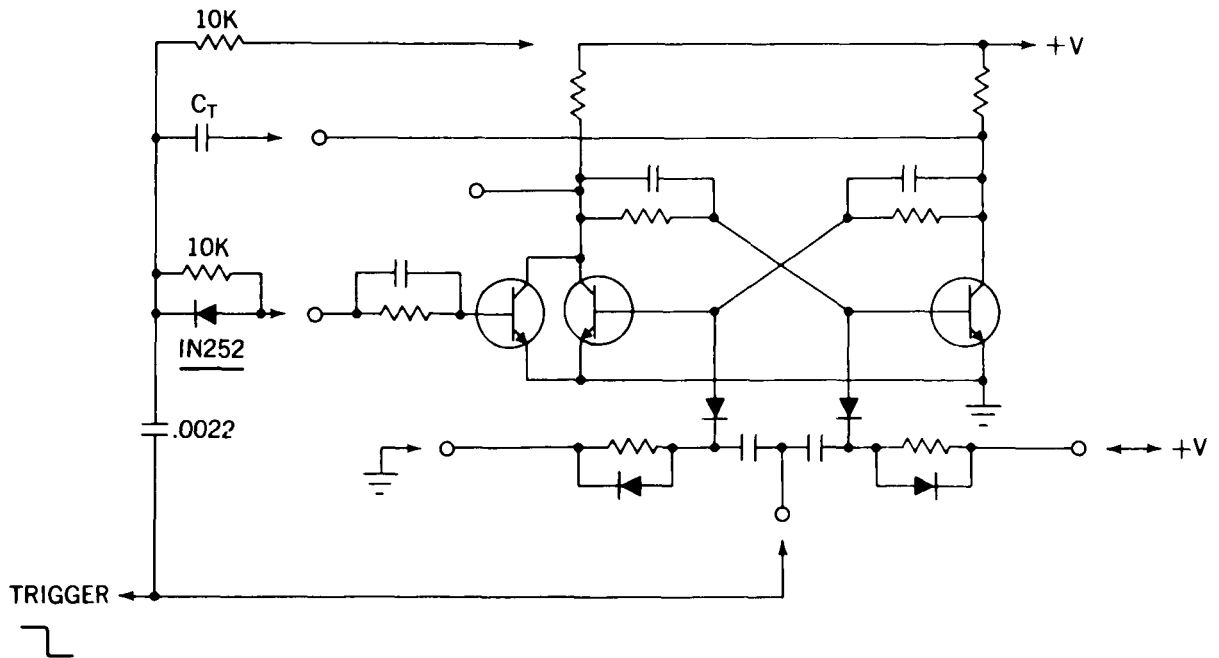


FIGURE 2.—Monostable multivibrator built with one type 510 solid circuit, two resistors, two capacitors, and one diode.

elements required would be the timing capacitors. A Schmitt Trigger circuit, the construction of which was unsuccessfully attempted with the standard Solid Circuits, could also be realized.

CLOCK READOUT

In past satellites that used the PFM (Pulse Frequency Modulation) encoding process there was no provision for an on-board clock readout. Such a readout would be useful in reducing data from certain experiments, such as those in which the time of an event of interest and in which satellite time rather than ground station time is of prime importance. For one application, it was desired to provide a 22-bit sequence counter and readout for use with an Explorer XII-type Telemetry Encoder. The resulting design was never breadboarded because the project for which the application arose was later shelved; however, the block diagram of the proposed counter is included here (Figure 3) to demonstrate the scale of subsystems which may readily be implemented using the Solid Circuits.

PFM CLOCK AND MATRIX

The clock and matrix for the Explorer XII series PFM encoders supplies timing signals to the rest of the telemetry and to the experiments. It consists merely of a frequency source, a count-down chain, and a number of gate circuits. A block diagram of such a system is shown as Figure 4. The output functions shown are the particular ones used in Explorer XIV, and obviously could be changed by a different gating arrangement. The only unusual feature is the function of gates 3A, 3B, 4A, 4B, and 9. The explanation for their presence is as follows.

In the PFM scheme currently in widespread use, frame synchronization is aided by width modulation of the first channel frequency burst. In particular, the first burst is 15 msec in duration, the first blank is 5 msec, and all other blanks and bursts are 10 msec. Existing encoders provide a signal to accomplish this width modulation by changing the bias on a magnetic core oscillator for one channel period. On the other hand, the Solid Circuit encoder accomplishes the width modulation by means of the aforementioned gates

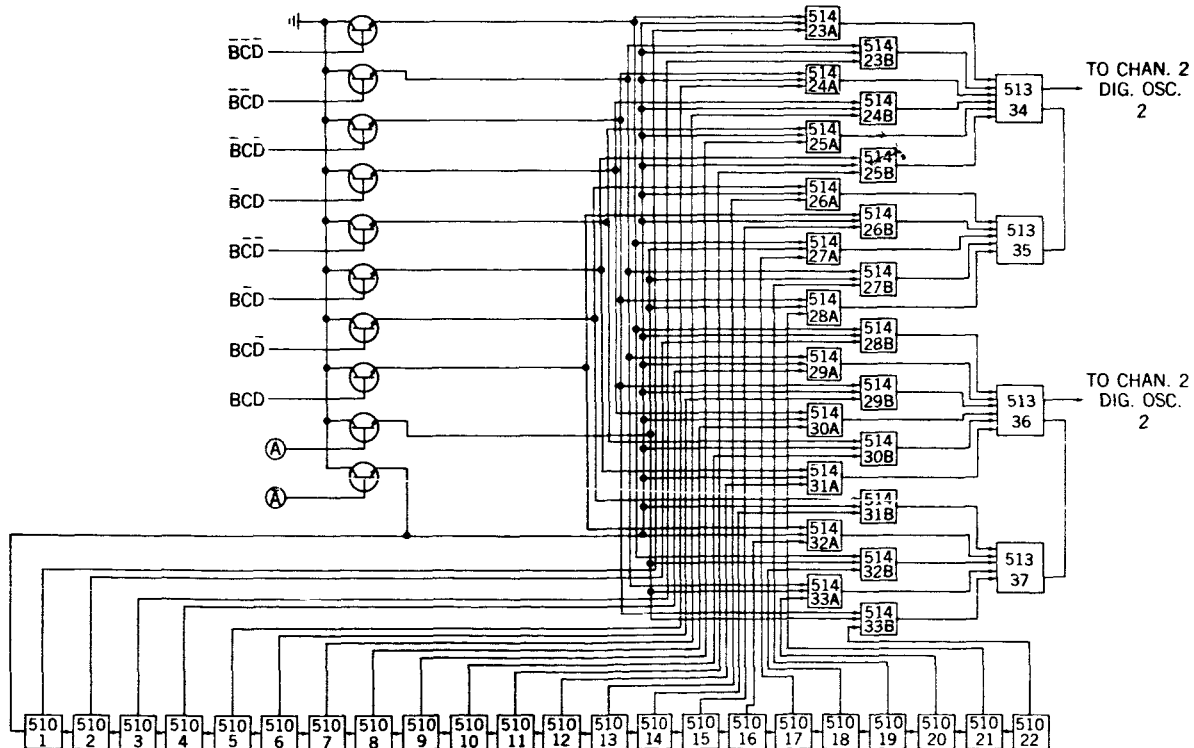


FIGURE 3.—Block diagram of a clock accumulator and matrix designed with T1 solid circuits.

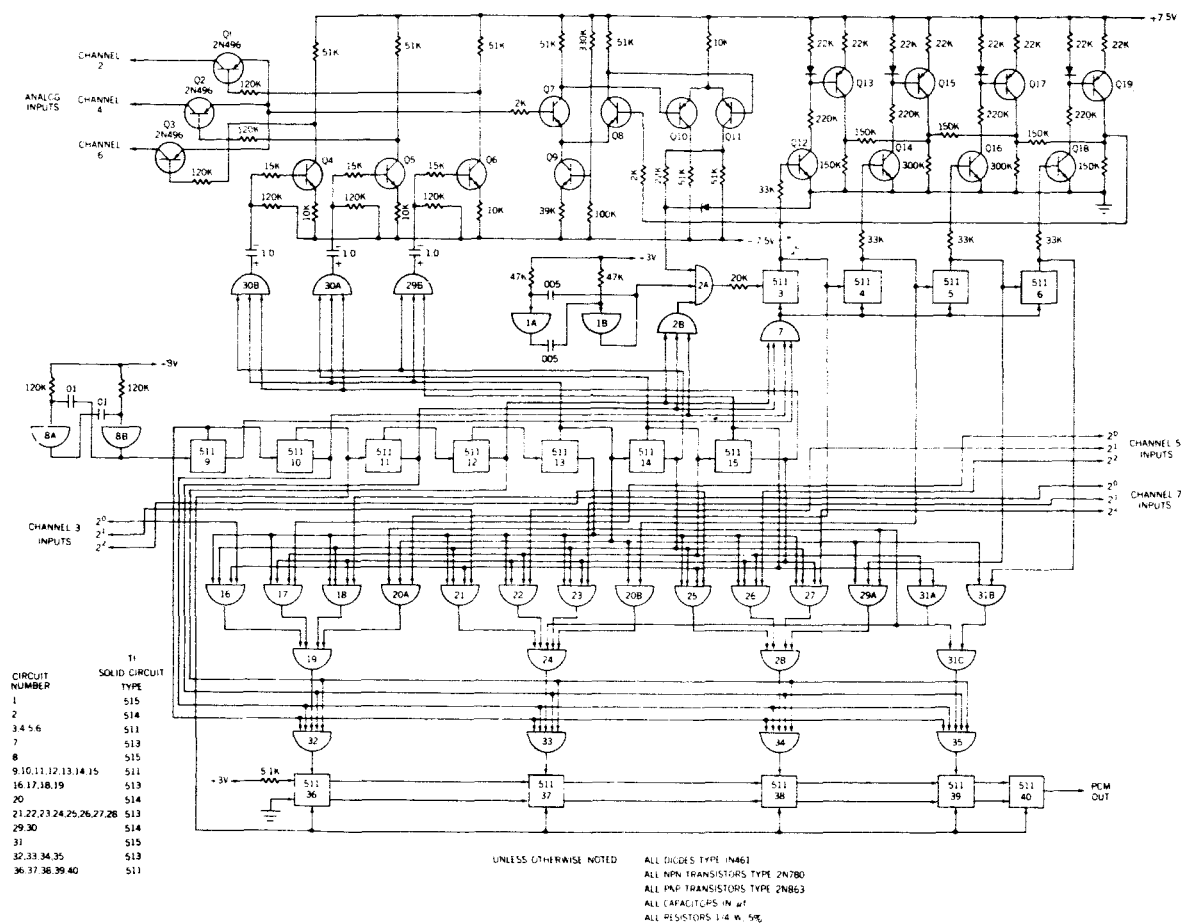


FIGURE 5.—Demonstration model of a PCM telemetry encoder designed with T1 solid circuits.

were used although the precision of four bits is ± 3 percent. The converter was checked for stability (as opposed to absolute accuracy) over a temperature range from -10°C to $+60^{\circ}\text{C}$ and was found to vary less than ± 2 percent even with the crude compensation used (diodes in the base circuits of Q13, Q15, Q17, and Q19 to compensate for V_{BE} changes). Also, the analog switching circuits (Q1 through Q6) were not optimized, and thus the converter output was one count high because of the V_{CE} of the switching transistors. The nominal analog input of the range is 0 to +5 volts; the digital inputs are 0 volts for logic zero and +3 volts for logic one.

CONCLUSIONS

The immediate advantages of using presently available Solid Circuits in these applications are

decreased size, weight, and cost. Because of the smaller number of interconnections, there is also the potential advantage of increased reliability, although this has not yet been proved. One disadvantage is the increased power required (Appendix A). For example, the eight-channel, 4 bit/channel PCM encoder discussed above requires about 110 mw of power, as compared with approximately 90 mw for the 272-channel, 7 bit/channel UK-2PFM encoder which uses conventional components in low-dissipation circuits. All of these remarks apply also, of course, to other types and of integrated circuits, such as the Fairchild Micrologic, Sperry Semi-Nets, Signetic Integrated Circuits, etc. In fairness it should be mentioned that the increased power requirements would be partially offset by increased converter efficiency. This would follow from the higher

power itself and also from the relaxed voltage regulation requirements (typically, ± 10 percent for the supply to the Solid Circuits in the digital portion of the system).

There are two basic reasons for the relatively high power requirements of all these circuits. First, present techniques limit the available resistance range of diffused resistors to less than 50 K, and even this limit is reached only at considerable expense to the consumer. Second, because the whole circuit is on a single semiconductor substrate, it is difficult to provide NPN and PNP transistors (which would permit construction of

complementary flip-flops). Several stopgap solutions (two packages, two chips in a single package) have been proposed for the second problem to avoid the more difficult techniques required to diffuse both type (PNP and NPN) transistors.

Future investigations should be directed first to reducing the power requirements for typical applications such as the ones shown here. This should be both from a device standpoint (e.g., higher resistance values) and from a configuration standpoint (e.g., series gate circuits instead of parallel gates).

APPENDIX A

POWER DISTRIBUTIONS

The power requirements given below are average, based on a 3 volt supply and a 2 milliwatt typical dissipation per collector node. All figures have been rounded to the nearest milliwatt.

Table A1

Power Requirements for the Clock Accumulator and Matrix Shown in Figure 3.

Circuit (see Figure 3)	Number of Solid Circuits @ 2 mw	Duty Cycle	Power Dissipation (mw)
Circuits 1 through 22	22	100%	44
Circuits 23 through 33	22	255/256	44
Circuits 34 through 37*	2	50% (avg)	2
Total for Solid Circuits			90
Input Transistors	10	7/8	18
Total power dissipation			108

*While there are four circuits, there are only two collector nodes.

Table A2

Power Requirements for the Clock and Matrix Shown in Figure 4.

Circuit (see Figure 4)	Number of Solid Circuits @ 2 mw	Duty Cycle	Power Dissipation (mw)
Circuits 1, 2, 5 through 8, and 18 through 21	10	100%	20
220-cps Oscillator	1	100%	2
Circuits 3A and 4B	2	50% (avg)	2
Circuits 3B and 4A	2	100%	4
Circuit 9	1	100%	2
Circuits 10A through 17A	10	15/16 (avg)	19
Circuit 17B	1	1/16	0
Circuits 22A through 25B	8	7/8	14
Circuits 26A and 26B	2	127/128	4
Total power dissipation			67

Table A3
Power Requirements for the PCM Encoder Shown in Figure 5.

Circuit (see Figure 5)	Number of Solid Circuits @ 2 mw	Duty Cycle	Power Dissipation (mw)
Circuits 1A, 1B, 8A, and 8B	4	100%	8
Circuits 2A and 2B	2	100% (max)	4
Circuits 3 through 6	4	100%	8
Circuit 7	1	15/16	2
Circuits 9 through 15	7	100%	14
Circuits 16 through 31B (less 19, 24, and 28)	14	15/16 (avg)	26
Circuits 19, 24, 28, and 31C	4	1/16 (avg)	1
Circuits 32 through 35	4	100% (avg)	8
Circuits 36 through 40	5	100%	10
Total for Solid Circuits			81
Transistors Q1 through Q19 and associated resistors			29
Total power dissipation			110

THE USE OF SOLID CIRCUITS IN SATELLITE INSTRUMENTATION*

EDGAR G. BUSH

An Optical Aspect Computer for the forthcoming Interplanetary Monitoring Probe has been designed and constructed with SN510 and SN514 solid-circuits used as binary counters, flip-flops, and inhibiting circuits.

The results obtained during the development and testing of the computer indicated that these devices have advantages over other types of logic components and are particularly suitable for spacecraft flight hardware applications.

INTRODUCTION

Integrated Circuits may be expected to have an increasing impact on all types of future electronic instrumentation. In particular, satellite instrumentation has much to gain from the use of integrated circuits: compact, reliable instrumentation can now be easily built that was beyond the realm of practicability before the event of these components.

This report is in the nature of a progress report on experiences and techniques derived during the development of what will probably be the first NASA unit using integrated circuits to be put into orbit. The integrated circuits to be discussed are Texas Instruments corporation solid circuits.

Many favorable features are offered by the various types of solid circuits, such as small size and weight and significantly fewer external connections needed for assembly on a printed circuit board or welded module as compared to the number required for a comparable transistorized circuit. For example, the assembly of a two transistor binary counter requires at least 22 connections on the board; the same function can be performed by a solid circuit with six. A strong argument for solid circuits is the fact that a complete system can be mounted on one printed circuit card that would require three such cards for conventional transistor circuitry. The elimination of numerous unreliable pressure-type contacts between cards is alone a very desirable feature.

The main *disadvantage* of solid circuits is their greater power drain over comparable transistor circuitry. A less tangible disadvantage is the lack of any precedent to follow in designing flight hardware circuitry using solid circuit devices—and a natural reluctance to substitute new devices for their flight-tested transistor counterparts.

BINARY COUNTERS

During the preliminary work with solid circuit binary counters, some difficulty was encountered in designing a transistorized input circuit. A positive input voltage signal is required by an SN510 binary counter (Figure 1), and triggering

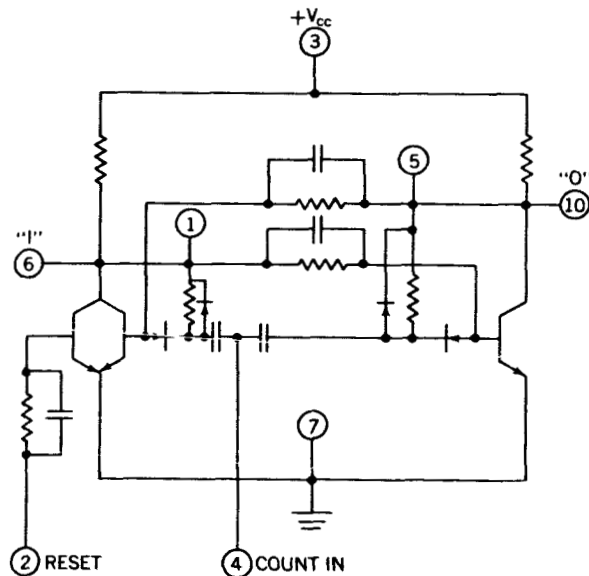


FIGURE 1.—SN510 solid circuit used as a binary counter.

*Published as NASA Technical Note D-1758, July, 1964.

occurs when the positive signal voltage drops toward ground potential. Reliable triggering requires that the signal voltage amplitude be somewhat less than the supply voltage, V_{cc} , to the binary. Figure 2 shows input circuitry that has actually been used in flight hardware. It can

be seen that the signal voltage requirement here is met by the voltage-divider action of R_1 and R_2 ; the voltage at the collector of Q_3 , which is the input signal, can never exceed $V_{cc}/2$.

In Figure 2, transistors Q_1 and Q_2 and capacitor C_1 produce a positive pulse to drive the input transistor to each counter. The purpose of the inhibiting circuit at the base of the input transistor is to stop the count at a time determined by gating signals. Since the collector-to-emitter voltage of the inhibiting transistors is in the order of 0.1 volt when the transistor is biased on, and since the offset voltage (the base-to-emitter voltage of the input transistor) is in the order of 0.5 volt, an inhibitor of this type is positive in action.

A unique type of counter shown in Figure 3 requires input circuits to both the first and second stages. The function of this counter is to count at the signal input rate until signal A is received and then count at half that rate until cut off by signal B. Since each input transistor of this counter has two independent inhibiting circuits, a single SN514 solid circuit was used to perform these functions instead of four transistors.

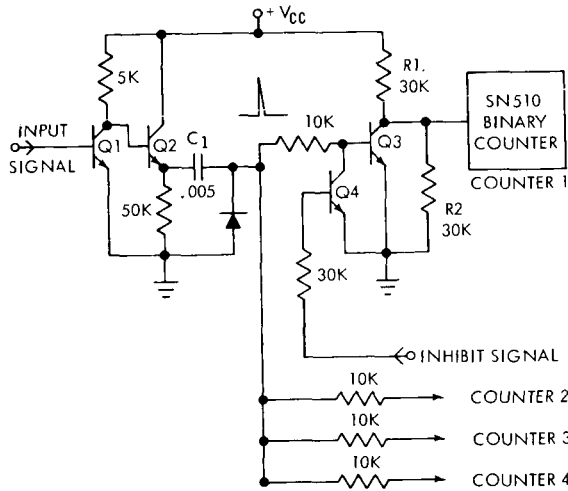
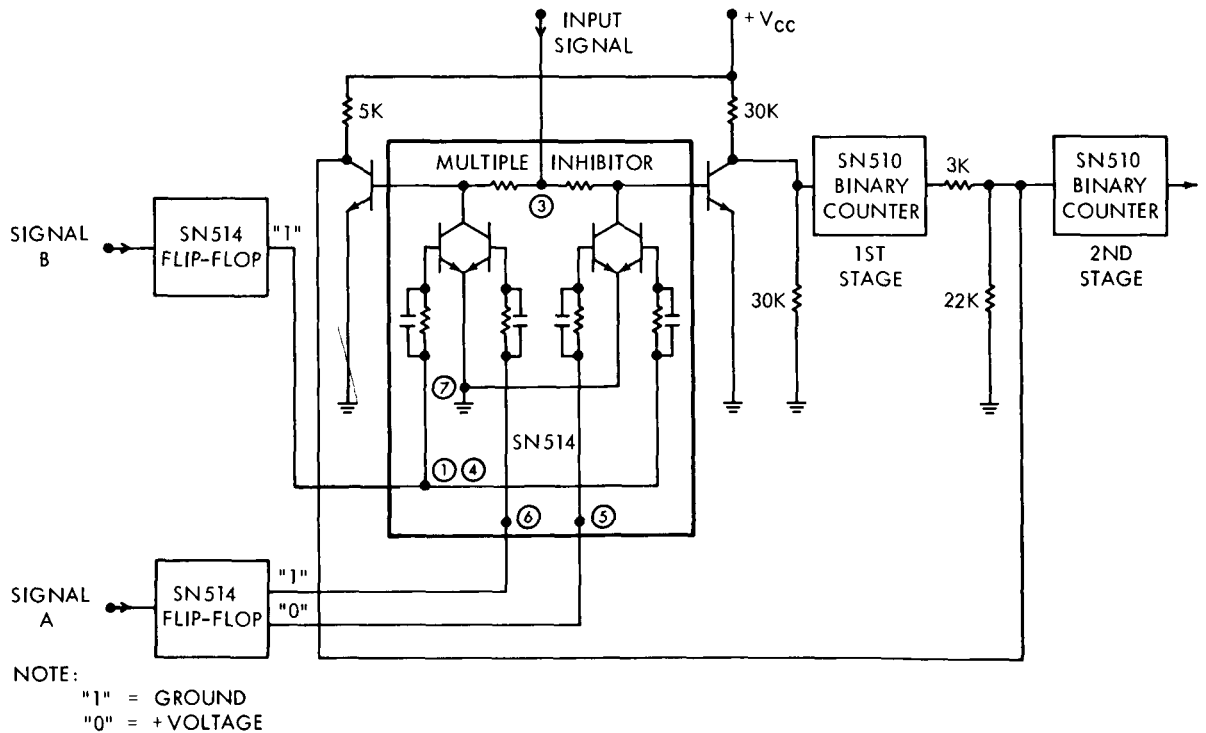


FIGURE 2.—Input circuit to SN510 binary counters.



NOTE:
 "1" = GROUND
 "0" = +VOLTAGE

FIGURE 3.—Full or half speed counter.

SET-RESET FLIP-FLOPS

A SN514 Texas Instrument solid circuit is shown in Figure 4. This circuit can be set to the "1" or "0" state by the application of a positive voltage to terminal 1 or 5 respectively. The set

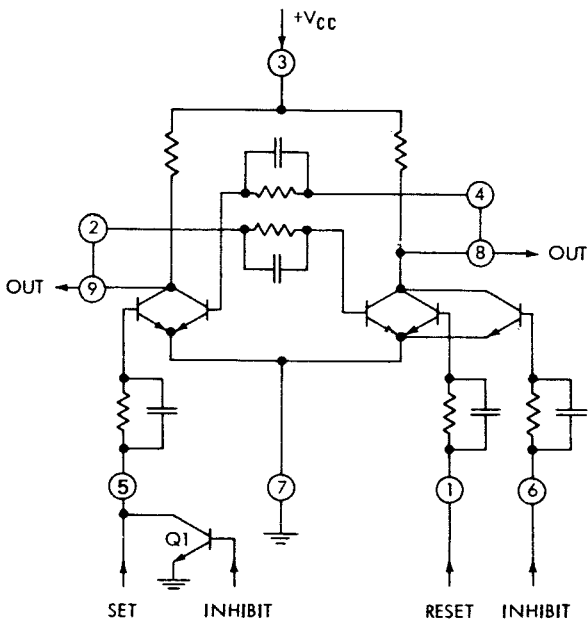


FIGURE 4.—SN514 solid circuit used as a flip-flop.

signal can be effectively gated or inhibited by the external transistor Q_1 . Alternatively, the transistor Q_1 can be eliminated and signal inhibition obtained by connecting the inhibiting voltage to terminal 6, if no inhibiting is allowed to occur after a signal pulse has set the flip-flop. If inhibiting is allowed during this period, the flip-flop will be returned to the reset state. Both inhibiting modes are useful.

THERMAL EFFECTS

Figures 5 and 6 show the value of the collector load resistor and the base input resistance respectively of an SN512 solid circuit as a function of temperature.* These diffused resistors have a large positive temperature coefficient, approximately doubling in resistance value from -50°C

*These measurements were made by R. Cliff and D. McDermond.

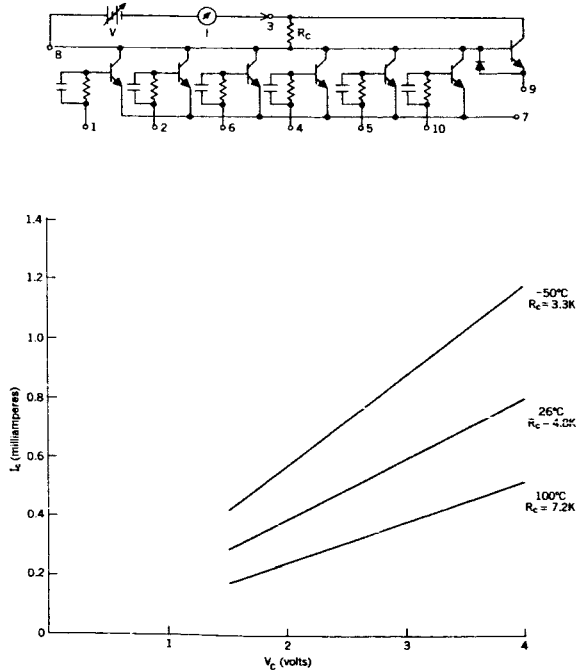
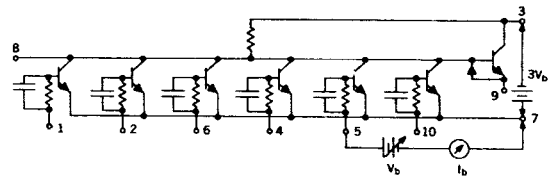


FIGURE 5.—Collector resistance of the SN512 solid circuit.

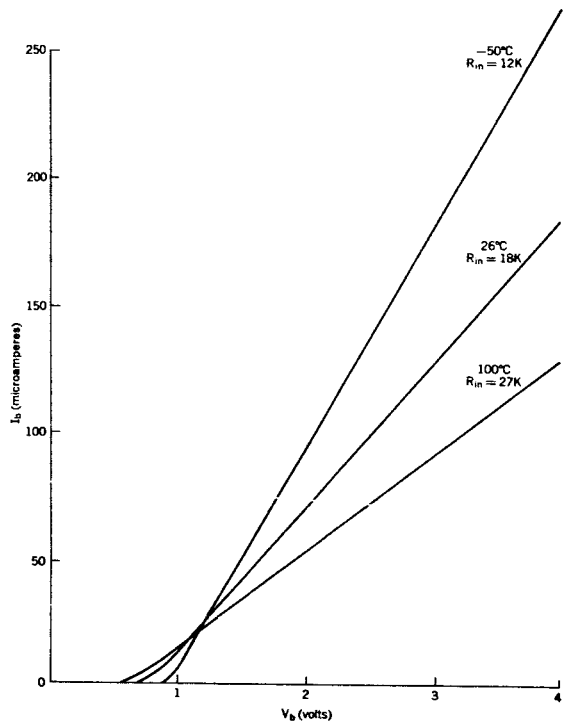


FIGURE 6.—Input resistance of the SN512 solid circuit.

to 100°C. Therefore, unlike conventional transistor circuitry these components consume more power cold than hot.

Figure 7 is a plot of the power dissipation vs. temperature of a SN510 solid circuit connected as a binary counter. Except for power dissipation, the temperature in the range -50°C to 80°C has little effect on the operation or stability of any of the previously described circuits and in no instance has temperature compensation been necessary.

BREADBOARDING TECHNIQUES

A breadboard circuit using solid circuits can easily be constructed by mounting each solid circuit on a card supplied by the manufacturer and

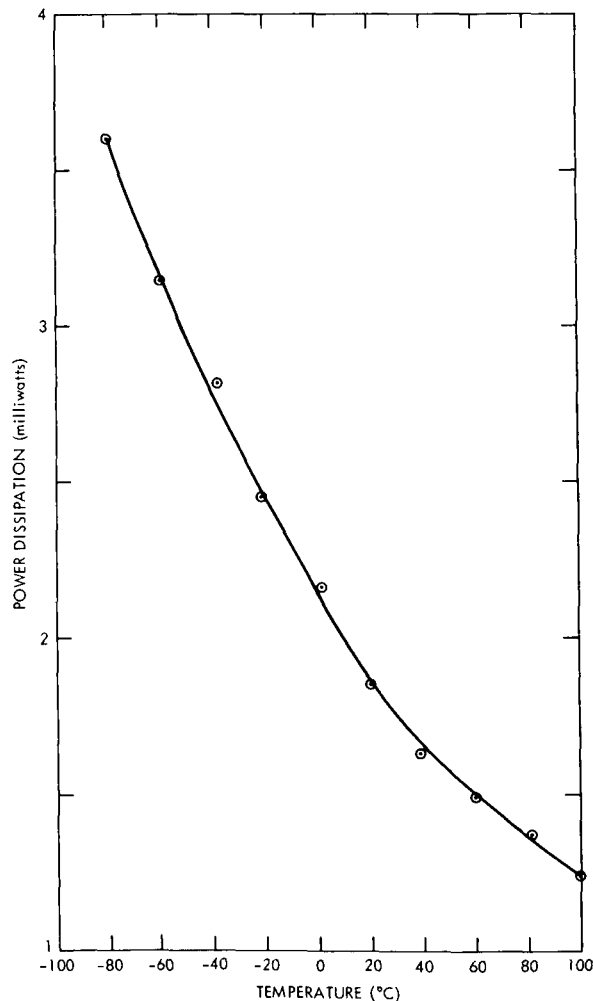


FIGURE 7.—Power dissipation of the SN510 solid circuit connected as a binary counter; $V_{cc}=3$ volts.

interconnecting these cards in a conventional manner. No special techniques, skills, or equipment are needed to work with solid circuits and the tabs for external connections will withstand a reasonable amount of flexing and stress. Certain advantages are gained by using solid circuits in circuit development or breadboarding for switching type circuits or logic networks. Since each solid circuit is a complete unit with components such as capacitors, resistors, and diodes diffused on the same silicon slab as the transistors, the assembly of any system which can employ such devices is greatly simplified. A system that would require three breadboards for conventional components can be mounted on one breadboard by using solid circuits, eliminating the need for interconnecting cables. The time required to assemble a circuit with conventional components is in the order of 10 times the period required to assemble the same circuit using solid circuits.

IMP OPTICAL ASPECT COMPUTER

All of the previously described circuits are used in the IMP optical aspect computer system, a block diagram of which is shown in Figure 8. Figure 9 shows the finished printed circuit card. The solid circuit devices are mounted on sub-modules (Figure 10) which are inserted into Digi-clips on the large printed circuit board; the Digi-clip connections are soldered before potting. This system can operate with a supply voltage tolerance of ± 50 percent from a nominal 3.5 volt supply, an input signal voltage range of ± 30 percent from a nominal 6 volt value, and over the temperature range -20°C to +80°C.

During testing, it has been noted that the solid circuit binaries and flip-flops are at least as good as their transistor counterparts with respect to stability in the presence of noise pulses, such as those generated by bench top temperature test chambers.

The majority of the active components used are contained in SN510 and SN514 solid circuit packages. Conventional transistor circuits were employed where it was not convenient to use solid circuits and where, in critical circuits, the question of reliability arose. The excellent quality of the SN510 and SN514 solid circuits is confirmed by the fact that although these components were assembled as received from the manufacturer

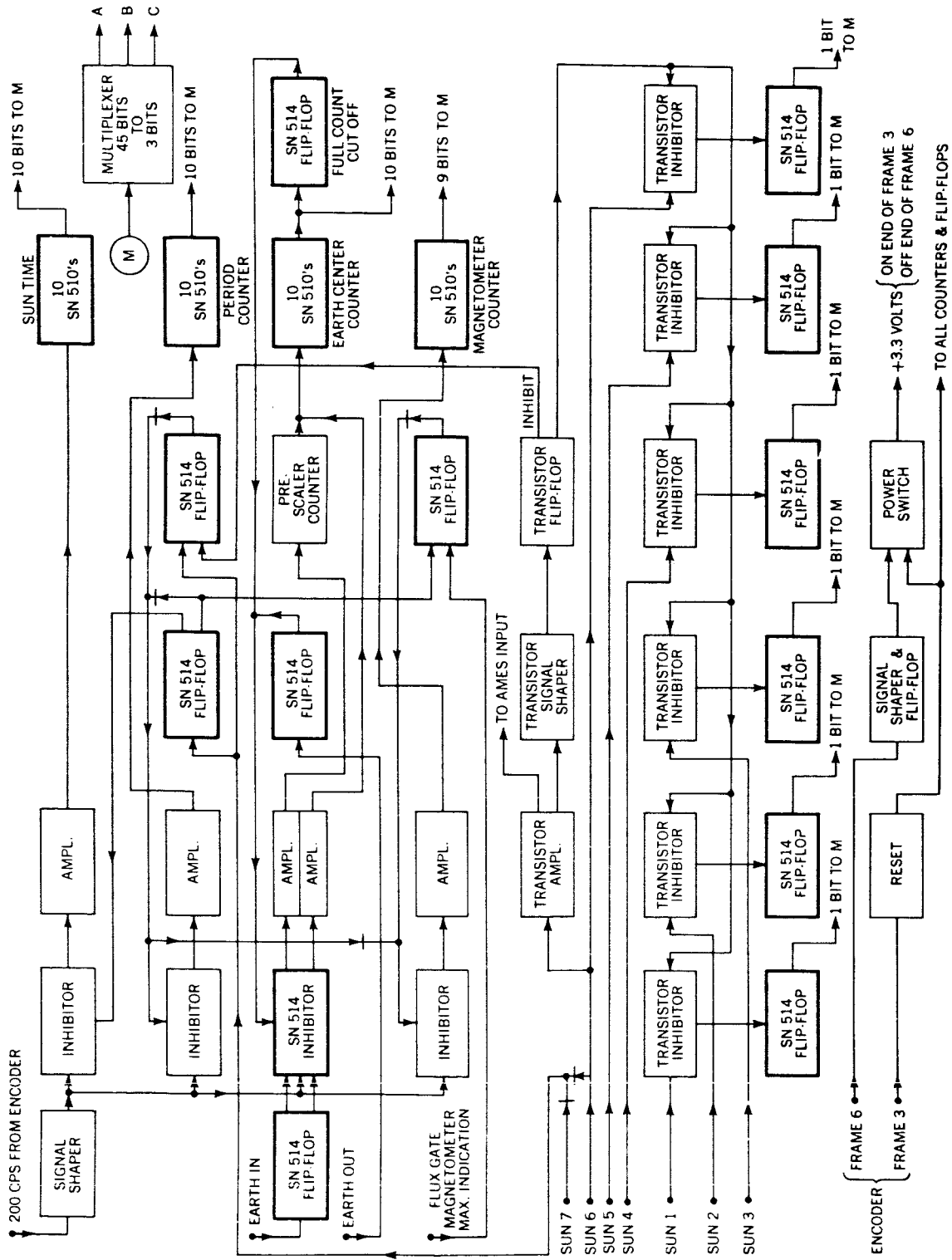


Figure 8.—IMP optical aspect computer.

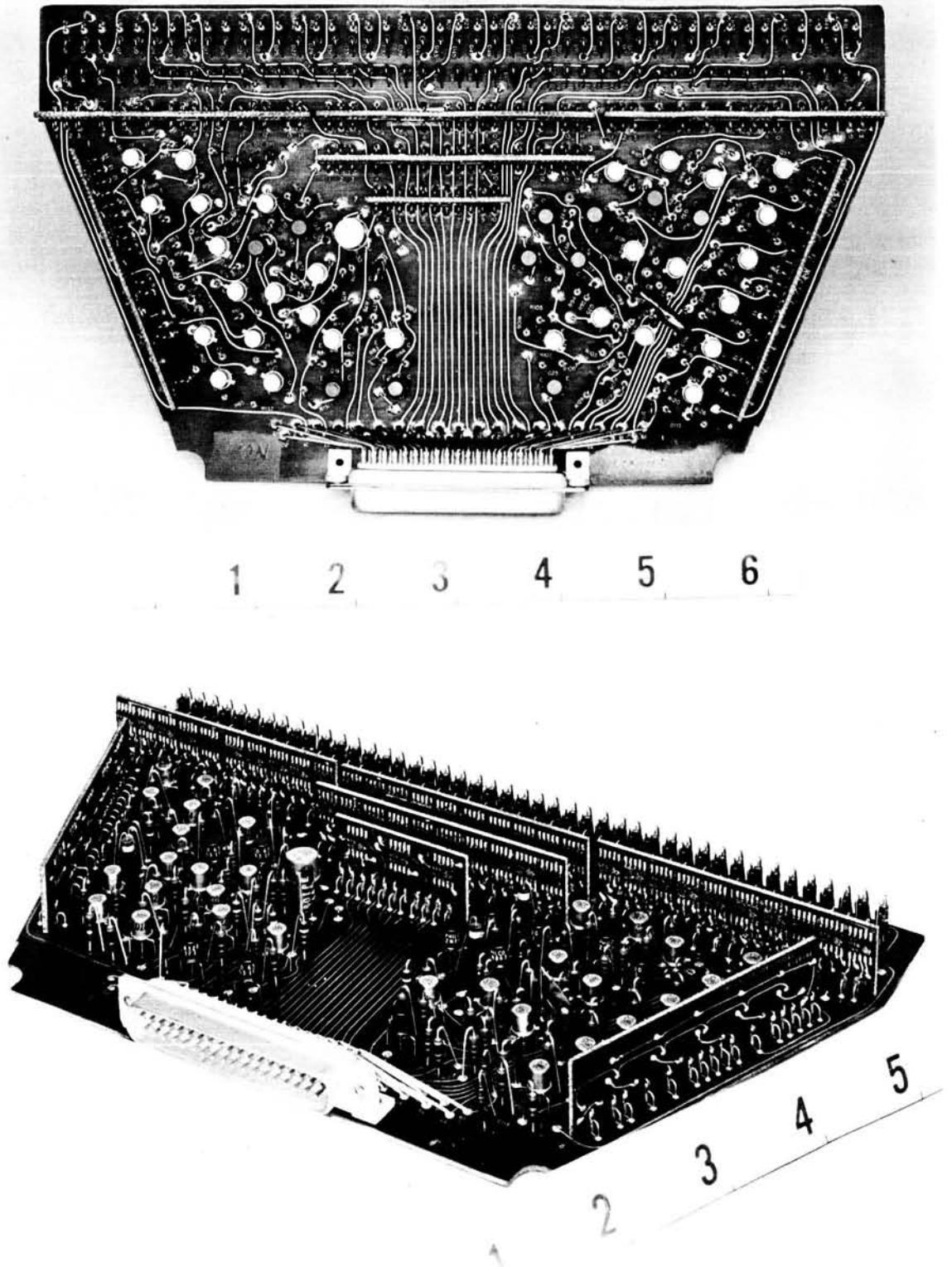


FIGURE 9.—Flight prototype model of the computer card.

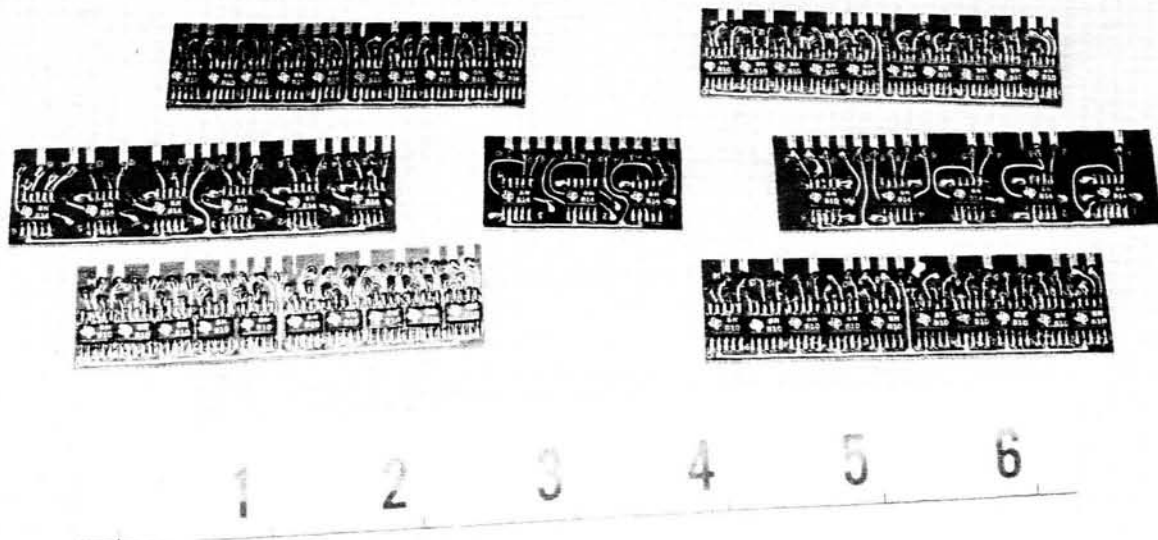


FIGURE 10.—Modules of the flight model computer card.

(without tests), in no instance was it necessary to replace a unit on the prototype card.

The fact that both the SN510 and the SN514 units dissipate 2 milliwatts with a 3 volt supply at 25°C as compared to $\frac{1}{2}$ milliwatt by their complementary-symmetry transistor counterparts is the main disadvantage in using these devices in flight hardware. In the design of the optical aspect computer a transistor switch disconnects the 3.5 volt supply from all circuits, excepting one flip-flop and two transistors for 4/5 of the time of one sequence of the encoder. This cuts the power

drain by a factor of 80 percent. The peak power dissipated by the system is approximately 150 milliwatts.

CONCLUSION

From the information acquired in using these devices and from quality assurance tests it appears that solid circuits are qualified both electrically and mechanically to be used as flight hardware components. The small size and weight of these logic networks should make possible in-flight computations which would not be practical with any other available type of circuitry.

THE EPE-D TWO YEAR TIMER*

T. D. CLEM

FOREWORD

This document is intended for general descriptive purposes only. It is a guide for those who need or wish to know the general circuitry and the specific functional operation of the Two Year Timer in relation to the spacecraft during testing and flight. No attempt has been made to present detailed theoretical or circuit analysis.

A solid state timing system used to disconnect the solar cells from the batteries after an interval of two years is described. The basic timing principle is discussed, as well as the circuits used to provide telemetry monitoring of the timing system performance. The circuitry for providing blockhouse testing and monitoring of the system is shown, along with an explanation of the precautions taken to isolate the test circuitry from the actual timing system after testing. Examples and the interpretation of the information received during testing are given. Finally, expected telemetry monitor readout during the complete two year period is tabulated.

GENERAL FUNCTION AND DESCRIPTION

The mission of the EPE-D Two Year Timer is to disable the payload so that transmissions will cease after a period of approximately 2 years from launch. This is accomplished by opening normally-closed magnetic latch-relay contacts which connect the solar cell panel and the battery. The timer consists of two identical, parallel, timing circuits which independently operate two latch relays. Connections are made to the relays such that both relays are required to operate before the solar cells are disconnected from the battery.

BASIC TIMING CIRCUIT OPERATION

A block diagram of the basic timing circuitry is shown in Figure 1. The power for the circuitry is obtained through a voltage regulator operating directly from the spacecraft +19.6 volts battery

*Published as Goddard Space Flight Center Document X-632-64-81, April 1964.

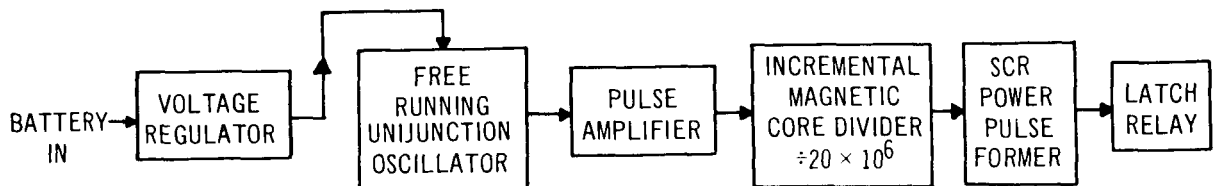


FIGURE 1.—Block diagram of basic timing circuit EPE-D.

line. The timing pulses are derived from a unijunction transistor acting as a relaxation oscillator. The pulse output of the unijunction is amplified before entering the magnetic core divider. An output from the divider occurs after 20×10^6 inputs. This output pulse then fires a silicon controlled rectifier, used as a power pulse-former, which in turn energizes a magnetic latch relay. Figure 2 shows this basic timing circuitry in a schematic diagram.

Figure 3 shows two basic timing circuits connected to form the Two Year Timer. The two timing circuits are completely independent except at the battery power input.

TELEMETRY MONITOR

Figure 4 is a block diagram showing the operation of the telemetry monitor incorporated into the timer units as well as the relay wiring. The monitor is a flip-flop which operates from certain intermediate outputs of the core dividers and

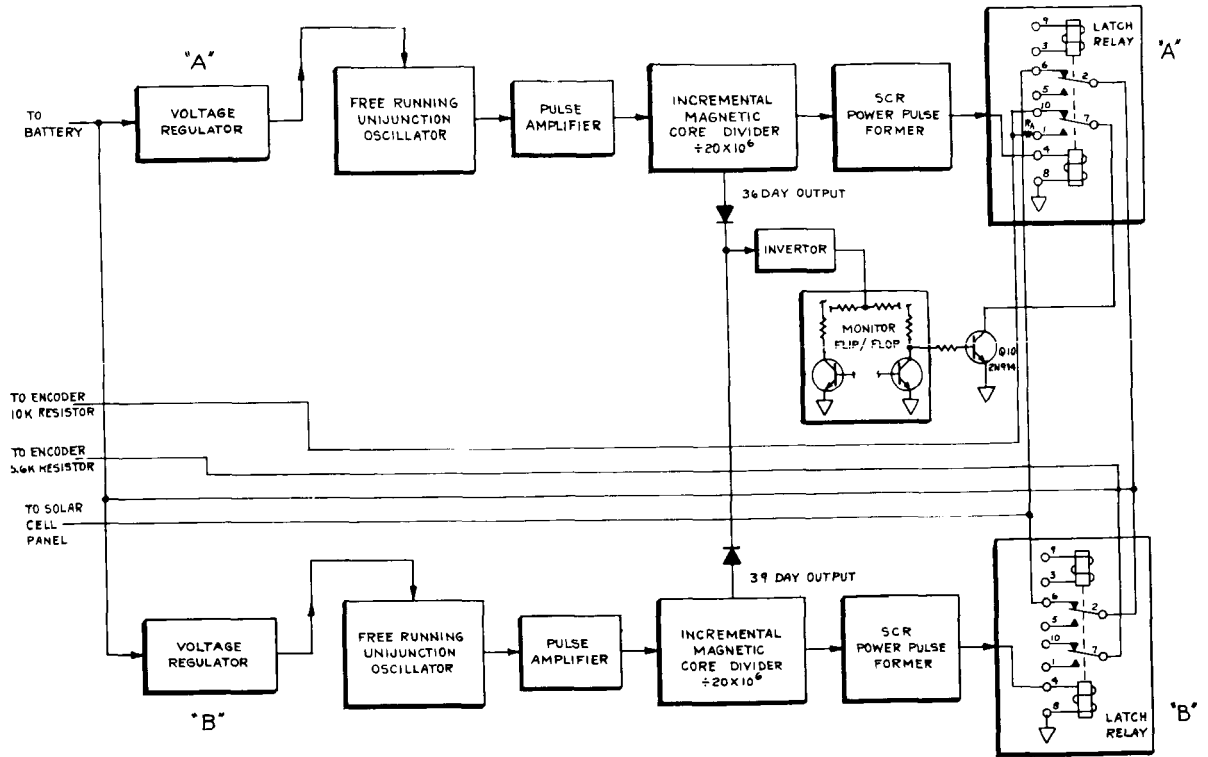


FIGURE 4.—Block diagram basic parallel timers showing redundant and telemetry monitor circuits EPE-D.

which acts to provide a shunt to ground to a section of a voltage divider network in the encoder. This divider network then furnishes voltage to a voltage controlled oscillator whose frequency is sampled and transmitted in the telemetry information.

Figure 5 depicts the operation of the telemetry monitor flip-flop upon the voltage divider network

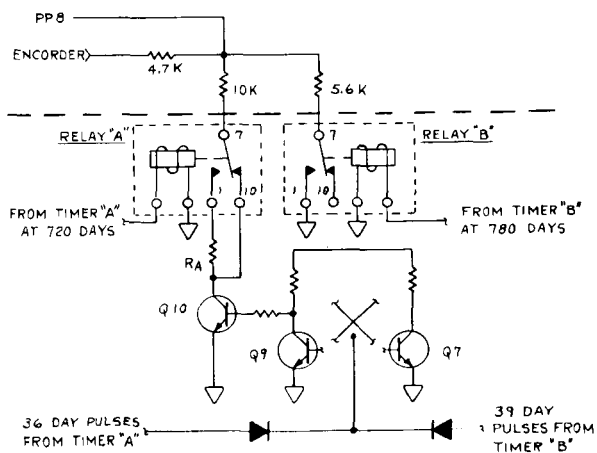


FIGURE 5.—Two year timer switching sequence EPE-D.

in the encoder. Either the 36 day pulses from timer A or the 39 day pulses from Timer B will operate the flip-flop of Q7 and Q9. Q8 of this flip-flop in turn causes Q10 to be on or off and hence short the 10K encoder resistor to ground or leave it open. This operation persists for 720 days. At this time, Relay A is actuated and the resistor R_A is connected in series with the 10K encoder resistor. The operation of Q10 and the flip-flop remain the same. The other contact of Relay A, acting as one of the parallel Solar Cell-Battery Connect switches, is also opened at this time.

At $T=780$ days, Relay B is actuated. With this, the 5.6K encoder resistor is connected to ground through one set of the relay contacts. The other set opens the remaining Solar Cell-Battery Connect switch and completes the disconnect of the Solar Cells and Battery. When this occurs, the life of the spacecraft depends solely upon the remaining capacity of the battery.

Table I summarizes the switching sequence of the Timer. The column labeled Resistance Switching is merely a list of the switching sequence

TABLE 1.

Time	Resistance Switching Figure 5	Voltage Level Switching Figure 6	Relay "A"	Relay "B"
T=0 to T=720 days	Open to 10K	1 to 3	Closed	Closed
T=720 days to T=780 days	Open to 10K+R _A 5.6K (10K+R _A)	1 to 2	Open	Closed
T=780 days to batt. death (≈2-3 hours)	5.6K to 5.6K+10K+R _A	4 to 5	Closed	Closed

of the resistors in the encoder voltage level network. The column labeled Voltage Level Switching refers to the curves shown in Figure 6. The reference voltage of the divider network in the encoder is derived from the -17.8 volt bus. Therefore, monitoring performance parameter 8

testing. This auxiliary circuitry is isolated from the timing circuit by diodes and resistors and uses separate power from the blockhouse. Therefore the timing circuitry is protected from any possible failures due to the auxiliary circuitry. The cable driver is a pulse stretcher and amplifier for monitoring the normal rate of the oscillator and the functioning of the pulseformer section of the magnetic core divider. The pulse outputs from the cable driver from both timing circuits are diode coupled and fed through the Normal Rate line to the blockhouse. The speed up rate for the timer is derived from a separate oscillator whose output is coupled into the pulse amplifier preceding the divider. Upon completion of the timing cycle at speed-up rate, the output of the SCR operating the relay fires an SCR which shunts the speed-up oscillator power to ground. The shunt SCR will stay on until the speed-up power is released. This operation insures that the timer does not pass the T₀ position of the timing cycle at the accelerated rate.

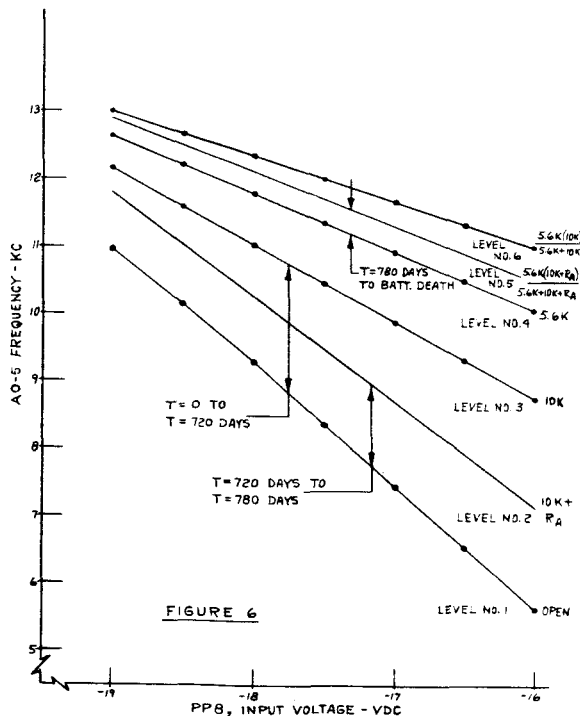


FIGURE 6.—Encoder divider resistance.

of the encoder sequence and comparing it to a calibration curve will give information concerning the -17.8 volt bus as well as the condition and operation of the Two Year Timer Unit.

SPEED-UP AND CABLE DRIVER OPERATION

Figure 7 is a block diagram of the auxiliary speed-up and cable driver circuitry provided for

DATA INTERPRETATION

Figure 8 shows the telemetry indication of the operation of the timer for the full two year interval. The indications are based on idealized timing oscillator function, that is, no error in accuracy or stability. Figure 8 also assumes that the telemetry monitor flip-flop is in such a position that the telemetry indication is at Level 1 at time T₀. (See Figure 6.) However, since the speed-up oscillators of the timer are not synchronized with the normal rate oscillators the telemetry indication may be at either Line 1 or line 3 at time T₀. Variations in normal oscillator rate that may be expected from normal variations in temperature, stability, and accuracy will be measured and estimated for each flight unit.

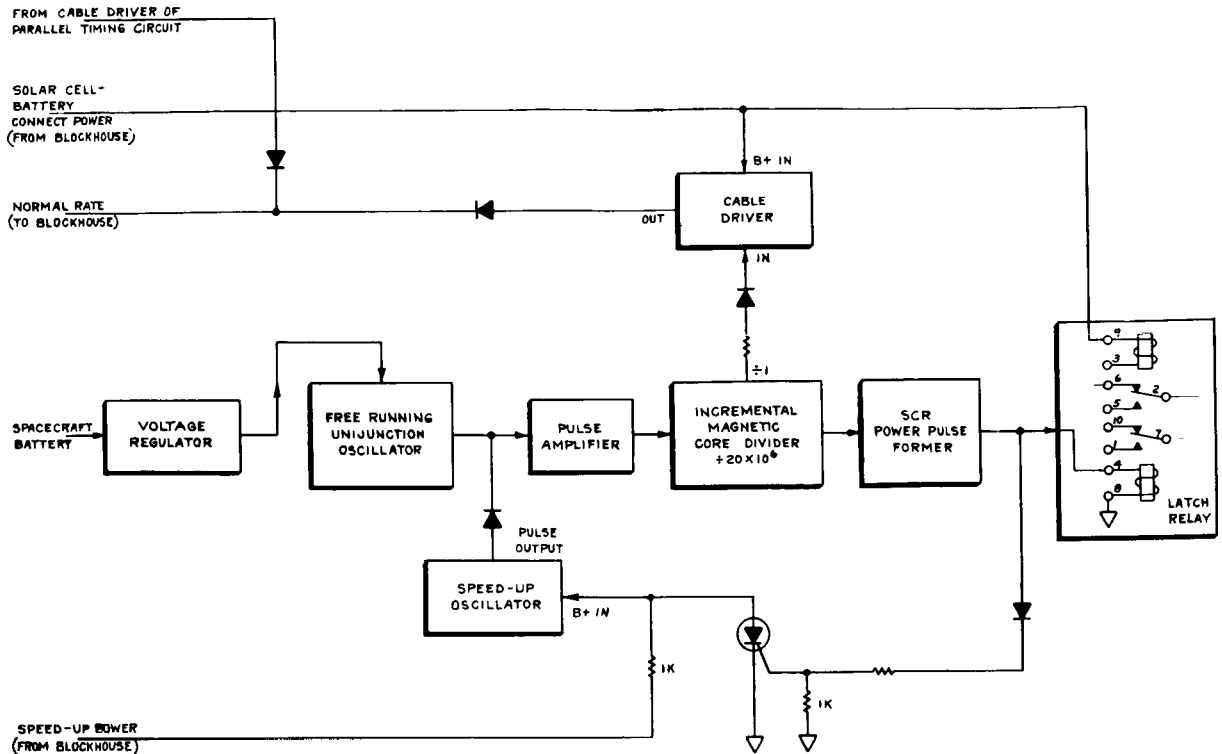


FIGURE 7.—Block diagram basic timing circuit showing speed-up and cable driver circuit EPE-D.

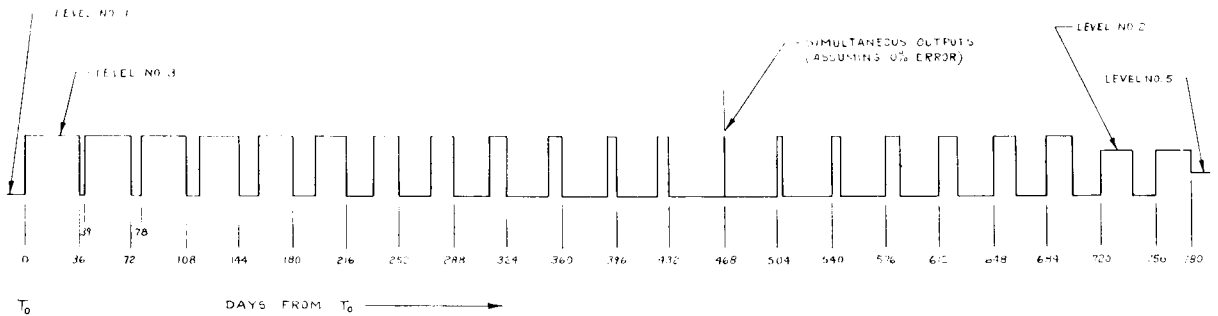


FIGURE 8.—Graph real time telemetry monitor readout EPE-D.

This data will then be available to aid in telemetry data interpretation.

POWER REQUIREMENTS

The two year timer will operate normally with battery voltage input range from +10.0 volts to +25 volts. At the normal spacecraft voltage of +19.6 volts, the unit draws 8 to 10 milliamperes.

ACCURACY AND STABILITY

The maximum components of error in the total accuracy of the Two Year Timer over the flight

lifetime of two years are tabularized below:

Component	Maximum Percentage of the Timing Error
Accuracy in initial oscillator rate	$\pm 1\%$
Stability at a given temperature	$\pm 3\%$
Stability over the temperature range of -20°C to $+70^{\circ}\text{C}$	$\pm 6\%$
Battery voltage variations	$< \pm 1\%$

Therefore, the total error of the two year interval will be ± 10 percent or less.

COMPONENT RELIABILITY CONTROL AND UNIT TESTING

The basic circuitry is relatively simple and has been flown successfully in several spacecraft. Therefore, extensive additional circuit analysis has not been necessary. However, considerable effort has been extended toward ensuring maximum component reliability. Filter and timing capacitors are high reliability solid tantalum types. Coupling capacitors are low temperature coefficient mylar. The capacitors are individually checked electrically and given burn-in tests in the lab before use. The unijunction transistors are selected and lab tested, primarily to ascertain low and stable peak point emitter current. The silicon controlled rectifiers are selected and given extensive tests for holding and firing current variations. The remaining transistors are high

reliability versions which have passed 200 hour burn-in tests.

The fabricated flight units will be given operating temperature tests in the lab to ensure overall stability. These tests will include a minimum of 20 temperature cycles between -30°C and $\pm 75^{\circ}\text{C}$ in addition to a minimum of 100 hours of operation at $+75^{\circ}\text{C}$.

PRELAUNCH TESTS AND MONITOR

Since the timing cycle of the Two Year Timer is continuous, it is important that the prelaunch tests and monitoring of the unit be performed with care and thought to ensure that the timer is at the beginning of its two-year cycle. These tests should be performed at any convenient time prior to launch. These tests should be performed at any convenient time prior to launch. The tests

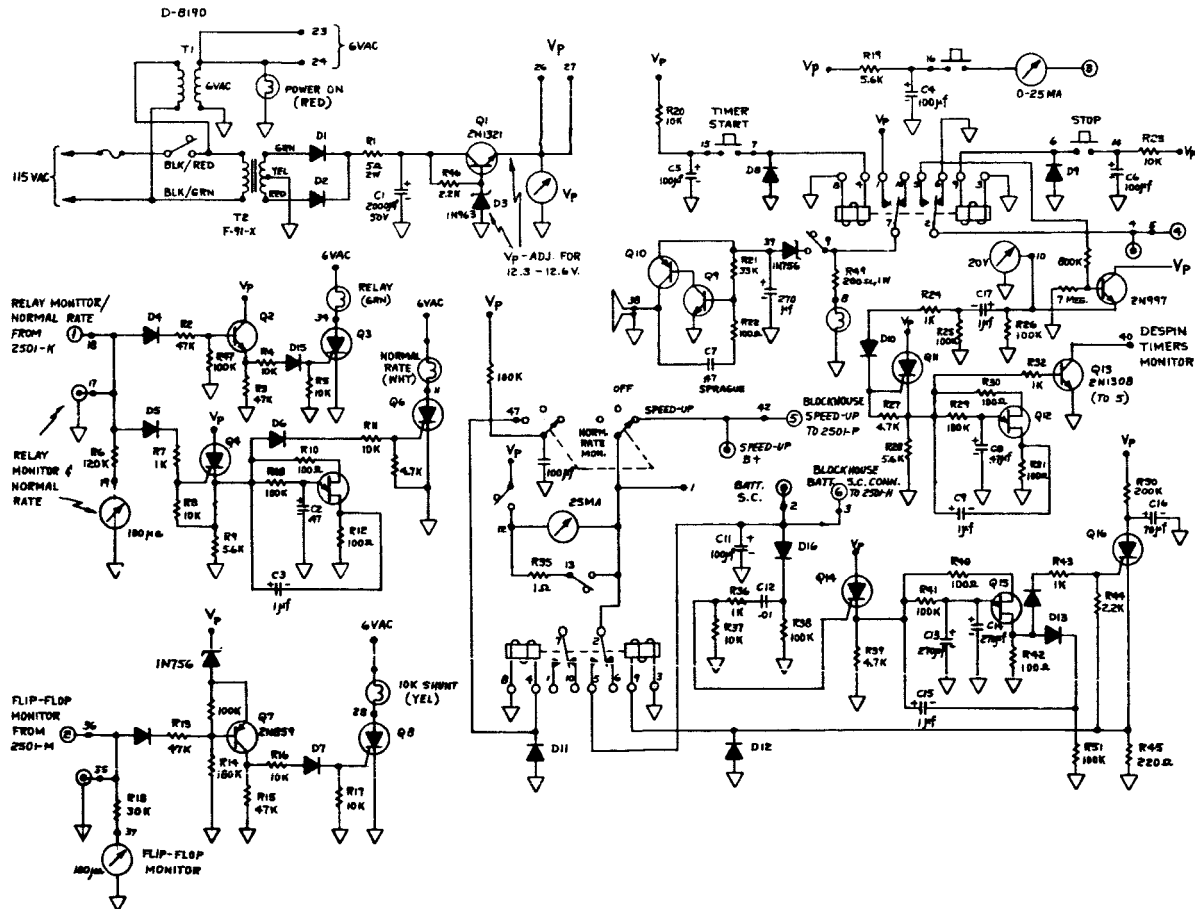


FIGURE 9.—Schematic diagram despin timer-progress switch—two year timer control and test panel S-3C.

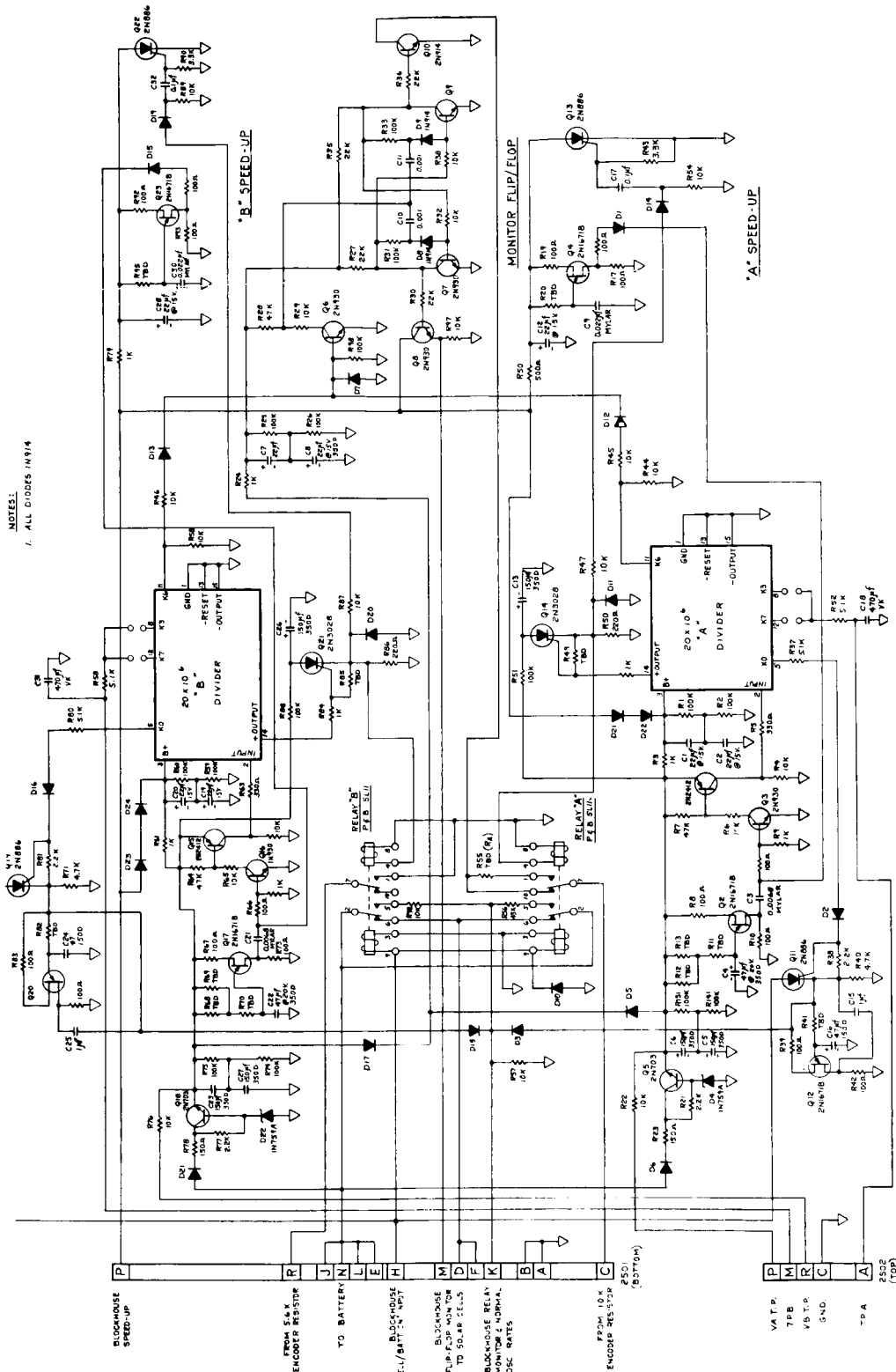


FIGURE 10.—Schematic diagram two year timer with monitor EPE-D.

will consist of running the timer through its cycle twice at the speed up rate.

The tests will include the following:

1. Timing of the interval length at speed-up rate.
2. Recording the flip-flop monitor output and observing the number and interval of the cross over points (See Figure 8).
3. Noting speed up power line current to ensure that the speed-up oscillators are turned off at the end of the timing cycle.
4. Ensuring operation of both relays in the timer at the end of the timing cycle.
5. Correlating the above information with decommutated telemetry readout.
6. Ensuring that the lines between the batteries and the solar cells are reconnected after the speed-up cycle.

7. Checking the normal oscillator rates.
8. Ensuring that speed-up current is not inadvertently applied before launch.
9. Recording position of telemetry monitor through telemetry readout.

All of this information will be available with proper monitoring of the Blockhouse lines from the Timer in the spacecraft. The complete test sequence and performance monitoring will be performed with the test panel provided with the Timer. This panel will also include test circuitry for the Program Switch and the Despin timer. The schematic of the test circuitry is shown in Figure 9. The prelaunch tests will be performed by Programmers Section personnel. The time required to perform the tests will be 2 hours or less.

The schematic of the complete Two Year Timer with Monitor is shown in Figure 10.

ANALYSIS OF NON-LINEAR NOISE IN FDM TELEPHONY TRANSMISSION OVER AN SSB-PM SATELLITE COMMUNICATION SYSTEM*

PAUL J. HEFFERNAN

Under high received signal conditions, voice channel quality in FDM telephony transmission over an SSB-PM satellite communication system is limited by dynamic non-linearity and down-link differential group delay. This report analyzes the noise produced by dynamic non-linearity in elements of the SSB up-link, spacecraft phase modulator, and ground receiver demodulator. Worst-case channel signal-to-non-linear-noise power ratios are developed in terms of the coefficients of a power series expressing the non-linear characteristic. The group delay problem is not treated. Four appendixes discuss in detail the calculation of the autocorrelation function, the evaluation of spectral convolutions, the determination of the power series coefficients, and CCIR terminology and multichannel loading procedures for FDM telephony.

INTRODUCTION

This report analyzes the effects of certain non-linearities in the transmission of frequency-division-multiplex (FDM) telephony over an SSB-PM communication link of the advanced Syncom† type.

Voice channel quality in an FDM SSB-PM link is determined by two factors: thermal noise, and intermodulation or non-linear noise. Non-linear noise is the unintelligible crosstalk in a voice channel due to the harmonics and intermodulations of the complex multichannel signal generated by system non-linearities. It resembles thermal noise to the ear but, unlike thermal noise, vanishes when the multichannel signal is removed.

With a weak received signal, thermal noise usually limits channel quality. However, with a strong received signal, thermal noise may be insignificant as compared with the non-linear noise; and in this case system non-linearities limit the channel quality. Under certain conditions, system parameters can be optimized by trading one type of interference for the other.

There are two important sources of non-linear noise in an SSB-PM system: dynamic nonlinearities and down link differential group delay. Dynamic non-linearities are encountered in the

amplitude characteristics of base-band multichannel amplifiers, klystrons, and phase modulators and discriminators.

Differential group delay is encountered on the down link when the instantaneous phase deviation $\phi(t)$ of the modulated carrier is modified non-linearly by a passive selective circuit as a function of the instantaneous frequency $\phi(t)$.

This paper analyzes the effects of dynamic nonlinearities in detail, and does not treat the group delay problem.

The principle source of non-linear noise in the SSB up-link is the dynamic characteristic of the ground transmitter power amplifier. Multichannel telephony has a peak-to-average ratio of

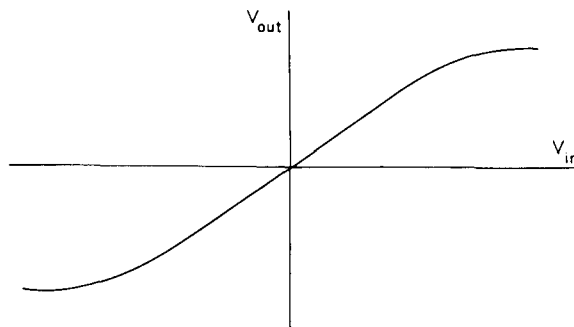


FIGURE 1.—Typical dynamic characteristic of a klystron power amplifier.

*Published as *NASA Technical Note D-2365*, July 1964.

†Synchronous Communications Satellite.

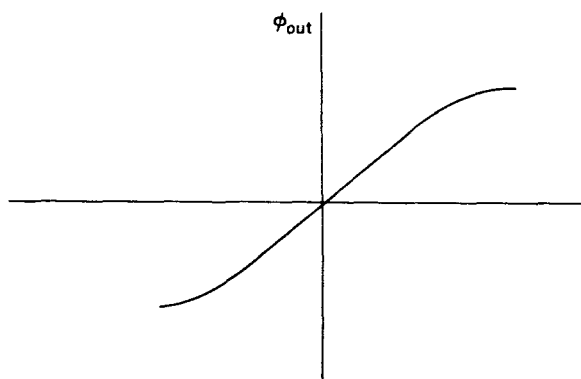


FIGURE 2.—Non-linear characteristic of a phase modulator.

approximately 13 db (Reference 1). Hence, the ground transmitter power amplifier must be linear over a considerable dynamic range, and in general is required to operate well below saturation.

In the spacecraft, all converters and amplifiers prior to the phase modulator must be dynamically linear. Because of the multiple access consideration, these units must be designed for considerably greater dynamic ranges than must the ground transmitter; thus linearity may be a problem despite the low levels at which these units operate.

In the PM down link, non-linearities in the characteristics of the spacecraft phase modulator and ground receiver demodulator introduce non-linear noise in exactly the same manner as do dynamic non-linearities in the up-link.

$$y(t) = \alpha x(t) + \beta x^2(t) + \gamma x^3(t) + \delta x^4(t) + \dots$$

For systems and devices which are essentially linear, and only deviations from linearity are being considered, the power series converges rapidly, and terms higher than the cubic may ordinarily be ignored. Where basically non-linear devices (rectifiers or limiters) are being treated, higher order terms must be included for an accurate representation.

In the frequency domain, the power spectrum of the output $y(t)$ may quite generally be obtained from a knowledge of $x(t)$ and the coefficients α , β , γ etc.

is defined as

$$R_x(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) x(t + \tau) dt = \overline{x(t) x(t + \tau)}$$

and is independent of time.

The CCIR recommends (Reference 2) that for purposes of analysis and testing, a multichannel telephony signal may be represented by a band of white Gaussian noise. Hence, either the instantaneous amplitude of the multichannel signal (on the SSB up-link) or the instantaneous phase deviation of the modulated carrier (on the PM down-link) may be represented by a real stationary random variable $x(t)$ of Gaussian statistics and narrowband white spectral density.

Where the non-linear characteristic of an element can be expressed as a power series of the input $x(t)$, the output ratio of channel test-tone power to the weighted non-linear noise power produced by the element can be computed in terms of the power series coefficients and the variance of $x(t)$. For a system of N channels, CCIR multichannel loading factors relate the mean square value of $x(t)$ to the power or mean square phase deviation corresponding to channel test-tone.

Hence, the performance of a non-linear element in the SSB-PM link can be specified in terms of its power series coefficients and how hard it is driven by a standard signal.

DESCRIPTION OF METHOD

In the time domain, the output $y(t)$ of a zero-memory non-linear device may be generally expressed as a power series of the input $x(t)$;

For the present case where $x(t)$ is a stationary random variable, the power spectrum of $y(t)$ can be calculated by using the Wiener-Khinchin theorem and the statistical properties of $x(t)$ (probability density function, autocorrelation function, spectral density, etc.)

The Wiener-Khinchin theorem states that the autocorrelation function of a sample function of a random process and the spectral density of the process form a Fourier transform pair. If $x(t)$ is a real sample function of a stationary random process, we have that its autocorrelation function

The theorem states that

$$S_x(f) = \mathcal{F}[R_x(\tau)]$$

$$R_x(\tau) = \mathcal{F}^{-1}[S_x(f)]$$

where $S_x(f)$ is the spectral density of the random process.

If $x(t)$ is the input to a non-linear device with output $y(t)$, we have the output autocorrelation function

$$R_y(\tau) = \overline{y(t)y(t+\tau)} ;$$

and we can operate on this to get the spectral density $G_y(f)$ of $y(t)$:

$$G_y(f) = \mathcal{F}[R_y(\tau)]$$

This method of determining the output spectrum of a non-linear device is quite general and very powerful.

For FDM telephony a knowledge of the output spectrum permits the determination of channel signal-to-noise ratios directly. The method used here may be considered an analytical equivalent of the noise-loading procedures widely used to test FDM equipment and systems (Reference 3).

The heart of the problem is to determine the output autocorrelation function $R_y(\tau)$. This demands thorough knowledge of the statistics of the input $x(t)$ and the power series coefficients. It is possible, but not always desirable, to develop a closed form expression for $R_y(\tau)$. For the power series case with Gaussian statistics, we can calculate $R_y(\tau)$ in terms of a power series in $R_x(\tau)$, considerably simplifying the mathematics.

$$y(t) = \alpha x(t) + \beta x^2(t) + \gamma x^3(t) + \delta x^4(t) + \dots$$

The output autocorrelation function is

$$R_y(\tau) = \overline{y(t)y(t+\tau)}$$

Truncating the power series, we have

$$y(t)y(t+\tau) = [\alpha x(t) + \beta x^2(t) + \gamma x^3(t)] \cdot [\alpha x(t+\tau) + \beta x^2(t+\tau) + \gamma x^3(t+\tau)]$$

and $R_y(\tau)$ will be the sum of the expectations of each of the terms produced in the multiplication.

CALCULATION OF OUTPUT SPECTRA AND S/N RATIOS

The input to the non-linear element of the SSB-PM system is a stationary Gaussian functional of zero mean and variance $\overline{x^2}$; i.e., we have

$$p(x) = \frac{e^{-x^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}$$

$$\overline{x} = 0$$

$$\overline{x^2} = \sigma^2$$

The autocorrelation function is

$$R_x(\tau) = \overline{x(t)x(t+\tau)}$$

and

$$R_x(0) = \overline{x^2(t)} = \overline{x^2(t+\tau)} = \sigma^2$$

The spectral density $S_x(f)$ of the multichannel signal is defined as the average power per unit bandwidth and is an even (two-sided) function of frequency.

From the Wiener-Khinchin theorem,

$$S_x(f) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j\omega\tau} d\tau$$

and

$$R_x(\tau) = \int_{-\infty}^{\infty} S_x(f) e^{j\omega\tau} df$$

Assuming a converging power series expresses the non-linear characteristic, the output is

It is at this point that the statistics of $x(t)$ enter the analysis. For $x(t)$ a Gaussian functional of

zero mean and variance \bar{x}^2 , the autocorrelation function (Appendix A) is

$$\begin{aligned} R_y(\tau) &= \beta^2 \bar{x}^2 \\ &+ \left[\alpha^2 + 6\alpha\gamma \bar{x}^2 + 9\gamma^2 (\bar{x}^2)^2 \right] R_x(\tau) \\ &+ \left[2\beta^2 + 24\beta\gamma \bar{x}^2 \right] R_x^2(\tau) \\ &+ \left[6\gamma^2 \right] R_x^3(\tau) \end{aligned}$$

The power spectrum $G_y(f)$ of the output $y(t)$ can now be obtained by forming the Fourier transform of the entire expression.

The first term represents dc energy, since the transform of a constant in the time domain is a delta function at zero frequency:

$$\int_{-\infty}^{\infty} k e^{-j\omega\tau} d\tau = k \lim_{T \rightarrow \infty} \left(\frac{\sin 2\pi f T}{\pi f} \right) = k \delta(f)$$

For the purpose of computing the output spectrum, this term can be ignored.

The terms in the coefficients of $R_x(\tau)$ and $R_x^2(\tau)$ proportional to \bar{x}^2 represent second order effects which may be disregarded for the purpose of computing the output spectrum.* The autocorrelation function is then written

$$R_y(\tau) \approx \alpha^2 R_x(\tau) + 2\beta^2 R_x^2(\tau) + 6\gamma^2 R_x^3(\tau)$$

The first term is transformed to yield a first order spectrum $G_1(f)$ of the freely transmitted signal.

$$G_1(f) = \int_{-\infty}^{\infty} \alpha^2 R_x(\tau) e^{-j\omega\tau} d\tau = \alpha^2 S_x(f)$$

Note that the effect is simply that of a linear bandpass filter of gain α .

The other terms yield second, third, and fourth order spectra proportional to convolutions of the

*This is so if the power series for $y(t)$ converges rapidly. However, terms like $6\alpha\gamma\bar{x}^2$ are interesting in their own right. They indicate that the apparent linear gain of the non-linear element depends on the variance of the Gaussian input. Lewin (Reference 4) has shown the instantaneous waveforms corresponding to these spectral terms are always in exact phase coherence with the freely transmitted waveform. Hence, these can hardly be called distortion spectra. For a single tone input, terms of this type mathematically represent the action of a tuned limiter, and provide a convenient means of determining the power series coefficients α , β , γ , etc., of a non-linear element (see Appendix B).

input spectrum $S_x(f)$ with itself.

$$G_2(f) = 2\beta^2 S_x(f) * S_x(f)$$

$$G_3(f) = 6\gamma^2 S_x(f) * S_x(f) * S_x(f)$$

These convolutions are evaluated in Appendix C.

It may be seen that the output signal to non-linear noise power ratio at any frequency is given by

$$\frac{P_s}{P_n} = \frac{G_1(f)}{G_2(f) + G_3(f)}$$

To evaluate the spectral convolutions, it is necessary at this point to define more completely the input spectral density $S_x(f)$ for the two situations of interest in this analysis, i.e., multichannel telephony spectra at baseband and IF. Figure 3 shows these two cases.

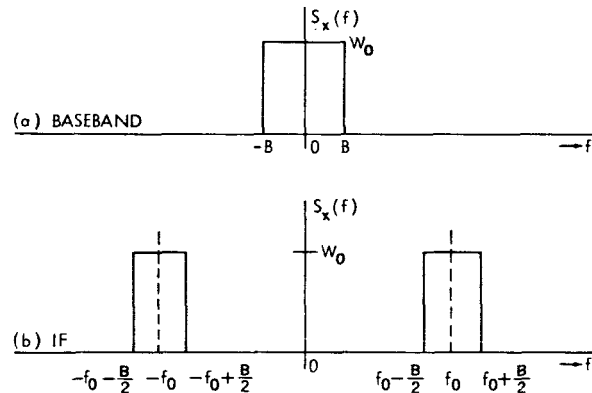


FIGURE 3.—Multichannel signal spectra at baseband and IF.

In either case, the spectral density is the average power per unit bandwidth, and has the value

$$W_0 = \frac{\bar{x}^2}{2B}$$

within the band and is zero outside the band. Spectral density is an even function of frequency

$$S_x(f) = S_x(-f)$$

and we have that

$$\int_{-\infty}^{\infty} S_x(f) df = \bar{x}^2$$

Baseband Case

The spectrum of the multichannel signal at baseband is shown in Figure 3(a). In reality, the spectrum does not extend to zero frequency; but for the purpose of this report no significant error is introduced by this representation, since dc levels produced by system non-linearities are filtered out by the natural system bandwidth limitations and need not be considered.

The convolutions $S_x(f) * S_x(f)$ and $S_x(f) * S_x(f) * S_x(f)$ for the baseband case are evaluated in Appendix C and shown in Figure C7. The baseband non-linear noise spectra $G_2(f)$ and $G_3(f)$ obtained on this basis are shown in Figure 4, along with $G_1(f)$, the spectrum of the freely transmitted signal.

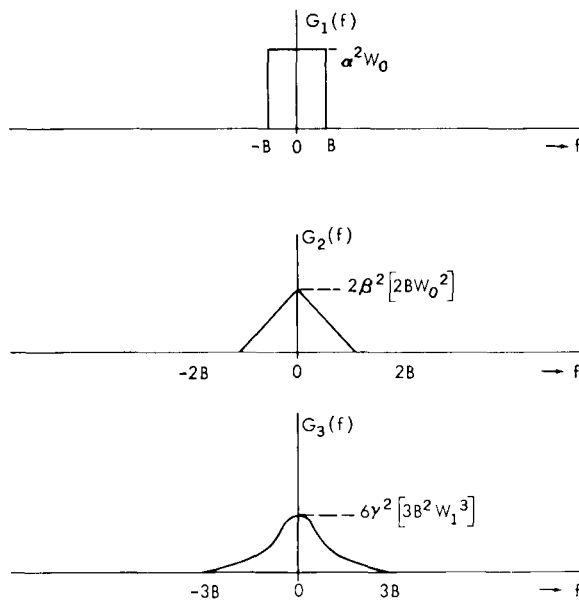


FIGURE 4.—The baseband output spectra $G_1(f)$, $G_2(f)$, and $G_3(f)$.

The ratio of signal power to non-linear noise power in a slot of width Δf at the bottom of the baseband (corresponding to the worst channel) is then

$$\frac{P_s}{P_n} \approx \frac{[G_1(0)] \Delta f}{[G_2(0) + G_3(0)] \Delta f}$$

and in general, in any channel,

$$\frac{P_s}{P_n} > \frac{\alpha^2 W_0}{4\beta^2 B W_0^2 + 18\gamma^2 B^2 W_0^3}$$

From the definition of the input spectral density

$$W_0 = \frac{\overline{x^2}}{2B}$$

we have

$$\frac{P_s}{P_n} \geq \frac{\alpha^2}{2\beta^2 \overline{x^2} + \frac{9}{2} \gamma^2 (\overline{x^2})^2}$$

Note that the signal-to-noise power ratio is given only in terms of the power series coefficients squared and the variance of the Gaussian input.

IF Case

The spectrum of the multichannel signal at IF is shown in Figure 3(b).

The convolutions $S_x(f) * S_x(f)$ and $S_x(f) * S_x(f) * S_x(f)$ for the IF case are evaluated in Appendix C and shown in Figures C10 and C11. The IF output non-linear noise spectra $G_2(f)$ and $G_3(f)$ obtained on this basis are shown in Figure 5, along with $G_1(f)$, the spectrum of the freely transmitted signal. It is immediately clear that for the IF case, all of $G_2(f)$ and the portion of $G_3(f)$ centered at $\pm 3f_0$ fall outside the region of the spectrum occupied by $G_1(f)$, and hence may be assumed to be filtered out by system bandwidth limitations. The portion of $G_3(f)$ centered at $\pm f_0$ is the spectrum of non-linear noise that cannot be filtered out.

The ratio of signal power to noise power in a slot of width Δf at f_0 (corresponding to the worst channel) is

$$\frac{P_s}{P_n} \approx \frac{[G_1(f_0)] \Delta f}{[G_3(f_0)] \Delta f}$$

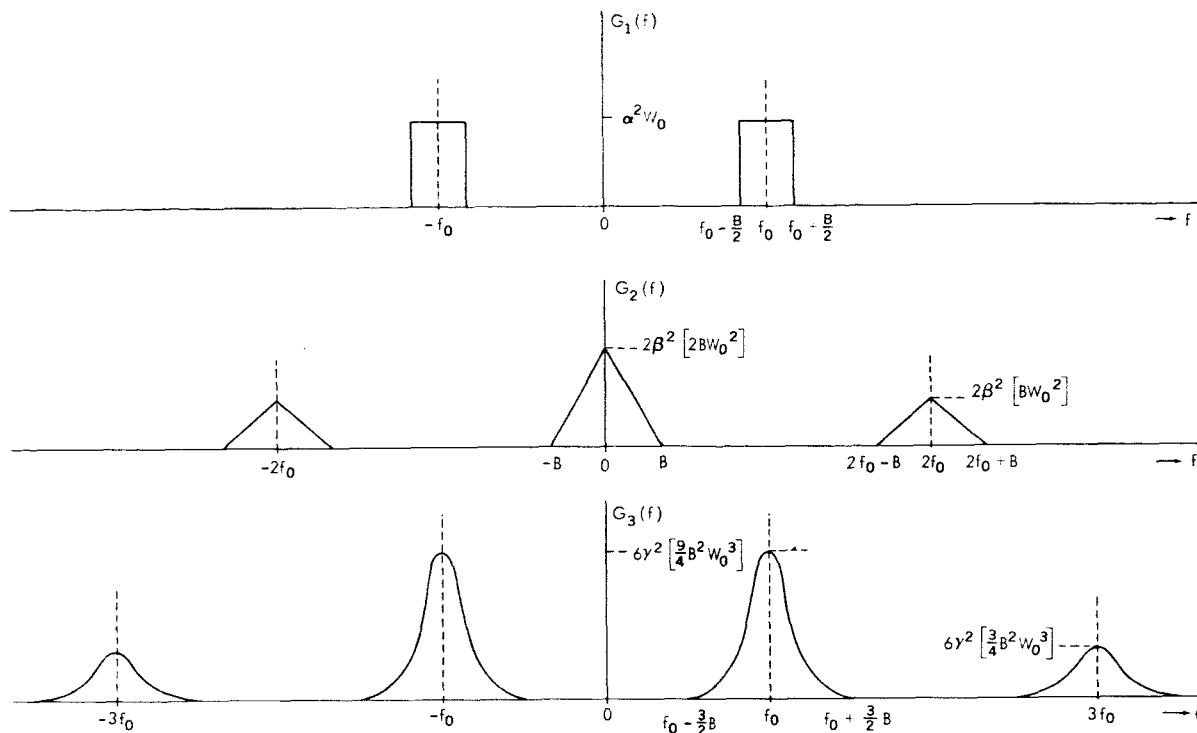
and in any channel

$$\frac{P_s}{P_n} \geq \frac{\alpha^2 W_0}{6\beta^2 \left[\frac{9}{4} B^2 W_0^3 \right]}$$

From the definition of W_0 ,

$$\frac{P_s}{P_n} \geq \frac{8}{27} \left(\frac{\alpha}{\gamma} \right)^2 \frac{1}{(\overline{x^2})^2}$$

Again, the signal to noise power ratio is in terms

FIGURE 5.—The IF output spectra $G_1(f)$, $G_2(f)$, and $G_3(f)$.

of the squares of the power series coefficients and the variance of the Gaussian input.

APPLICATION TO ADVANCED SYNCOM

The formulas developed above are directly applicable to certain elements of the SSB-PM FDM telephony mode of the proposed advanced Syncom communication satellite system.

The result for the baseband case

$$\frac{P_s}{P_n} \geq \frac{\alpha^2}{2\beta^2 \bar{x}^2 + \frac{9}{2} \gamma^2 (\bar{x}^2)^2}$$

applies to baseband amplifiers at the ground transmitter and ground receiver, and also the spacecraft modulator* and ground receiver phase demodulator.

The result for the IF case,

$$\frac{P_s}{P_n} \geq \frac{8}{27} \left(\frac{\alpha}{\gamma}\right)^2 \frac{1}{(\bar{x}^2)^2}$$

*It should be pointed out that the phase modulator proposed for the advanced Syncom is unique in that it effectively performs a down-conversion of the multichannel signal in the modulation process. The input to the modulator is at IF, but the phase spectrum of the modulated carrier is at baseband, and demodulation of the carrier yields the multichannel signal at baseband, not at IF.

applies to the SSB exciter chain and power amplifier of the ground transmitter, and the front-end and IF amplifier of the spacecraft receiver.

These signal-to-noise power ratios are analytical equivalents of the "noise-to-noise" power ratio (NPR) measured in a noise loading test as described by White and Whyte, (Reference 3). Since it is customary to characterize the performance of international telephone circuits in terms of a channel test-tone to psophometrically-weighted-noise power ratio, it is necessary to convert the formulas developed above into this form. This is readily done by applying the appropriate CCIR loading factor and NPR conversion factor as discussed in Appendix D.

To illustrate, we consider the linearity requirements of a SSB ground power amplifier for 600 channels. We use the formula for the IF case, and interpret \bar{x}^2 as the mean multichannel signal power dissipated through unit resistance into the amplifier:

$$\frac{P_s}{P_n} \geq \frac{8}{27} \left(\frac{\alpha}{\gamma}\right)^2 \frac{1}{(\bar{x}^2)^2}$$

The NPR conversion factor for 600 channels is 19.2 db, meaning that the noise-to-noise ratio in a 3.1 kc slot measured in a noise loading test or calculated analytically is 19.2 db poorer than the corresponding channel test-tone to psophometrically-weighted-noise power ratio:

$$\left. \frac{P_{tt}}{P_n} \right|_{\substack{3.1 \text{ kc} \\ \text{weighted}}} \geq (10^{1.92}) \frac{8}{27} \left(\frac{\alpha}{\gamma}\right)^2 \frac{1}{(\bar{x}^2)^2}.$$

The loading factor of 600 channels is 12.8 dbm₀; meaning that the mean multichannel signal power is 12.8 db above the power of a channel test-tone:

$$\begin{aligned} \left. \frac{P_{tt}}{P_n} \right|_{\substack{3.1 \text{ kc} \\ \text{weighted}}} &\geq (10^{1.92}) \frac{8}{27} \left(\frac{\alpha}{\gamma}\right)^2 \frac{1}{(10^{1.28} P_{tt})^2} \\ &\geq 10^{-.64} \left(\frac{8}{27}\right) \left(\frac{\alpha}{\gamma}\right)^2 (P_{tt}^2)^{-1}. \end{aligned}$$

This result now specifies how hard the power amplifier may be driven in terms of its power series coefficients; and conversely, for a given power level and allowable non-linear noise level, the permissible ratio of the power series coefficients is specified.

As a second illustration, we consider the linearity requirements of a ground receiver phase demodulator for 1200 channels. We use the formula for the baseband case and interpret \bar{x}^2 as the mean square phase deviation ϕ^2 of the modulated carrier into the demodulator:

$$\frac{P_s}{P_n} \geq \frac{\alpha^2}{2\beta^2 \bar{x}^2 + \frac{9}{2} \gamma^2 (\bar{x}^2)^2}.$$

The NPR conversion factor for 1200 channels is 19.3 db, and the loading factor is 15.8 dbm₀.

$$\begin{aligned} \left. \frac{P_{tt}}{P_n} \right|_{\substack{3.1 \text{ kc} \\ \text{weighted}}} &\geq \frac{(10^{1.93}) \alpha^2}{2\beta^2 (10^{1.58} \phi_{tt}^2) + \frac{9}{2} \gamma^2 (10^{1.58} \phi_{tt}^2)^2} \\ &\geq 10^{0.35} \alpha^2 \left[2\beta^2 \phi_{tt}^2 + \frac{9}{2} \gamma^2 10^{1.58} (\phi_{tt}^2)^2 \right]^{-1}. \end{aligned}$$

Again, this result relates the allowable non-linear level in the worst channel to the power series coefficients and the level (in this case, mean-square phase deviation) of the standard channel test-tone.

It should be noted that the results developed here are for a single non-linear element in the SSB-PM system. Non-linear noise accumulates as the signal progresses through the system (Reference 5) and the noise from different elements tends to add up on something between a voltage basis (correlated noise) and a power basis (uncorrelated noise). Worst-case performance can always be calculated by assuming correlation between the non-linear noise sources.

REFERENCES

- HOLBROOK, B. D., and DIXON, J. T., "Load Rating Theory for Multi-Channel Amplifiers," *Bell Syst. Tech. J.* 18:624-644, October 1939 p. 635, fig. 4.
- BRAY, W. J., "The Standardization of International Microwave Radio-Relay Systems," *IEE Proc. Paper* 3412E, 108B:180-200, March 1961 p. 196.
- WHITE, R. W., and WHYTE, J. S., "Equipment for Measurement of Interchannel Crosstalk and Noise on Broadband Multi-Channel Telephone Systems," *Post Off. Elect. Eng. J.* 48(3): 127-132, October 1955.
- LEWIN, L., "Interference in Multi-Channel Circuits," *Wireless Eng.* 27:294-304, December 1950.
- YEH, L. P., "Consideration of Nonlinear Noise and Its Testing in Frequency Division Multiplex Voice U.H.F. Radio Communication Systems," *IRE Trans. Commun. Syst.* CS-9(2):115-129, June 1961.

APPENDIX A

COMPUTATION OF OUTPUT AUTOCORRELATION FUNCTION

The output is

$$y(t) = \alpha x(t) + \beta x^2(t) + \gamma x^3(t) + \dots$$

and its autocorrelation function is

$$R_Y(\tau) = \overline{y(t)y(t+\tau)} .$$

Truncating the power series, we have

$$y(t)y(t+\tau) = [\alpha x(t) + \beta x^2(t) + \gamma x^3(t)] \cdot [\alpha x(t+\tau) + \beta x^2(t+\tau) + \gamma x^3(t+\tau)]$$

Carrying out the indicated multiplication,

$$\begin{aligned} y(t)y(t+\tau) = & \alpha^2 x(t)x(t+\tau) + \alpha\beta x(t)x^2(t+\tau) + \alpha\gamma x(t)x^3(t+\tau) \\ & + \beta\alpha x^2(t)x(t+\tau) + \beta^2 x^2(t)x^2(t+\tau) + \beta\gamma x^2(t)x^3(t+\tau) \\ & + \gamma\alpha x^3(t)x(t+\tau) + \gamma\beta x^3(t)x^2(t+\tau) + \gamma^2 x^3(t)x^3(t+\tau) . \end{aligned}$$

The autocorrelation function is the sum of the expectations of each of these terms.

Middleton* gives a formula for the expectation of the product of Gaussian variables which is directly applicable since $x(t)$ is Gaussian of zero mean and variance $\overline{x^2}$.

$$\begin{aligned} \text{(a)} \quad \overline{x^{2N}} &= \overline{x_1 x_2 \dots x_{2N}} = \sum_{\substack{\text{ALL} \\ \text{PAIRS}}} \left(\prod_{j \neq k}^N \overline{x_j x_k} \right) \\ \text{(b)} \quad \overline{x^{2N+1}} &= 0 \end{aligned}$$

where N is a positive integer. Formula (b) is intuitively obvious, since

$$\overline{x} = 0$$

and

$$\overline{x^{2N+1}} = \overline{x(x^{2N})}$$

Adopting the notation

$$\phi = x(t) ,$$

$$\theta = x(t+\tau) .$$

so that $R_x(\tau) = \overline{\phi\theta}$, and observing that

$$\overline{x^2(t)} = \overline{x^2(t+\tau)} = \overline{\phi^2} = \overline{\theta^2} ,$$

we note that $\overline{\phi^M \theta^N} = 0$ when $M+N = \text{an odd integer}$. This immediately does away with several

*Middleton, D., "An Introduction to Statistical Communication Theory," New York: McGraw-Hill Book Company, 1960 p. 343, eq. 7.28.

of the terms of $R_y(\tau)$. Those which exist are

$$\begin{aligned} R_y(\tau) &= \alpha^2 \overline{\phi\theta} \\ &+ \alpha\gamma \overline{\phi\theta^3} + \alpha\gamma \overline{\phi^3\theta} \\ &+ \beta^2 \overline{\phi^2\theta^2} \\ &+ \gamma^2 \overline{\phi^3\theta^3} \end{aligned}$$

Clearly, the first term is identically $\alpha^2 R_x(\tau)$. The other terms can be written out in terms of $\overline{x^2}$ and $R_x(\tau)$. The other expansions will be worked through completely, following Middleton:

$$\begin{aligned} \overline{\phi\theta^3} &= \overline{\phi\theta_1\theta_2\theta_3} = \overline{\phi\theta_1} \cdot \overline{\theta_2\theta_3} + \overline{\phi\theta_2} \cdot \overline{\theta_1\theta_3} + \overline{\phi\theta_3} \cdot \overline{\theta_1\theta_2} \\ &= 3\overline{\theta^2} \cdot \overline{\phi\theta} = 3\overline{x^2} R_x(\tau); \\ \overline{\phi^3\theta} &= \overline{\phi_1\phi_2\phi_3\theta} = \overline{\phi_1\phi_2} \cdot \overline{\phi_3\theta} + \overline{\phi_1\phi_3} \cdot \overline{\phi_2\theta} + \overline{\phi_1\theta} \cdot \overline{\phi_2\phi_3} \\ &= 3\overline{\phi^2} \cdot \overline{\phi\theta} = 3\overline{x^2} R_x(\tau); \\ \overline{\phi^2\theta^2} &= \overline{\phi_1\phi_2\theta_1\theta_2} = \overline{\phi_1\phi_2} \cdot \overline{\theta_1\theta_2} + \overline{\phi_1\theta_1} \cdot \overline{\phi_2\theta_2} + \overline{\phi_1\theta_2} \cdot \overline{\phi_2\theta_1} \\ &= \overline{\phi^2} \cdot \overline{\theta^2} + 2\overline{\phi\theta} \cdot \overline{\phi\theta} = (\overline{x^2})^2 + 2R_x^2(\tau); \\ \overline{\phi^3\theta^3} &= \overline{\phi_1\phi_2\phi_3\theta_1\theta_2\theta_3} \\ &= \overline{\phi_1\phi_2} \cdot \overline{\phi_3\theta_1\theta_2\theta_3} + \overline{\phi_1\phi_3} \cdot \overline{\phi_2\theta_1\theta_2\theta_3} \\ &+ \overline{\phi_1\theta_1} \cdot \overline{\phi_2\theta_2\theta_3\phi_3} + \overline{\phi_1\theta_2} \cdot \overline{\phi_2\phi_3\theta_1\theta_3} \\ &+ \overline{\phi_1\theta_3} \cdot \overline{\phi_2\phi_3\theta_1\theta_2} \\ &= \overline{\phi^2} [3\overline{\theta^2} \cdot \overline{\phi\theta}] + \overline{\phi^2} [3\overline{\theta^2} \cdot \overline{\phi\theta}] \\ &+ \overline{\phi\theta} [\overline{\phi^2} \cdot \overline{\theta^2} + 2\overline{\phi\theta} \cdot \overline{\phi\theta}] + \overline{\phi\theta} [\overline{\phi^2}\theta^2 + 2\overline{\phi\theta} \cdot \overline{\phi\theta}] \\ &+ \overline{\phi\theta} [\overline{\phi^2} \cdot \overline{\theta^2} + 2\overline{\phi\theta} \cdot \overline{\phi\theta}] \\ &= 9\overline{\phi^2} \cdot \overline{\theta^2} \cdot \overline{\phi\theta} + 6(\overline{\phi\theta})^3 \\ &= 9(\overline{x^2})^2 R_x(\tau) + 6R_x^3(\tau); \end{aligned}$$

Thus, finally,

$$\begin{aligned}
 R_y(\tau) &= \beta^2 \overline{x^2} \\
 &+ \left[\alpha^2 + 6\alpha\gamma \overline{x^2} + 9\gamma^2 (\overline{x^2})^2 \right] R_x(\tau) \\
 &+ [2\beta^2] R_x^2(\tau) \\
 &+ [6\gamma^2] R_x^3(\tau) .
 \end{aligned}$$

This is the complete expression for the auto-correlation function $R_y(\tau)$ of the output of a non-linear device with a third order power series representation and a Gaussian input of zero mean and variance $\overline{x^2}$.

The output spectrum is the Fourier transform of $R_y(\tau)$:

$$G_y(f) = \mathcal{F}[R_y(\tau)] = \int_{-\infty}^{\infty} R_y(\tau) e^{-j\omega\tau} d\tau$$

APPENDIX B

DETERMINATION OF POWER SERIES COEFFICIENTS

The coefficients α , β , γ , etc., of the power series representation of the non-linear device may be determined either analytically or experimentally in a number of ways. For the case of the IF spectra, where the ratio γ/α determines the non-linear noise, there are two simple ways to get γ/α experimentally.

Harmonic Margin Measurement (small signal case)

If the input to the non-linear device is a single tone

$$x(t) = V \cos \omega t$$

the output $y(t)$ will contain harmonics with relative amplitudes which can be readily determined

by a wave analyzer:

$$y(t) = A_1 \cos \omega t + A_2 \cos 2\omega t + A_3 \cos 3\omega t + \dots$$

The harmonic margins A_2/A_1 and A_3/A_1 can be used to evaluate the relative values of α , β , and γ . We have, generally,

$$\begin{aligned} y(t) &= \alpha x(t) + \beta x^2(t) + \gamma x^3(t) + \dots \\ &= \alpha V \cos \omega t + \beta V^2 \cos^2 \omega t + \gamma V^3 \cos^3 \omega t \\ &= \alpha V \cos \omega t + \beta V^2 \left[\frac{1}{2} + \frac{1}{2} \cos 2\omega t \right] + \gamma V^3 \left[\frac{3}{4} \cos \omega t + \frac{1}{4} \cos 3\omega t \right] + \dots \end{aligned}$$

or, combining terms and dropping the dc contribution,

$$y(t) = \left[\alpha + \frac{3\gamma V^2}{4} \right] V \cos \omega t + \frac{\beta V^2}{2} \cos 2\omega t + \frac{\gamma V^3}{4} \cos 3\omega t + \dots$$

If $V \ll 1$ (small signal case),

$$y(t) \approx \alpha V \cos \omega t + \frac{\beta V^2}{2} \cos 2\omega t + \frac{\gamma V^3}{4} \cos 3\omega t + \dots$$

and the harmonic margins are

$$\begin{aligned} \left| \frac{A_2}{A_1} \right| &= \frac{V}{2} \left| \frac{\beta}{\alpha} \right|, \\ \left| \frac{A_3}{A_1} \right| &= \frac{V^2}{4} \left| \frac{\gamma}{\alpha} \right|; \end{aligned}$$

or conversely

$$\begin{aligned} \left| \frac{\beta}{\alpha} \right| &= \frac{2}{V} \left| \frac{A_2}{A_1} \right|, \\ \left| \frac{\gamma}{\alpha} \right| &= \frac{4}{V^2} \left| \frac{A_3}{A_1} \right|. \end{aligned}$$

Gain Variation Measurement (large signal case)

where the normalized gain is defined as

Using the same expression as above,

$$G = \frac{y(t)}{ax(t)}$$

$$y(t) = \left[a + \frac{3\gamma V^2}{4} \right] V \cos \omega t + \frac{\beta V^2}{2} \cos 2\omega t \\ + \frac{\gamma V^3}{4} \cos 3\omega t + \dots$$

If the magnitude of the input signal is varied by ΔV , there is a change in gain

$$\Delta G = \frac{3}{4} \left(\frac{\gamma}{a} \right) (\Delta V)^2$$

assuming harmonic terms are filtered out, and removing restrictions on the magnitude of V we have that

Under certain conditions this is an excellent way to measure the ratio γ/a . This ratio may be either positive or negative in sign; e.g., for a limiting amplifier, the sign would be negative.

$$\frac{y(t)}{x(t)} = a + \frac{3\gamma V^2}{4} = a \left[1 + 3 \frac{\gamma}{a} \frac{V^2}{4} \right]$$

APPENDIX C

EVALUATION OF SPECTRAL CONVOLUTIONS

To evaluate the Fourier transform of the output autocorrelation function as computed in Appendix A, we must compute the transforms

$$F(f) = \mathcal{F} [R_x^2(\tau)] = \int_{-\infty}^{\infty} R_x^2(\tau) e^{-j\omega\tau} d\tau$$

and

$$H(f) = \mathcal{F} [R_x^3(\tau)] = \int_{-\infty}^{\infty} R_x^3(\tau) e^{-j\omega\tau} d\tau,$$

where $R_x(\tau)$ is given by

$$R_x(\tau) = \int_{-\infty}^{\infty} S_x(f) e^{j\omega\tau} df = \overline{x(t)x(t+\tau)},$$

$$R_x(0) = \int_{-\infty}^{\infty} S_x(f) df = \overline{x^2(t)}.$$

We have that

$$F(f) = \int_{-\infty}^{\infty} R_x^2(\tau) e^{-j\omega\tau} d\tau = \int_{-\infty}^{\infty} R_x(\tau) \left[\int_{-\infty}^{\infty} S_x(f) e^{j\omega\tau} df \right] e^{-j\omega\tau} d\tau$$

or, more concisely, introducing a new dummy variable λ to avoid confusion,

$$\begin{aligned} F(f) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_x(\tau) S_x(\lambda) e^{j2\pi\lambda\tau} e^{-j\omega\tau} d\lambda d\tau \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_x(\tau) S_x(\lambda) e^{-j(\omega-2\pi\lambda)\tau} d\lambda d\tau; \end{aligned}$$

$$S_x(f-\lambda) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j(\omega-2\pi\lambda)\tau} d\tau;$$

$$F(f) = \int_{-\infty}^{\infty} S_x(\lambda) S_x(f-\lambda) d\lambda = S_x(f) * S_x(f)$$

This result is known as the Borel or convolution theorem. Similarly,

$$\begin{aligned} H(f) &= \int_{-\infty}^{\infty} S_x(\theta) d\theta \int_{-\infty}^{\infty} S_x(\lambda) S_x(f-\lambda) d\lambda \\ &= S_x(f) * S_x(f) * S_x(f) \end{aligned}$$

It is required to compute $S_x(f) * S_x(f)$ and S_x

$(f) * S_x(f) * S_x(f)$ for baseband and IF input spectra as shown below in Figures C1 and C2.

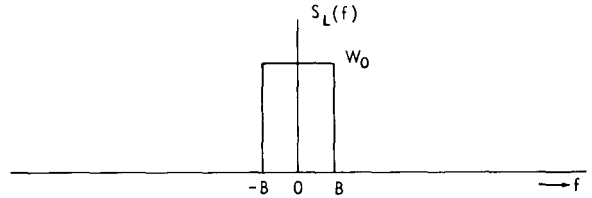


FIGURE C1.—The input baseband spectrum

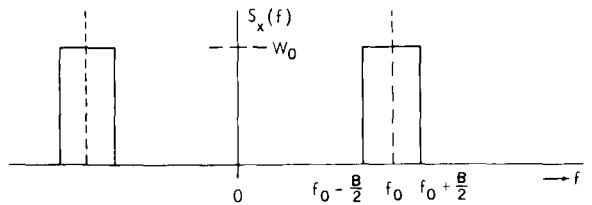


FIGURE C2.—The input IF spectrum

To do this, we start with the baseband spectrum (which we now designate $S_L(f)$, and compute $S_L(f) * S_L(f)$ and $S_L(f) * S_L(f) * S_L(f)$. These results are then used to get the results for the IF spectrum.

To get $S_L(\lambda-f)$, $S_L(f)$ is displaced to the left (for positive λ) and then folded about the point $f=0$ (Figure C3).

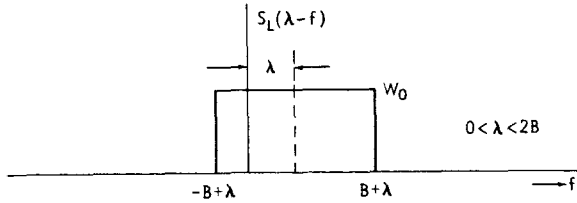


FIGURE C3.—The spectrum $S_L(\lambda-f)$ for a small positive displacement λ .

For values of displacement less than $-2B$, there is no overlap and the convolution is zero. At $\lambda = -2B$, the displaced function starts to overlap $S_L(f)$, and the convolution has a value

$$S_L(f) * S_L(f) = \int_{-B}^{B+\lambda} (W_0)(W_0) df = W_0^2(2B + \lambda).$$

This expression holds for values of displacement

$$-2B < \lambda < 0.$$

For positive values of displacement, the limits of the integral must be changed, and the convolution (Figure C4) is

$$S_L(f) * S_L(f) = \int_{-B+\lambda}^B W_0^2 df = W_0^2(2B - \lambda).$$

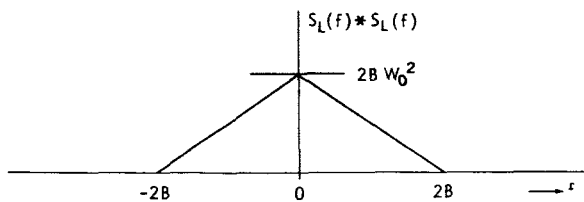


FIGURE C4.—The spectrum of the convolution $S_L(f) * S_L(f)$.

The abscissa is frequency, which we label f (instead of λ), and the ordinate is spectral density. Note that the symmetry of $S_L(f)$ produces a convolution which is an even function of frequency.

To evaluate the iterated convolution $S_L(f) * S_L(f) * S_L(f)$, the piecewise technique described by Lee* is used. The convolution $S_P(f) * S(f)$ as computed above is broken into two auxiliary functions $g(f)$ and $h(f)$, each is convolved separately

*Lee, Y. W., "Statistical Theory of Communication," New York: John Wiley & Sons, 1960 p. 24.

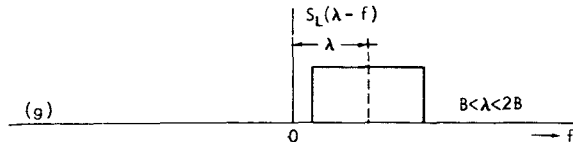
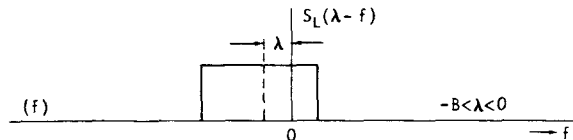
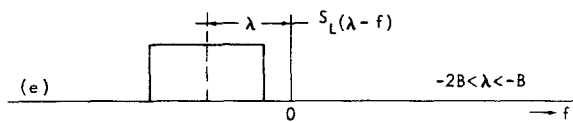
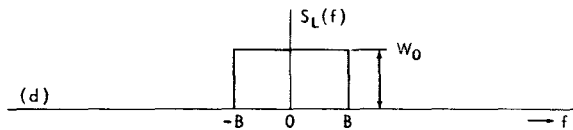
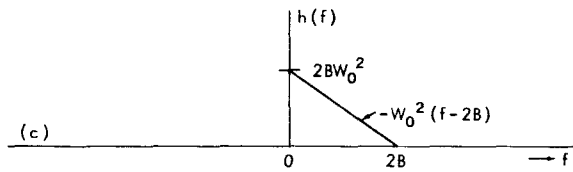
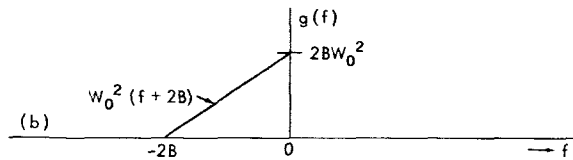
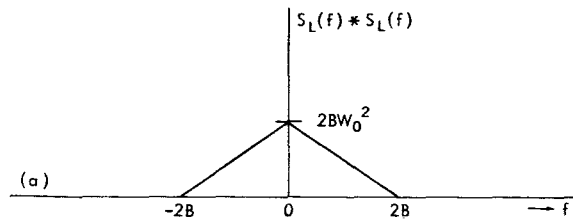


FIGURE C5.—The auxiliary functions $g(f)$ and $h(f)$, and three conditions of displacement of $S_L(\lambda-f)$.

rately with $S_L(f)$, and the results are superposed. In Figure C5 (a through g) the auxiliary functions

$g(f)$ and $h(f)$ are shown, and three conditions of displacement of $S_L(\lambda-f)$ are shown. The "box-car" slides from left to right as λ goes from $-\infty$ to ∞ .

By careful inspection, it is observed that the convolution of $S_L(f)$ with $g(f)$ is given by

$$g(f) * S_L(f) = I_1 + I_2$$

where

$$I_1 = \int_{-2B}^{\lambda+B} (W_0) (W_0^2) (f+2B) df, \quad -3B \leq \lambda \leq -B,$$

where

$$I_3 = - \int_0^{\lambda+B} (W_0) (W_0^2) (f-2B) df, \quad -B \leq \lambda \leq B,$$

$$I_4 = - \int_{\lambda-B}^{2B} (W_0) (W_0^2) (f-2B) df, \quad B \leq \lambda \leq 3B;$$

and is zero for values of displacement

$$-B > \lambda$$

$$3B < \lambda.$$

The iterated convolution is then

$$S_L(f) * S_L(f) * S_L(f) = I_1 + I_2 + I_3 + I_4.$$

Particular attention must be paid to the exact domain of definition of each of the integrals.

The four integrals are evaluated, graphed, and superposed to yield the result. The first two are

$$I_1 = \int_{-2B}^{\lambda+B} W_0^3 (f+2B) df = W_0^3 \left[\frac{f^2}{2} + 2Bf \right]_{-2B}^{\lambda+B};$$

$$I_1 = \frac{1}{2} W_0^3 [9B^2 + 6B\lambda + \lambda^2], \quad -3B \leq \lambda \leq -B;$$

$$I_2 = \int_{\lambda-B}^0 W_0^3 (f+2B) df = W_0^3 \left[\frac{f^2}{2} + 2Bf \right]_{\lambda-B}^0;$$

$$I_2 = \frac{1}{2} W_0^3 [3B^2 - 2B\lambda - \lambda^2], \quad -B \leq \lambda \leq B;$$

and

$$I_2 = \int_{\lambda-B}^0 (W_0) (W_0^2) (f+2B) df, \quad -B \leq \lambda \leq B;$$

and is zero for values of displacement

$$-B > \lambda$$

and

$$3B < \lambda.$$

Likewise, we have that the convolution of $h(f)$ with $S_L(f)$ is given by

$$h(f) * S_L(f) = I_3 + I_4$$

in which we note that

$$I_1(-3B) = I_2(B) = 0.$$

$$I_1(-B) = I_2(-B) = 2B^2 W_0^3,$$

$$I_2(0) = \frac{3}{2} B^2 W_0^3.$$

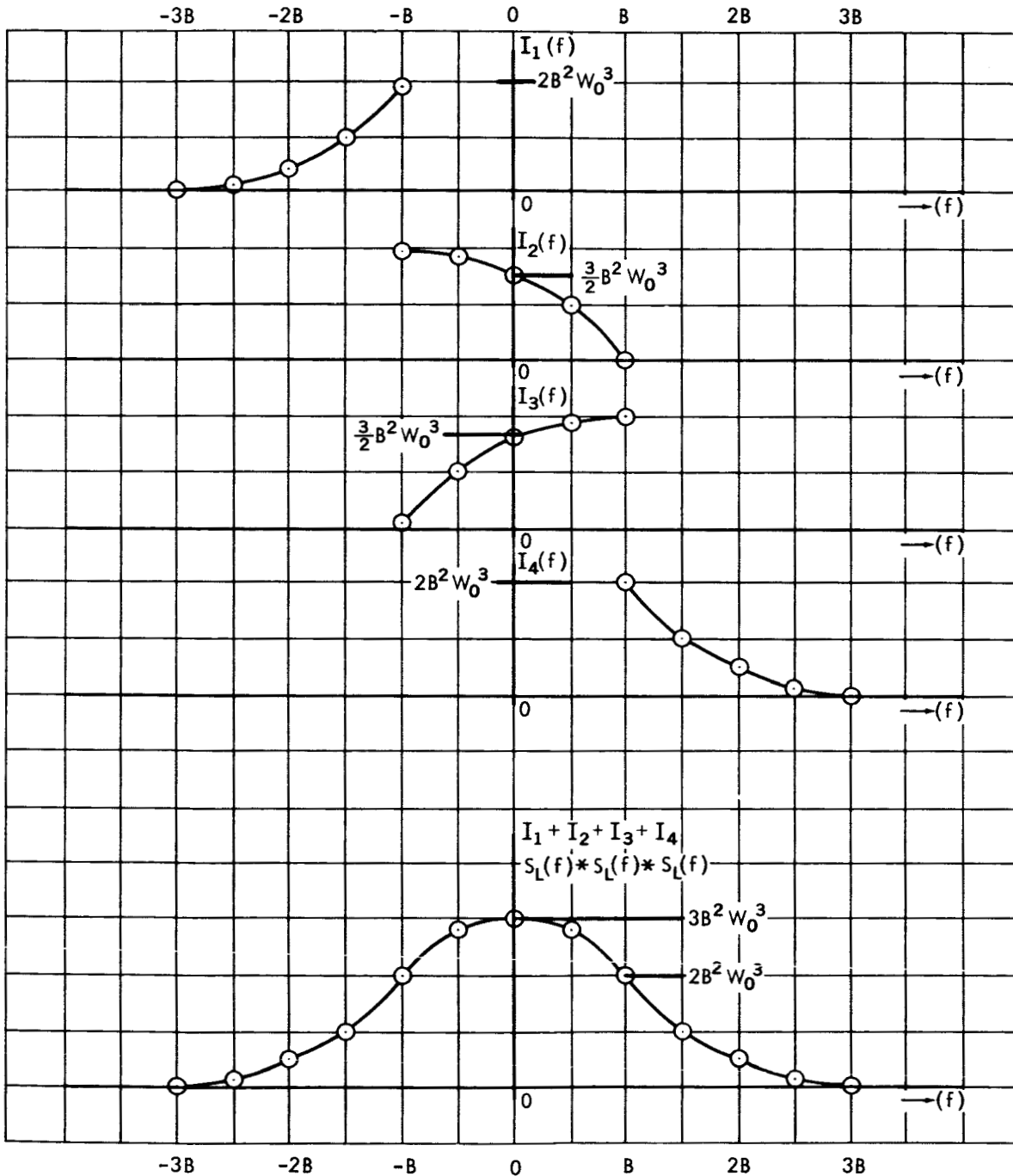


FIGURE C6.—Superposition of $I_1, I_2, I_3,$ and I_4 to give the iterated convolution $S_L(f) * S_L(f) * S_L(f)$.

and the second two integrals are

$$I_3 = - \int_0^{\lambda+B} W_0^2 (f - 2B) df = W_0^3 \left[\frac{f^2}{2} - 2Bf \right]_0^{\lambda+B};$$

$$I_3 = - \frac{1}{2} W_0^3 [-3B^2 - 2B\lambda + \lambda^2], \quad -B \leq \lambda \leq B;$$

$$I_4 = - \int_{\lambda-B}^{2B} W_0^3 (f - 2B) df = W_0^3 \left[\frac{f^2}{2} - 2Bf \right]_{\lambda-B}^{2B};$$

$$I_4 = - \frac{1}{2} W_0^3 [-9B^2 + 6B\lambda - \lambda^2], \quad B \leq \lambda \leq 3B;$$

in which we note that

$$I_3(-B) = I_4(3B) = 0,$$

$$I_3(B) = I_4(B) = 2B^2 W_0^3,$$

$$I_3(0) = \frac{3}{2} B^2 W_0^3.$$

The superposition of the four integrals to give the iterated convolution $S_L(f) * S_L(f) * S_L(f)$ is shown in Figure C6.

Using these results, we can now evaluate the result for the IF spectrum by inspection. The calculated convolutions are shown in Figure C7.

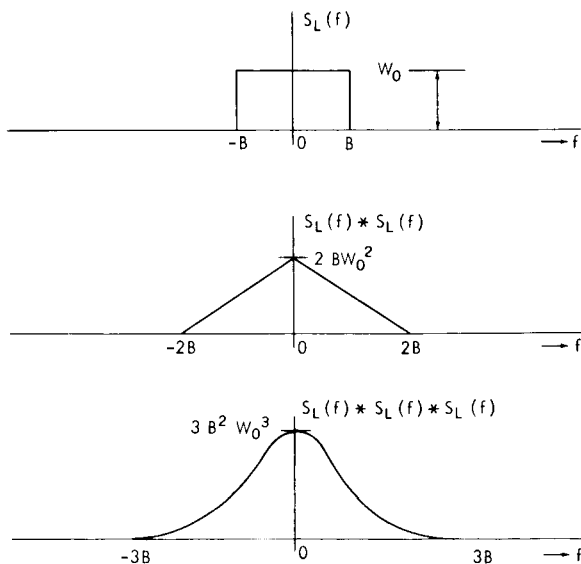


FIGURE C7.—The spectra of the convolutions $S_L(f)$, $S_L(f) * S_L(f)$, and $S_L(f) * S_L(f) * S_L(f)$.

Auxiliary convolutions in slightly different notation can be derived by inspection and are shown in Figure C8.

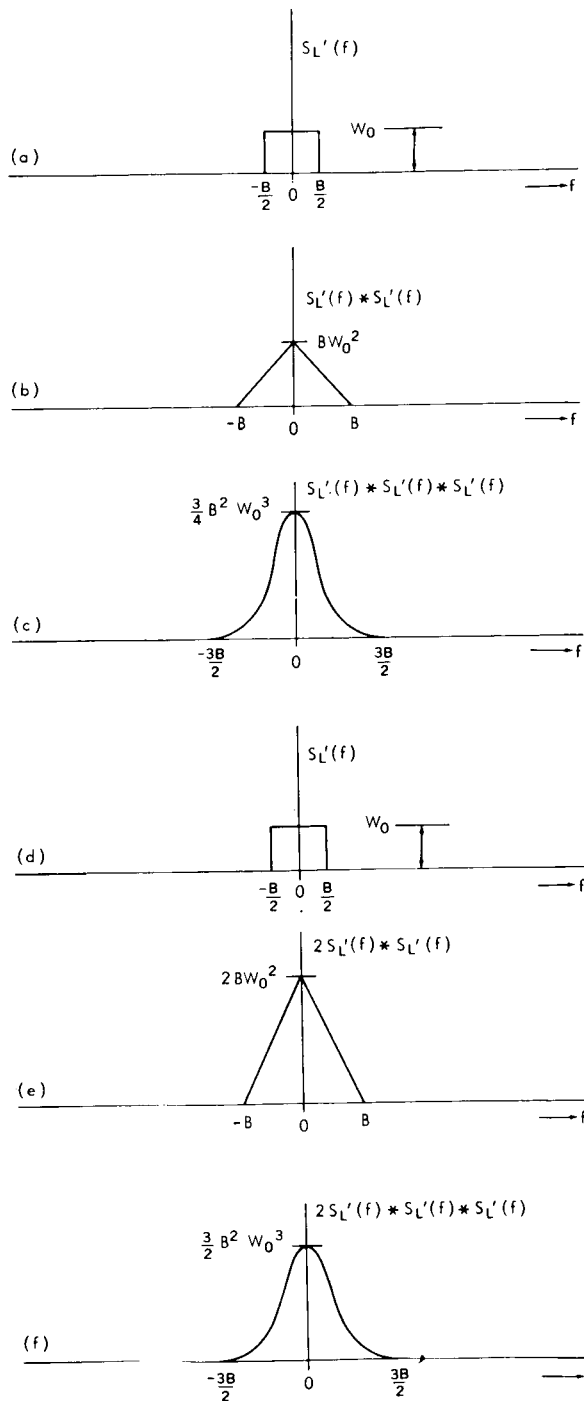


FIGURE C8.—Auxiliary convolutions.

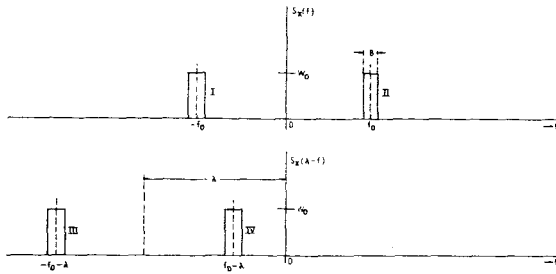


FIGURE C9.—The input IF spectra $S_X(f)$ and $S_X(\lambda - f)$.

The IF spectra $S_X(f)$ and $S_X(\lambda - f)$ are shown in Figure C9. As λ increases from $-\infty$, boxcar IV will overlap boxcar I as $\lambda \rightarrow -sf_0$ generating a triangular spectrum (Figure C8-b). As the displacement nears zero, boxcars III and IV will overlap boxcars I and II, respectively, at the same time, generating a triangular spectrum centered at zero frequency (Figure C8-e). As λ increases positively, a triangular spectrum (Figure C8-b) is generated at $\lambda = 2f_0$, and the spectrum of the convolution is shown in Figure C10.

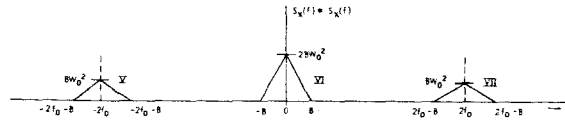


FIGURE C10.—The spectrum of $S_X(f) * S_X(f)$.

If we now convolve this second order spectrum with $S_X(f)$, we see that as λ increases from $-\infty$ and nears $-3f_0$, boxcar IV (Figure C9) overlaps triangle V, generating a bell-shaped spectrum (Figure C8-c) at $-3f_0$. As λ increases to $-f_0$, boxcar IV overlaps triangle VI simultaneously as boxcar III overlaps triangle V, generating a bell shaped spectrum (Figure C8-f) at $-f_0$.

For positive λ , a mirror image results, and we have the result shown in Figure C11.

It can be seen that a further convolution of $S_X(f)$ will produce fourth order distortion spectra at 0 , $\pm 2f_0$, and $\pm 4f_0$. This is not computed because such spectra may be readily filtered out, as was noted earlier.

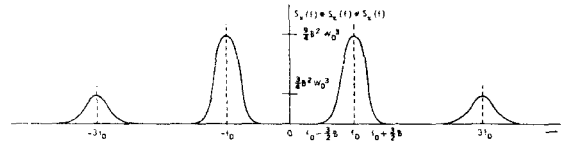


FIGURE C11.—The spectrum of $S_X(f) * S_X(f) * S_X(f)$.

APPENDIX D

DISCUSSION OF CHANNEL LOADING FACTOR, NPR CONVERSION FACTOR, AND MULTICHANNEL PEAK FACTOR

To avoid ambiguity, it is convenient to specify the intensity of a multichannel signal in terms of the power of a single channel test-tone at a point of zero relative level. The CCIR recommends desirable multichannel signal levels and allowable noise levels in dbm_0 , or decibels with respect to one milliwatt at zero relative level. The phrase "zero relative level" refers to some physical point

in the communication system where channel test-tone level happens to be exactly 1 mw, or 0 dbm_0 . The statement that for 1200 channels the multichannel signal is 15 dbm_0 means that at any point in the system, the multichannel signal power is 15 db above the channel test-tone power, and at zero relative level it is exactly 15 dbm_0 or 31.6 milliwatts. The CCIR recommended loading

factors for busy-hour traffic are

$$10 \log_{10} \frac{P_s}{P_{tt}} = -1 + 4 \log_{10} N \quad \text{dbm}_0, \quad 12 \leq N \leq 240 ;$$

and

$$10 \log_{10} \frac{P_s}{P_{tt}} = -15 + 10 \log_{10} N \quad \text{dbm}_0, \quad N > 240 .$$

These equations are graphed as curves (a) and (b) in Figure D1.

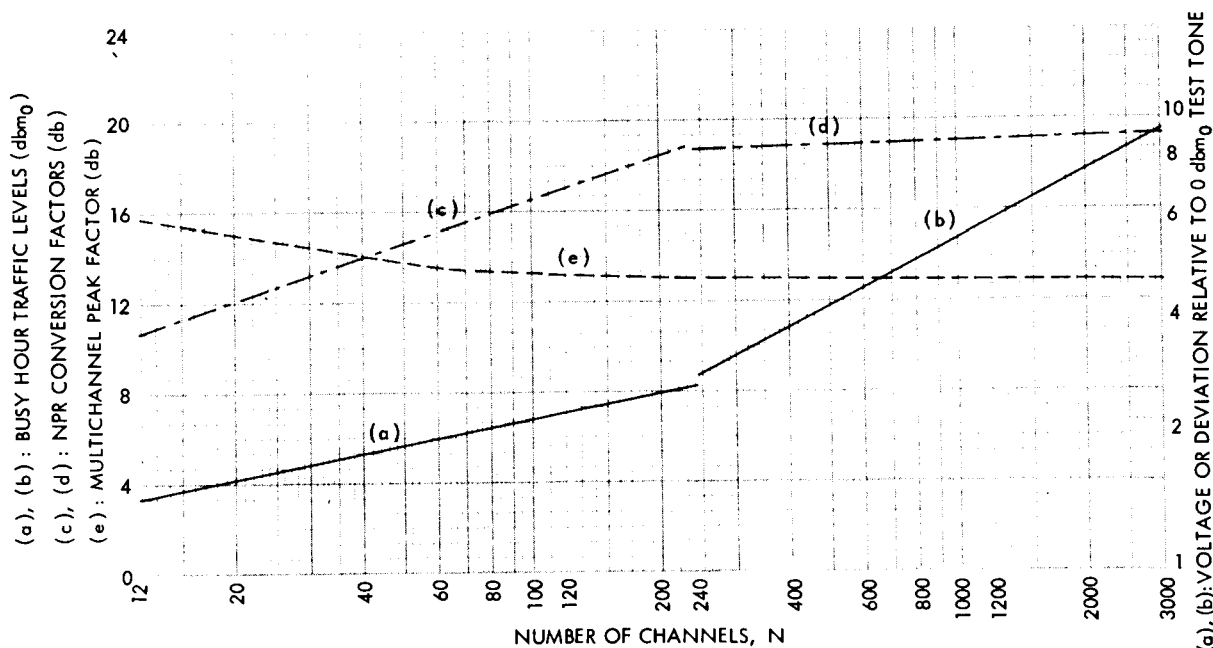


FIGURE D1.—CCIR channel loading factors and conversion factors, and Holbrook-and-Dixon peak factors. The conversion factor is added to the noise power ratio to give the S/N ratio of 1 mw of tone and noise power in a 3.1 kc band, psophometrically weighted. The unweighted noise power in a 3.1 kc band is 2.5 db greater than the psophometrically weighted noise.

The CCIR noise-power-ratio conversion factor may be regarded as an allowance for the difference between channel test-tone level and the portion of the noise-power loading which is effectively applied to any one channel in a noise-loading test; i.e., for a 1200 channel system, the loading factor is 15.8 dbm₀, and the baseband is 5.6 Mc wide. The noise power falling into a 3.1 kc slot is 3.1/5600 of the applied power or -32.6 db with respect to +15.8 dbm₀; and it is -32.6+15.8 = -16.8 db with respect to 0 dbm₀, the channel test-tone power at zero relative level. Psophometric filtering reduces the noise by 2.5 db, so the

NPR conversion factor for 1200 channels is 16.8 +2.5 or 19.3 db. CCIR NPR conversion factors are given as curves (c) and (d) of Figure D1.

For a single speech channel, peaks as high as 19 db above the rms level are encountered. As many speech channels are multiplexed together, the composite signal peak-to-average ratio falls off, and for more than 800 channels it levels off at about 13 db. Multichannel peak factors from Holbrook and Dixon* are given as curve (e) in Figure D1.

*Holbrook, B. D., and Dixon, J. T., "Load Rating Theory for Multi-Channel Amplifiers," *Bell Syst. Tech J.* 18:624-644, October 1939.

THE AUTOMATIC PICTURE TRANSMISSION SYSTEM ON TIROS VIII*

DAVID W. HOLMES† AND CHARLES M. HUNTER

The automatic picture transmission (APT) system was developed to provide immediate readout of local cloud-cover data from weather satellites. This camera system takes pictures regularly during the sunlit portion of the orbit and transmits them automatically; thus suitably equipped ground stations are able to receive cloud photographs for the area immediately surrounding the station. The receiving and recording equipment required is relatively simple and inexpensive. Ground tests of the APT equipment were made and a flight test is being conducted with TIROS VIII. The objective of this test is to determine the feasibility of the APT concept. Some problems have been encountered and corrective action has been taken or initiated. The results of this first experiment are encouraging and indicate that when the APT picture information is regularly available it will be useful in local analysis and forecasting programmes.

The automatic picture transmission system (APT), developed by the United States for use on meteorological satellites, is a slow scan television system designed to take and immediately transmit cloud pictures for reception by a relatively inexpensive television ground station consisting of a rotatable antenna, a receiver, and a recorder. The APT sub-system is completely automatic and is independent of other sub-systems on the spacecraft except for power and a frequency reference. Many types of standard mobile communications equipment are readily adaptable for use in receiving the signals transmitted from this sub-system on the satellite.

OBJECTIVE

The APT camera system, designed originally for more advanced meteorological satellites, is receiving extensive evaluation in a preliminary test aboard TIROS VIII, launched 21 December 1963. The objective of this test is to determine the feasibility of the APT system concept, both satellite and ground receiving equipment. Some 50 to 60 ground stations, located throughout the world, are acquiring the pictures. Meteorologists at these locations are evaluating the quality and

utility of the data while engineering and scientific personnel at certain key ground stations are evaluating system operation. As a result, some modifications and improvements have been made and others may be recommended.

SATELLITE EQUIPMENT^{1**}

The APT camera sub-system shown in Figure 1 includes the optics, shutter, a one-inch, 800-line electrostatic storage vidicon, camera electronics, and an FM transmitter. The camera is programmed for automatic repetition of the prepare, expose, and readout cycle. When the shutter is triggered, the image is projected on the

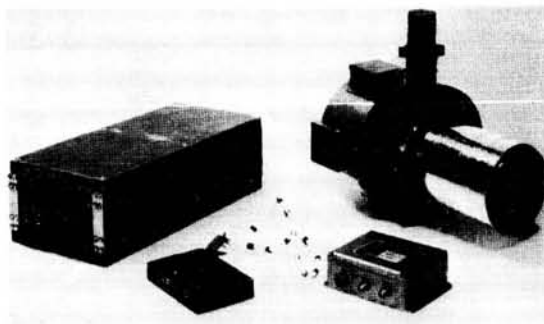


FIGURE 1.—APT camera subsystem.

*Published in *World Meteorological Organization Bulletin*, 13(3):128-134, July, 1964.

†National Weather Satellite Center of the U.S. Weather Bureau.

**See references.

photoconductive layer of the vidicon through a 108 degree lens (measured on the picture diagonal) and then electronically *developed*; that is, the is transferred to the polystyrene storage layer. This occurs during the first eight seconds of every picture sequence. During the remaining 200 seconds, the picture information is read from the vidicon at a scanning rate of four lines per second. The video signal is used to amplitude modulate a 2400 cps subcarrier which in turn frequency modulates the 136.950 Mc/s, 5 watt, solid state FM transmitter. The picture information has approximately seven shades of grey with a change in density of 0.12 per step. At the beginning of the picture sequence a 300 cps, lasting three seconds, and then five seconds of phasing pulses are transmitted to start and phase the ground station facsimile equipment automatically.

The ground stations (see Figure 2) receive the pictures in the form of facsimile reproductions of those transmitted by the automatic picture transmission equipment. This equipment consists basically of a receiving antenna, preamplifier, FM receiver, and a facsimile recorder. The

antenna is an 8-turn helix, 14 feet (4.3 m) long, 27 inches (68 cm) in diameter. It has a beam-width of 34° at the half power points and a gain of 12.5 db. Position control and indicator units, located in the ground station console, provide independent rate control of both azimuth and elevation and continuous display of the antenna position. The preamplifier, mounted on the antenna pedestal, is a two-stage RF amplifier with a gain of 22 db. The preamplifier is used to compensate for losses in the radio frequency cable when the antenna and the receiver must be located some distance apart. The FM receiver is crystal-controlled and can be tuned over a range of 150 kc/s on either side of the operating frequency of 136.950 Mc/s. It has a selectable bandwidth of 50 or 100 kc/s and has a noise figure of 4 db. The facsimile recorder, a helix and writing-blade-type machine operating at 240 rpm, uses electrosensitive (wet) paper that forms the picture by depositing ions on the paper. The recorder will produce an eight-inch-square picture with ten shades of grey and 100 line-per-inch resolution. The ground station can acquire pictures during

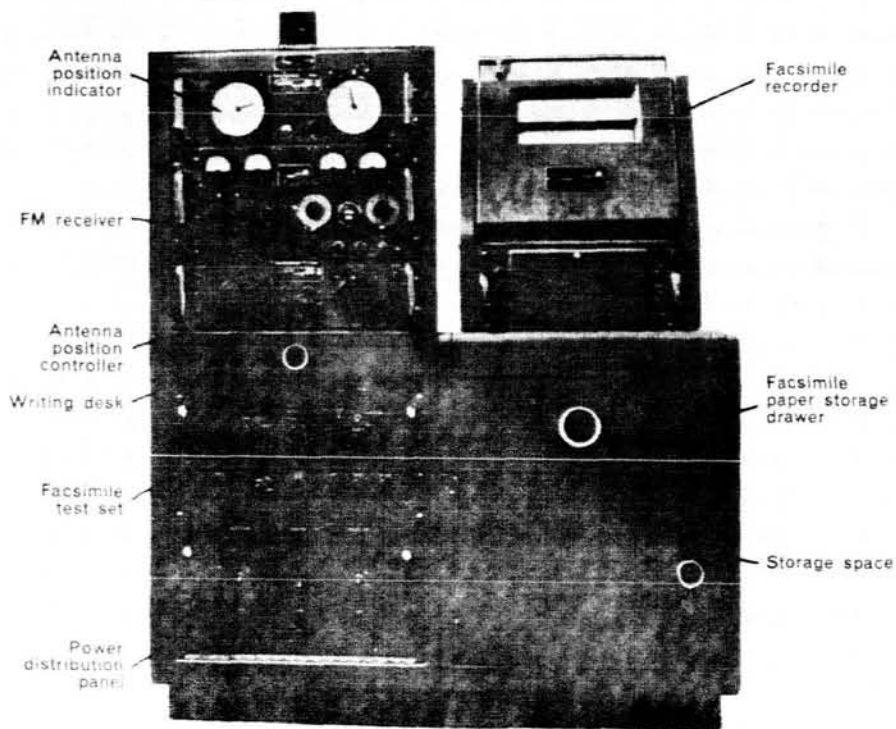


FIGURE 2.—APT ground station console.

the time the satellite is at least 2 to 5 degrees above the horizon of the receiving station. The signals received while the ground station antenna is directed toward the satellite are amplified and demodulated for reproduction by a facsimile recorder. The recorder reconstitutes the transmitted picture line-by-line during the 200 seconds of reception.

The United States has purchased a number of ground stations which have been located at major forecasting offices for evaluation during the test of APT on TIROS VIII and on the first Nimbus meteorological satellites. Ground equipment has been installed, and is being operated in Australia, Canada, Denmark, France², Hong Kong, India, Indonesia, and the United Kingdom. Most of these have been fabricated locally.

A daily message transmitted over the meteorological teletypewriter circuits, and a weekly message sent by mail, provide the operational data required by the meteorologists using the APT ground station. The weekly message is used by APT stations only when the daily messages are not received for a prolonged period of time; it contains essentially the same information as the daily message, but for an extended period of about two weeks. This message is mailed two weeks before its effective date. The data in these messages are used to determine the azimuth and elevation angles for following the satellite with the antenna, and for subsequent orientation and analysis of the picture. A polar stereographic latitude-longitude projection for a complete hemisphere plus special transparent overlays are used in plotting the satellite track.

PICTURE COVERAGE AND USAGE

Pictures taken by the APT camera when it is looking straight down from the 380 n.-mile altitude of TIROS VIII encompass an area of about 800 n. miles on a side. Pictures about 1050 n. miles on a side can be acquired from a satellite at an altitude of 500 n. miles and about 1400 n. miles on a side from a satellite at an altitude of 600 n. miles. The same limitations for latitudinal and longitudinal picture coverage apply to the APT on TIROS VIII as to the present TIROS cameras³. In addition, power limitations restrict the APT on TIROS VIII to about seven pictures per orbital pass over the sunlit portion of the orbit.

During each orbital pass within acquisition range of a ground station (1800 n. miles from the sub-satellite track) one to four pictures can be obtained (with a probable average of between two and three pictures per pass). At any given location, pictures can be obtained from one to three TIROS passes per day.

The number of TIROS passes acquired depends on the latitude of the station and the sun angle. Fewer passes can be acquired by stations located near the equator than by those located near latitude 50°. If the camera equipment is flown on future meteorological satellites orbiting in polar or quasi-polar orbits, any station with APT ground equipment can expect to acquire pictures from at least one pass each day. On the average, about three pictures per pass will be acquired. In optimum latitudinal locations, i.e., latitudes of 45 degrees or more from the equator, a ground station may expect to receive pictures from at least two passes per day.

Timely, detailed cloud pictures of an area provide supplementary observational data for use in local weather analysis. Data extracted from the APT pictures can be integrated with synoptic data such as surface observations, PIREPS, RAREPS,* etc. Over local areas and over areas where data are sparse, analyses can be derived by relating cloud features to established synoptic models. Cloud pictures permit identification of weather systems, and provide an observation on which to base the local forecast when emergencies preclude normal receipt of data from other stations. The APT picture data acquired from the satellite can also be used in the verification of previously issued forecasts.

RESULTS

On the fourth orbit of the TIROS VIII satellite, or about seven hours after launch, the first APT pictures were received at ground stations in the eastern United States. These pictures showed almost complete cloud cover from the Great Lakes to the Atlantic Ocean and also showed landmarks along the east coast of the United States.

Most of the stations have reported satisfactory equipment operation with good signal reception. However, degradation of the pictures by a

*Coded reports from aircraft in flight (PIREP) and of ground radar observations (RAREP).

scalloping effect has limited their meteorological usefulness. The scalloping effect is caused by interaction between the earth's magnetic field and the vidicon scanning beam as the satellite spins through approximately 33 revolutions during the 200-second picture scanning and transmission period. Despite this effect, reports received here state that significant meteorological data have been extracted from the pictures and used to modify analyses and forecasts.

Two approaches were made toward compensating at the ground station for the earth's magnetic field modulation of the vidicon scanning beam. The first approach was an electromechanical system which would resolve both the vertical and horizontal error. The error, a sinusoidal signal at a frequency of 1/6 cps, was detected and used as a reference by appropriate electronic circuits to vary the helix motor drive to obtain the correct phase relationship for horizontal correction. Vertical correction was obtained by mechanically moving the writing blade sinusoidally about its normal position. This method was very difficult to use automatically because of the need for synchronization of the facsimile printer bar and helix.

Another approach to the problem was to use the 1/6 cps error signal with an electronic system to obtain horizontal correction. Intensity correction was provided by using a variable gain video amplifier in the facsimile machine. The detected 1/6 cps error signal was also used to control the video gain. The vertical displacement was not reduced by this method. The retention of this effect results in a one to two per cent error in location of cloud or ground objects in the vertical direction on the picture.

Figure 3 is an uncorrected picture. Figure 4 has electronic correction for both horizontal displacement and intensity error. This method appears to be a better solution than electronic horizontal correction only; however, it is fairly complex and the amount of improvement obtained does not appear to make it worth while for use at field stations.

Figure 5 illustrates the worst case of scalloping and the *venetian blind* effect while Figure 7 shows very little of these effects. The latter picture was taken near latitude 22°N on a north-bound pass of the satellite. In this case the interaction



FIGURE 3.—Uncorrected APT photograph.



FIGURE 4.—As Figure 3, but with electronic horizontal and intensity correction.

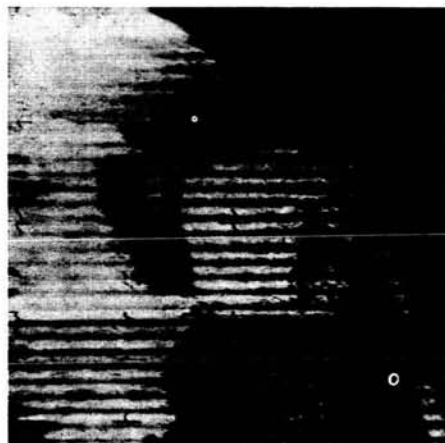


FIGURE 5.—APT picture received at National Weather Satellite Center from pass 674 of Tiros VIII at 2129 GMT, 5 February 1964.

between the earth's magnetic field and the vidicon scanning beam was at a minimum because of the satellite's position and attitude.

Figure 5 shows a photograph acquired at 2129 GMT on 5 February, 1964 from pass 674 of TIROS VIII, at several stations on the east coast of the United States. Although the picture is badly scalloped, useful meteorological information was derived from it. A nephanalysis based on this picture (Figure 6) was transmitted over the facsimile circuits in the United States. This storm,



FIGURE 6.—Nephanalysis of APT picture shown in Figure 5.

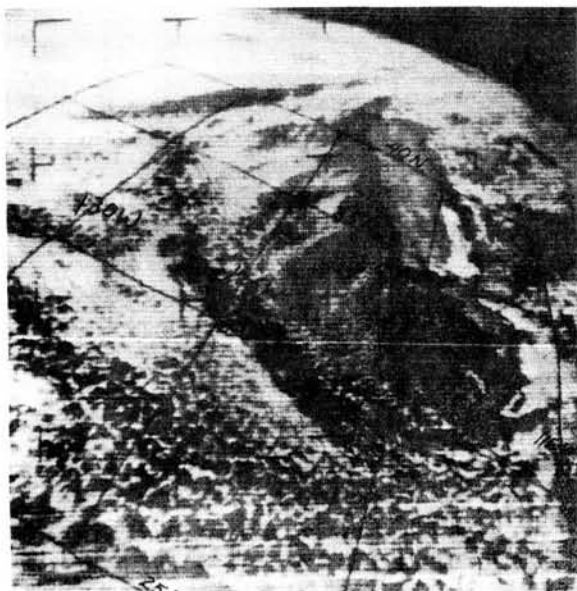


FIGURE 7.—APT picture received from pass 832 of Tiros VIII at 1900 GMT, 16 February 1964.

whose vortex is shown in Louisiana, moved north-eastward and produced widespread rain and snow along the east coast of the United States. Figure 7 was acquired from pass 832 of TIROS VIII at 1900 GMT 16 February 1964. The most significant cloud features are the frontal cloud band extending across the upper portion of the picture from a low pressure centre not visible on the horizon, and the spectacular cumuliform clouds to the south-west of the ridge line represented by the nearly clear band extending from 35°N, 130°W to 30°N, 120°W. This latter cloud formation is characteristic of the leading edge of a high-pressure area with its flow of cold air over the warm water surface. In this case the high pressure is centered at 37°N, 138°W. The bright image in the area of the San Joaquin Valley of California is low stratus and fog, typical for the area at this time of year. Both clouds and snow are observed on the Rocky Mountains.

Approximately 50 to 60 per cent of the pictures have been analysed and useful meteorological information obtained from them. One station reported that 80 per cent of the received photographs were useable. The New York International Airport Weather Bureau personnel report that there is little question as to the usefulness of APT pictures over areas of sparse data. Its usefulness over the U.S.A., even with the relatively dense observational network, cannot be disregarded. Many cases have been cited to indicate that the possibility exists for the development of techniques utilizing APT readout for use in terminal forecasting.

A photograph acquired at New York and at other locations, from orbit 845 of TIROS VIII at 1716 GMT on 17 February 1964, indicated considerable cloudiness in the Gulf of Mexico with an apparent wave centre near 24°N, 92°W. The analyses were carrying a warm front south of that area but there were no surface reports available to indicate the position of the front or the development of the wave. The National Meteorological Center prognostic charts forecast a wave development for the next day near New Orleans. This photo provided information that a development did indeed exist and considerably increased confidence in the forecasts. This particular storm affected the north-eastern part of the United States 36 to 48 hours later.

CONCLUSIONS

We are encouraged by the results of this first experiment and feel that after necessary corrective actions have been taken and further improvements made in the system, it will afford a unique opportunity for obtaining cloud-cover information on a routine basis for a substantial area around each receiving station, at a relatively low cost. We feel this information will be significantly useful to meteorologists in local analysis and forecasting programmes.

FUTURE

The scalloping and *venetian blind* effect on the pictures now being received will be corrected on APT systems flown in future meteorological satellites. The next meteorological satellite to carry the APT will be the Nimbus A. It is expected that this earth-oriented research and development satellite will be launched in early summer 1964 from the Pacific Missile Range, into a 500 n.-mile sun-synchronous orbit. In this orbit Nimbus A will cross the equator at local noon on each of its orbital passes. APT will also be carried on the TIROS Operational System (TOS) series of satellites which will be the wheel configuration of TIROS. The spin axis of the wheel configuration will be perpendicular to the plane of the satellite orbit. The cameras will be mounted looking out of the rim of the spacecraft. A horizon-sensing system will trigger the cameras when they are looking straight down. The first

of the series of wheel configuration satellites, known as TIROS I (eye) is expected to be launched later in 1964. It will not, however, contain an APT. The initiation of the TOS programme is expected to be in autumn 1965. The Nimbus and TOS satellites will provide pictures to the meteorologist that will not require attitude determination as in the case of the pictures from the space-oriented TIROS satellites.

It must be emphasized that the APT is in an experimental stage and some modification of components on the satellite and in the ground station may be required after evaluation of the results of these first experiments. It is suggested that those considering purchase or construction of APT ground stations be fully aware that modifications to the equipment may become necessary upon completion of the tests. It is not expected that APT will be in regular operational use until late 1965, but it will be tested thoroughly in the interim.

REFERENCES

1. STAMPGL, R. A., and STROUD, W. G. (1963). *The Automatic Picture Transmission (APT) Camera System for Meteorological Satellites*. NASA Technical Note NASA TND-1915. Goddard Space Flight Center, Greenbelt, Maryland.
2. WMO (1964). *Reception in France of TIROS Transmissions*. WMO Bull. Geneva. XIII, 2 pp. 71-73.
3. WMO (1963). *Reduction and Use of Data Obtained by TIROS Meteorological Satellites*. WMO, Geneva. Technical Note No. 49. pp. 7-13.

TELEMETRY INSTRUMENTATION OF THE ECHO II PASSIVE COMMUNICATIONS SATELLITE*

HAROLD S. HORIUCHI

The beacon telemetry system used on the Echo II satellite was to provide a tracking signal and to monitor the skin temperature and measure the internal pressure of the satellite, with particular emphasis on data acquisition during the initial inflation stages. A faithful TM system having very sensitive sensors with very short time constants was required since the inflation of the balloon (the satellite) was to take place in a matter of minutes.

In addition to the data acquisition requirement, there were the inevitable environmental and physical requirements that had to be satisfied. First of all, the size was limited by the space available in the overall payload package. The weight of the

The telemetry beacon that was finally designed, constructed and flown was packaged into a square, flat pancake, 14 inches square by $\frac{3}{4}$ inch thick. It weighed approximately 3 pounds of which almost half was accounted for by the two battery packs. The solar modules consisted of four triangular pancakes weighing approximately $\frac{3}{4}$ pound each. The total weight of a complete beacon system weighed just under six pounds. There were two of these systems on board the Echo II. The 12 pounds of electronic gear amounted to only two percent of the total payload.

Figure 1 shows the beacon being attached to the skin of the satellite with pressure sensitive adhesive and tapes. Figure 2 shows a better view of the solar panels in place. Figure 3 shows the balloon folded into the bottom half of the canister. A plastic enclosure will be placed over this package and the air evacuated to reduce the height of the folded balloon so that the top half of the canister can be put into place.



FIGURE 1.—Attaching beacon.

beacon system was also limited because of overall weight limit of the payload package and the concentrated mass of the beacons themselves.

*Published as *Goddard Space Flight Center Document X-621-64-281*, October 1964.



FIGURE 2.—Solar panels.



FIGURE 3.—Folded sphere.

Figure 4 is a schematic of the folded balloon packaged into the payload canister. Upon injection of the spacecraft into the orbital path, the canister opens and the folded sphere is released. It expands partially due to the residual air within the folds; later it is stressed into a rigidized sphere by means of the controlled inflation system consisting of a number of bags of a subliming material mounted strategically inside the folded balloon.

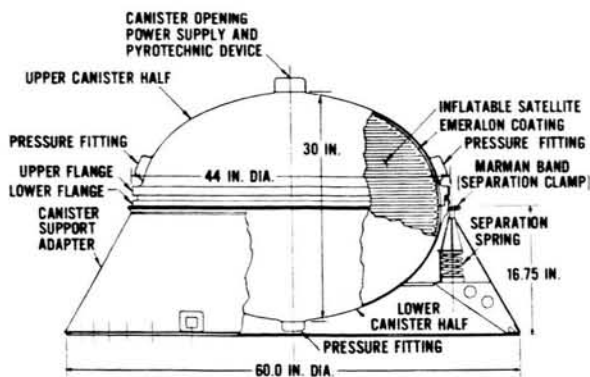


FIGURE 4.—Echo II Spacecraft (canister assembly).

Figure 5 shows a fully inflated Echo II. The photo was taken during the static inflation test held in the summer of 1963 in the famous dirigible hangar at the Lakehurst Naval Air Station, New Jersey.

Data wise, the beacons were required to sense balloon skin temperatures which were expected to vary anywhere between minus 120 degrees Centi-

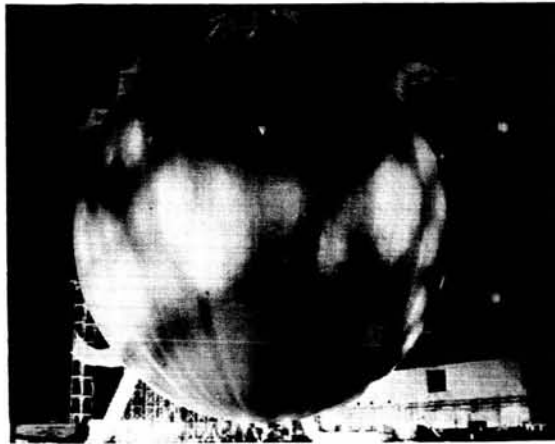


FIGURE 5.—Inflated Echo II.

grade and plus 160 degrees, a tremendous range to be covered. The internal pressure to be monitored was expected to range from a minimum of 10^{-5} mm of Hg to 0.5 mm of Hg under the expected orbital conditions. The five orders of magnitude of pressure range monitoring was also a tough requirement. Time does not permit a detailed discussion of the specific electrical, mechanical and environmental requirements and tests. (References 2, 3 and 4) Those interested may obtain this information from the Project Echo Office at the Goddard Space Flight Center. However, some interesting applications of circuits and schemes employed will be briefly described later.

Figure 6 is a block diagram of the telemetry beacon employed. Basically, the system consisted of two redundant amplitude modulated

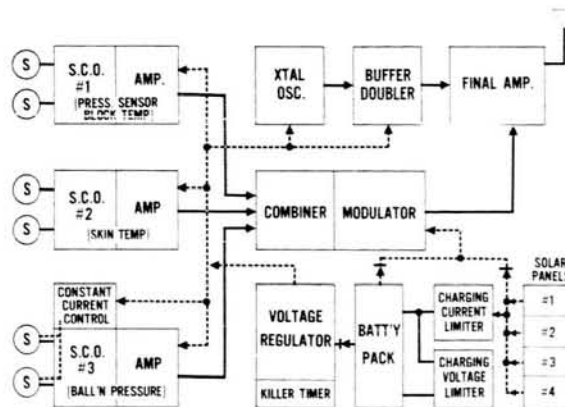


FIGURE 6.—Echo A12 telemetry beacon.

transistorized RF transmitters operating at the Minitrack frequencies of 136.020 and 136.170 megacycles, respectively. The effective radiated RF power of the two beacons was approximately 35–39 milliwatts each. Each transmitter consists of a crystal oscillator working at 68 megacycles on the fifth harmonic of a 13.6 Mc quartz crystal, followed by a frequency doubler-buffer stage and the final amplifier. Tight restrictions on the stability of the carrier frequency made the design effort extremely difficult, especially in view of the fact that temperature control of the crystal oscillator was precluded by space, weight, and other limitations. It is interesting to note, however, that a doppler frequency check was made on one of the beacons three months after launch and the carrier was found to have shifted only about 250 cycles from the last prelaunch check made a month before the actual launch. It will be noted that the specifications allowed a plus or minus 0.002 percent drift or a maximum of 2,600 cycles. During the three months in flight the beacons had gone through a complete cycle of extreme environmental temperatures and had been turned on and off at periodic intervals of about 100 seconds during this time.

Because of the extreme stability required, all of the circuits except the series modulator operated from a regulated voltage supply (See Figure 6). The dotted lines show the power connections. The solid lines indicate the signal paths. The modulator contained a waveform control circuit which employed its own feedback system and consequently had its own regulation system.

Because of the size and weight limitations again, the battery packs were limited to a total of 16 (8 each) nickel-cadmium cells with a maximum voltage of 21 volts and a capacity of 600 milliampere hours when fully charged. The regulated output was to be held at 19.2 volts by design. The four solar panels were in parallel and were to charge the batteries through a current limiting circuit whenever any solar panel voltage rose above 22 volts. The maximum voltage output from the solar panels was expected to be around 30 volts with a total power capacity of 16 watts. The beacon power requirement was only 2.8 watts including the battery charging requirement. A string of zener diodes was placed across the batteries to hold the charging voltage to approximately 24

volts to prevent damage to the batteries during the charging cycles. The solar panels were also placed in the overall circuitry such as to power the beacon in the event of battery rundown or failure.

The temperature and pressure data were to be covered through subcarrier channels, which were fed into a combiner-modulator stage. The three subcarrier circuits were essentially identical despite the different types of information to be conveyed. The subcarrier oscillators are of the Wien bridge type employing matched thermistors for frequency control.

Figure 7 shows both the Wien bridge circuit and the beacon subcarrier oscillators. The conven-

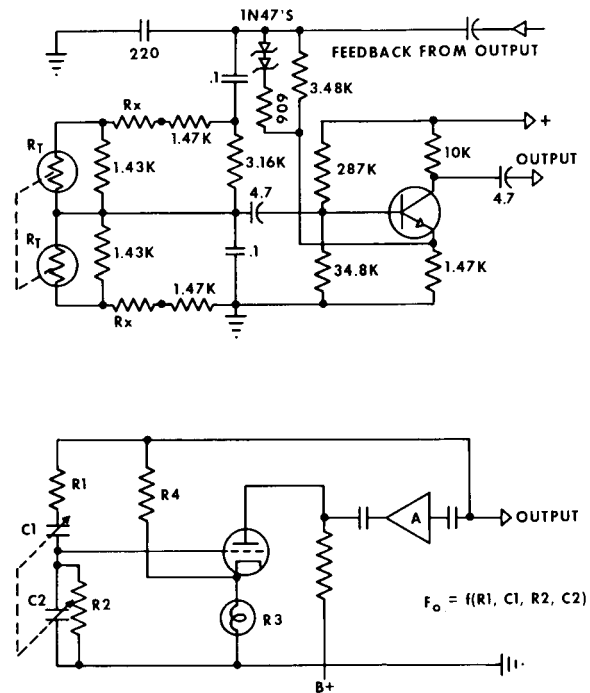


FIGURE 7.—(a) Beacon sub-carrier oscillator. (b) Wien bridge oscillator.

tional Wien bridge employs a RC circuit with capacitors in the two bridge arms being ganged and variable. In the Echo beacon the resistance elements were used as the frequency control. Since it was impossible to gang the thermosensitive elements mechanically, the thermistors were carefully matched for the nominal resistance values and temperature coefficients. Instead of employing a lamp type degeneration for output control, a feedback system was employed.

The pressure sensing, however, was slightly more involved. Here a scheme similar to the Hasting airflow gage was used but using fine thermistor beads instead of thermopiles. The design for an accurate low pressure measurement was first suggested by Ainsworth and Flanick (Reference 1) of the Goddard Space Flight Center.

Figure 8 is an attempt to illustrate in a simplified manner the temperature sensor assembly and mounting. The thermistors were mounted in a metallic cavity with a small opening exposed to the gas being monitored. The opening was protected by a radiation shield to prevent the effect of direct heat radiation from affecting the thermistor temperature. Space does not permit a detailed discussion of the principles. Bombardment of the thermistors in the cavity by the gas molecules lowers the thermistor temperature and hence the resistance and the oscillator frequency. The higher the pressure the greater the molecular activity of the gas particles and the greater probability of collision with thermistors and hence the cooling effect.

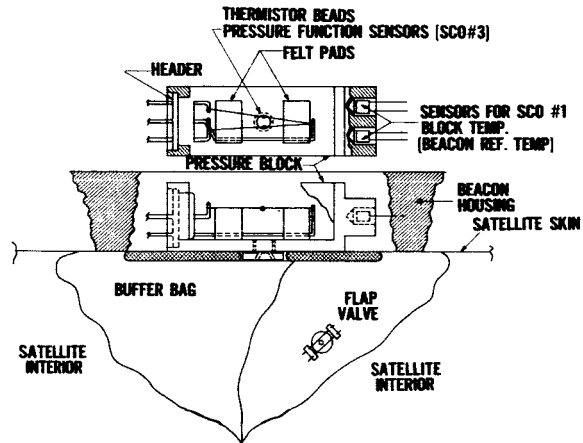


FIGURE 8

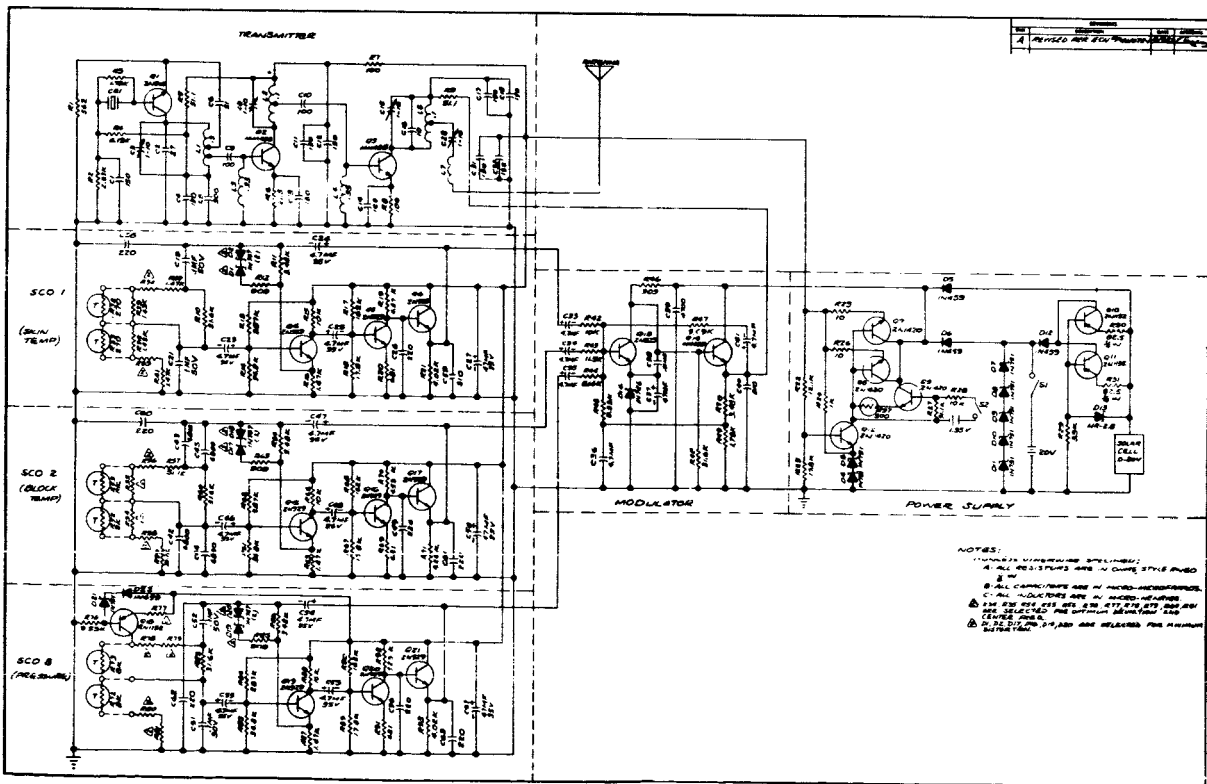


FIGURE 9.—Echo II beacon transmitter schematic.

The reference temperature was monitored by sensing the temperature of the cavity block.

Calibration of both the temperature and pressure sensors posed a problem in that relatively wide range of temperature and pressures had to be monitored. Fortunately, it was possible to select thermistors with proper temperature coefficients and by using appropriate trimming resistances, it was possible to fit the data coverage into the allotted telemetry channels. Figure 9 illustrates the entire schematic diagram of the beacon circuitry. Figure 10 shows a typical calibration curve for the skin temperature sensor. Note three curves almost parallel to each other, indicating the temperature sensitivity of the subcarrier. In this case, the dependence is very slight. One can also see the wide range of temperature squeezed into a narrow range of the subcarrier channel (IRIG 3).

Since the pressure sensor subcarrier was also temperature sensitive, only more so, the third

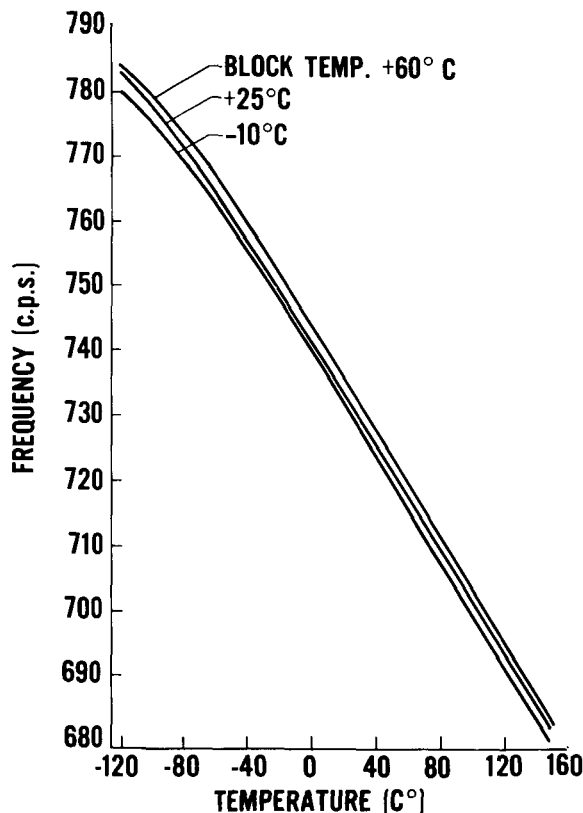


FIGURE 10.—Typical Echo A12 skin temp. sensor sub-carrier calibration.

subcarrier (IRIG 2) was used to monitor the temperature of the pressure sensor cavity block for calibration reference. Figure 11 illustrates a

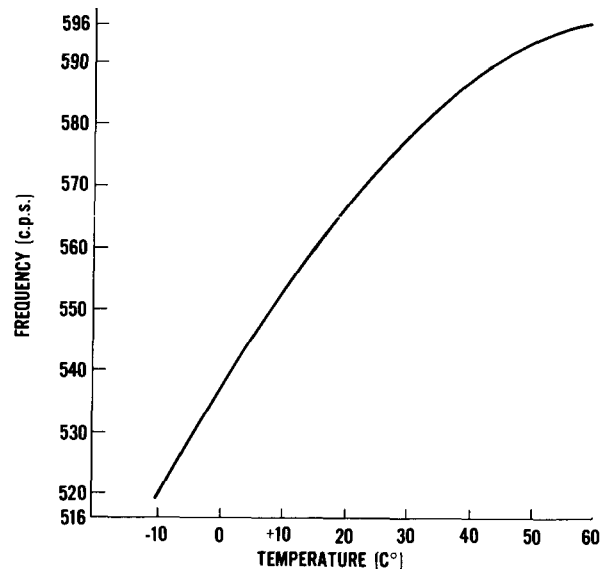


FIGURE 11.—Typical Echo A12 beacon temperature (pressure sensor block) sub-carrier calibration.

typical block temperature calibration curve. Since the entire beacon was potted with foam plastic and engineering tests had indicated that the temperature rise of the beacon under continuous operation was not greater than 25 degrees Centigrade, the calibration range of the reference temperature was limited to 70 degrees, from minus 10 to plus 60 degrees Centigrade covered within a subcarrier frequency range of 80 cycles. During the approximately six months that the beacons have been monitored to date, the reference block temperature has not been lower than minus 10 degrees nor higher than 30 degrees. The mean block temperature has ranged between 15 and 20 degrees. In this temperature region the slope is essentially linear which simplified the interpolation of the skin temperature and pressure. The extreme block temperatures were encountered only during the deepest eclipse and maximum sunlight periods.

The calibration of the pressure sensor posed other problems in addition to the wide range to be covered. First was the question of calibration with the actual gas to be used in the inflation system. Since there were several possible choices of

the inflation material, a scheme that would be compatible to all of the possible materials was required.

Figure 8 also shows the over pressure sensing scheme. It was decided to calibrate the sensors with air using a buffer chamber between the sensor and the main chamber . . . the balloon. The buffer bag became a part of the balloon, and consisted of a tetrahedron shaped polyethylene bag into which the pressure sensor "looked" through the cavity opening. A small one way pressure release valve was provided on the buffer chamber for pressure equalization as the balloon pressure increased during the inflation process; this valve was primarily for monitoring the decreasing balloon pressure.

Figure 12 is a typical pressure calibration chart. One can see the heavy dependence of the calibration curves on the block or the beacon reference temperature.

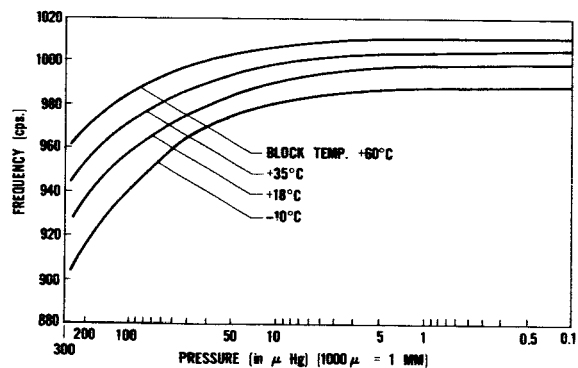


FIGURE 12.—Typical Echo A12 pressure function sub-carrier calibration.

Figure 13 shows how the pressure and temperature read during the initial inflation stage of the Echo satellite. The chart represents data for two orbits or about three and a half hours of flight after deployment. Dotted portions of the curves have been approximated and represent data missing owing to the separation distance between the tracking stations, which precluded a continuous orbital coverage.

The internal pressure has long since been dissipated. The temperatures of the skin and of the beacon, however, continue to be monitored and they reflect the mean values as the respective sensors rotate about a yet to be determined spin

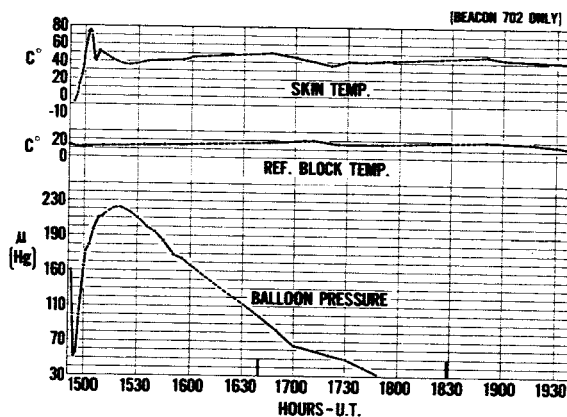


FIGURE 13.—Beacon telemetry data Echo II launch January 25, 1964.

axis of the satellite. Skin temperatures as high as 90 degrees Centigrade have been recorded at times. The lowest temperature recorded was minus 115 degrees just moments after the satellite emerged from the earth's shadow and the beacon turned on. Minus 115°C. is just about the lower limit of the subcarrier calibration. The average skin temperature of the Echo II during the past six months in orbit has been about 50 degrees Centigrade.

The telemetry beacons have provided certain other information not originally intended to be obtained. One of considerable interest in many quarters is the rotation speed of the satellite. Another is the location of the spin axis on the satellite and its attitude. One finds that the rotation period has been varying during the past six months; the variations have been very slight however. Figure 14 shows the rotation period in seconds plotted versus time in increments of 10

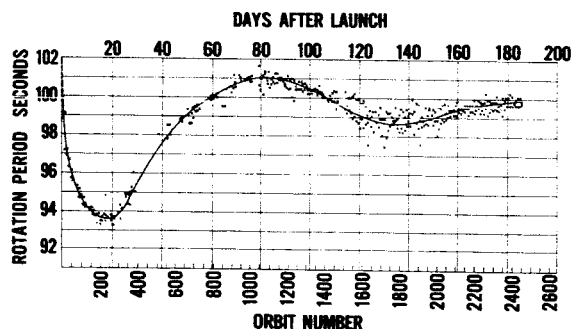


FIGURE 14.—Rotation period of Echo II vs Time.

orbits. The orbit numbers are listed below and the days since launch at the top. The smooth line is the approximate mean spin period for the ten-orbit intervals. Since the time span of the curve has been deliberately compressed in order to fit into the figure, it appears as if the spin period is varying considerably. Actually, if this curve is shown in its proper time perspective, the station to station variations are less than $\frac{1}{4}$ second in a given orbit and the orbit to orbit variations or changes are within one-half of a second at most.

Another set of information which will be consolidated shortly is that data from which the pointing direction of the satellite spin axis will be determined.

What is interesting is the source of information that led to the rotation period and spin axis data. This is the behavior of the beacons as the operating supply voltage varies. The supply voltage began to vary very soon after the orbit injection because the battery voltage dropped below the regulation level and its capacity dropped below the recovery point, with the result that the beacon operation became dependent on solar power. Interestingly enough, what caused these power supply changes is the rotation of the balloon itself.

The subcarrier stability was first affected by the power supply changes. For example, IRIG channels 2 and 3 (the temperature sensors) would shift two or three cycles when the supply voltage dropped below the regulation level of 19.2 volts. The frequency would then hold there until the supply voltage dropped to about 14 volts when both subcarriers quit operating. However, IRIG 4 (the pressure sensor) which contained the constant current heating device would continue to shift frequency steadily until the supply voltage dropped to 14 volts; then it would also quit. The carrier oscillator on the other hand would shift frequency with supply voltage also, but it would not quit oscillating until the voltage dropped to about 10 volts.

Figure 15 shows a typical carrier frequency shift as a function of beacon temperature and supply voltage. Interesting enough, in the flight units the carrier frequency remained within the receiver bandwidth and would continue to be received by the tracking stations but without any subcarrier signal during a part of a rotation period and during parts of an entire pass at times. Normally

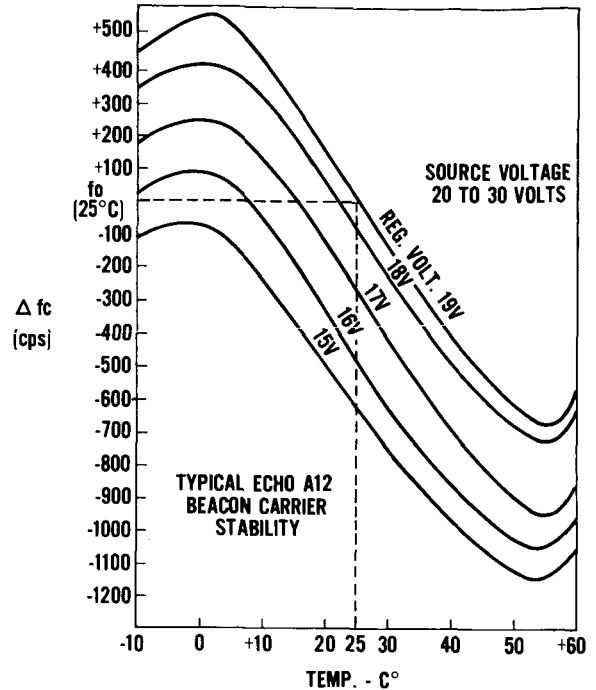


FIGURE 15.—Frequency deviation vs beacon temp. and supply voltage.

one would depend on the carrier level or the AGC level to determine the spin rate but in this instance it was more advantageous to use the subcarrier shifts since they were purely dependent on the solar cell voltage which in turn is a function of the incident illumination angle. But on the other hand, data have been obtained during some of the passes over certain stations in which the subcarrier from one beacon is present continuously for three to four rotations or about seven minutes, while the other beacon signal fades in and out with the rotation period. These items should be of interest to experimenters who are monitoring the Echo II telemetry beacons.

Figure 16 shows the carrier frequency versus time as the satellite passed over the Blossom Point, Maryland tracking station at 1400 hours UT on April 17. The smooth curve is the predicted doppler shift for this pass. It is based on a nominal carrier frequency of 136.170 Mcs. The other, rollercoaster type curve is the actual frequency shift detected as the satellite passed over the station. One can see several interesting events taking place. First, there is the relative stability

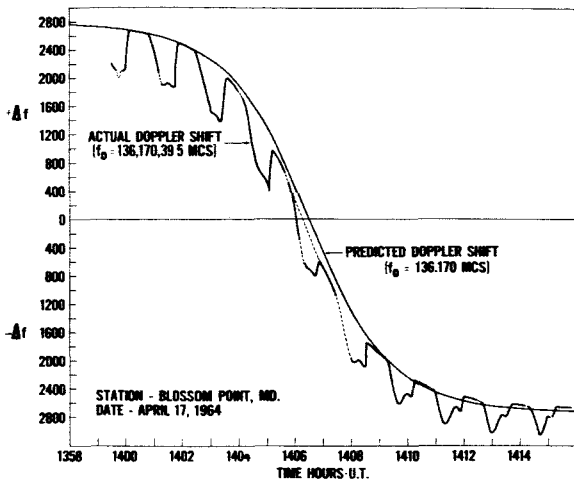


FIGURE 16.—Typical Doppler shift curve of Echo II telemetry beacon (#701).

of the carrier during portions of the pass. Next there is the repetition of the ups and downs, which in this case occurs at 100 second intervals. Finally, the signal dropouts as the antenna pattern passed out of range or the solar power dropped to point such that the carrier quit. The beacon stability with adequate operating supply voltage is attested to by the near superposition of the two curves. The actual shift is based on the carrier frequency of 136.170,395 Mc/s, which was the last check made on this beacon before flight. The

dotted portions of the curves indicate the missing signals and have been approximated.

The subcarrier data on strip charts recorded at five millimeters per second speed show these details much more clearly. However, they are not suitable for illustrations. The details will not be noticeable when reduced in size. It is regretted that they cannot be shown.

This concludes the report on the Echo II beacon telemetry instrumentation. In conclusion, the Project wishes to extend its acknowledgment to the numerous people who have participated in the Echo II telemetry program for their excellent effort and cooperation. Their participation has ensured a highly successful data source aboard the Echo II satellite.

REFERENCES

1. NASA TN D-504, "A Thermistor Pressure Gage," by J. AINSWORTH and A. P. FLANICK (GSFC), Nov., 1960.
2. NASA, GSFC Publication A12-0-102, Echo A12 Environmental Qualification Test Specification for Beacon-Telemetry Subsystem, dated Dec. 19, 1961 with revisions through June 28, 1962.
3. NASA, GSFC Publication A12-0-202, Echo A12 Environmental Acceptance Test Specification for Beacon-Telemetry Subsystem, dated Dec. 19, 1961 with revisions through June 28, 1962.
4. NASA, GSFC Publication A12-0-362, Performance Specification for Beacon-Telemetry Subsystem-Echo A12, dated Dec., 1961, with revisions through June 28, 1962.

EXPERIMENTS TO DETERMINE COMMUNICATION CAPABILITY OF THE ECHO II SATELLITE*

W. C. NYBERG, R. L. KAISER† AND W. E. LEAVITT‡

This paper describes some of the experiments being conducted on the Echo II satellite, which was launched by GSFC on January 25, 1964. The objective of these experiments, being performed under the direction of Goddard Space Flight Center, is to determine the communication capability of the Echo II satellite, and also to provide information about the shape and surface characteristics of the satellite as a function of time.

EXPERIMENT PARTICIPANTS

The stations that have participated in the experiments described in this paper are at Collins Radio Company in Richardson, Texas, a suburb of Dallas, and the Naval Research Laboratory at Stumpneck, Maryland, 30 miles south of Washington, D. C., Ohio State University at Columbus, has also begun participating in the Goddard experiments and it is anticipated that the Navy Electronics Lab at San Diego will begin to participate soon.

SYSTEM CONFIGURATION

The basic circuit configuration employed in the experiments is shown in Figure 1. The Collins

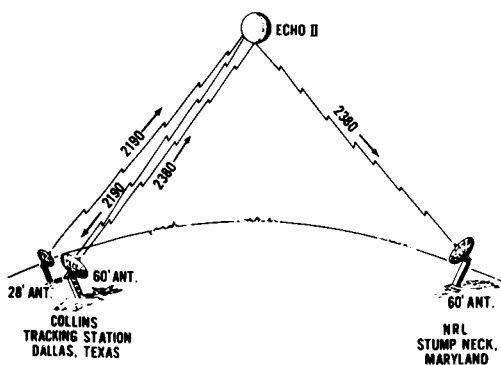


FIGURE 1.—Circuit configuration, Echo II communication experiments, Dallas-Stump Neck.

*Published as *Goddard Space Flight Center Document X-621-64-291*, October 1964.

†Collins Radio Co., Richardson, Texas.

‡Naval Research Laboratory, Washington, D.C.

facility utilizes a solid-surface, 60-foot parabolic antenna and a 10-kw transmitter to illuminate the Echo II satellite at 2380 Mc. The signal reflected from the satellite is received with a similar 60-foot parabolic antenna at the NRL facility. The signals are circularly polarized. To provide accurate pointing of the 60-foot antenna at Dallas, an automatic tracking radar system, operating at 2190 Mc, is employed. This consists of a separate 28-foot dish antenna with a 10-kw transmitter for illumination at 2190 Mc, together with the 60-foot antenna, which is equipped with an amplitude monopulse feed and phase-locked tracking receiver to receive the radar signals.

Figure 2 is a picture of Collins' 60-foot antenna. The transmitter and receiver front end are mounted on the antenna—just behind the reflector. Acquisition of the satellite is achieved by means of Goddard—furnished predictions combined with an automatic acquisition system.

The NRL 60-foot antenna is normally controlled by a digital drive tape through an antenna programmer. An optical tracker is used to correct errors in the drive tape when the satellite is optically visible. Because of slight inaccuracies in the predictions, most of the experiments have been conducted at night so that signal fluctuations caused by mispointing of the antenna can be eliminated through the use of the optical tracker.

Figure 3 is a picture of NRL's antenna. The optical tracker cabin is the large box seen behind the reflector. The NRL receiver system utilizes a low noise TWT and a phase-locked i-f loop at



FIGURE 2

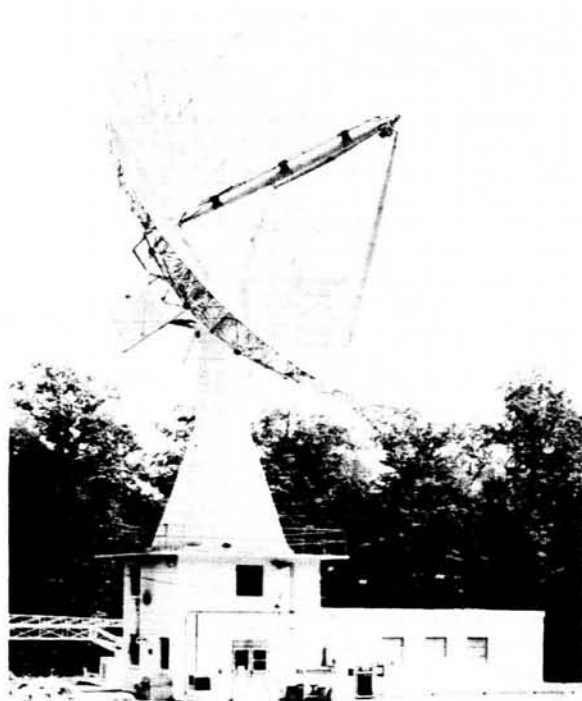


FIGURE 3

60 Mc to provide high sensitivity. Modified R-390 A/URR communication receivers serve as the primary envelope demodulators for the experiments.

Both the NRL and Collins' facilities are equipped with a variety of strip-charts and magnetic tape recorders for recording the experimental data. Both facilities are also equipped with bore-sight cameras so that the pointing errors can be accurately determined.

EXPERIMENTS

The experiments are divided into two categories: the first, to determine directly and quantitatively, the characteristics of the Echo II satellite as a radio frequency reflector, and the second, to verify by demonstration, some of the communication capabilities of the Echo II satellite. The tests in the first category include measurement of the received signal level. This test provides data for determining the effective cross-sectional area of the satellite and signal fading and scintillation characteristics.

Figure 4 is a typical record of the received signal strength versus time as received at the NRL facility. As you can see, the peak signal level is about -100 dbm. Average fluctuations in the signal caused by the satellite are on the order of

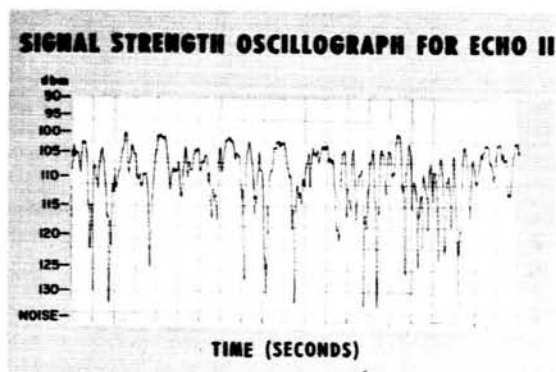


FIGURE 4.—Signal strength oscillograph for Echo II.

± 5 db; short deep fades of up to 20 db occur frequently. The scintillations or (amplitude-variations) are due to the irregularities of the satellite's surface. To date the various experimental measurements have not indicated any change in the satellite's structural characteristics between the time when it was first observed on the sixth revolution and the present time. This is in contrast to Echo I which showed significant changes over the same time period. Correlation

has not been observed during these experiments between the reflected r-f signal fluctuations and the rotation of the satellite as determined from the satellite's beacon telemetry signals.

The bandwidth capability of the satellite is indicated by determining the degree of correlation of the received signal level of two carriers separated in frequency. Each carrier is received separately with narrow band receivers, and the amplitude variations of the receiver outputs are then compared to determine the degree of correlation. Up to the present time, the degree of correlation between the fluctuations in the amplitude of the carriers, separated by 12 Mcs, has been very good, (close to 100 percent) as is shown in Figure 5. The two lower traces are the two received signals. The upper trace is the difference of the two signals and illustrates the high degree of correlation. This experiment has been conducted at spacings of 70 and 190 Mcs, but the acquired data has not yet been analyzed. Quick look analysis, however, indicates as expected, a lower degree of correlation at 70 Mcs, and still lower degree of correlation at 190 Mcs than at the 12 Mcs spacing.

Demonstrational type experiments include an audio-frequency transmission test utilizing tones, voice and music. Facsimile and digital data transmissions also are included in the demonstrational experiments.

Figure 6 shows examples of facsimile transmission that has been obtained over the Echo II satellite circuit. The facsimile signals were pre-recorded on magnetic tape and speeded up by a factor of 4 during transmission to produce an increased information bandwidth and shorten the



FIGURE 6

length of time of picture transmission. The picture on the right was made directly from the master tape. The other on the left was reproduced from a magnetic tape recording of the received signal over the satellite circuit. The vertical streaks are caused by the deep fades in the signal. Voice and music transmissions have been made using FM with modulation frequencies of 30 to 15,000 cycles and with a deviation of ± 15 kes. Except

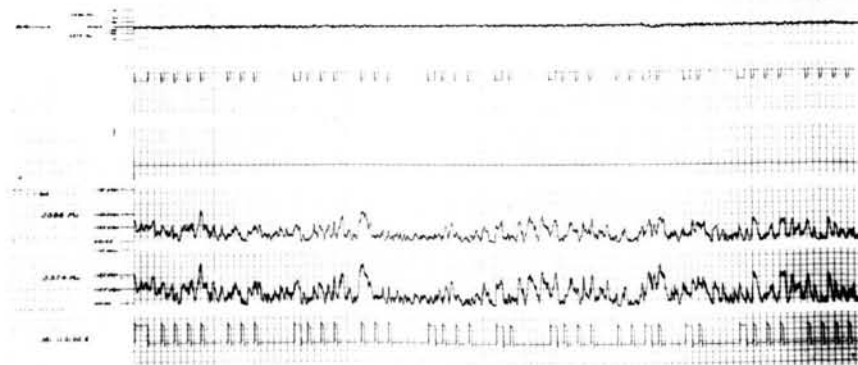


FIGURE 5

for an occasional hiss of noise caused by fades, the received voice and music is a duplicate of that on the original tape.

It must be pointed out that large quantities of data are being acquired from these experiments, a good portion of which will be reduced and analyzed in detail. The results will be made available after the conclusion of the experiment program.

The conclusions reached from the experiments so far are as follows:

1. These experiments indicate no change in the reflectivity characteristics of the satellite since the initial experiment conducted on the sixth orbit. This verifies that the

satellite has met the design objective of becoming a rigidized structure.

2. Echo II has a bandwidth capability of at least 12 Mcs.
3. Echo II has good voice, music, and facsimile transmission capability.
4. The transmission characteristics of the Echo II circuit of over 1,000 miles, compare favorably with those of tropospheric scatter circuits and micro-wave point-to-point circuits of considerably less distance.
5. Tracking and station operation quickly become routine with no major problems involved.

SLOPE DETECTION AS A METHOD OF DETERMINING THE PEAK POWER POINT OF SOLAR ARRAYS*

JOHN PAULKOVICH

Long term satellite missions have been primarily dependent on solar cells as a source of electrical power. The power generating capability of the solar cells used in the power supply are sensitive to hard particle irradiation, temperature and illumination. Typical voltage-current characteristics under differing conditions of temperature and irradiation are shown in figure 1. Figures 14 and 15 depict the effect of illumination on cell

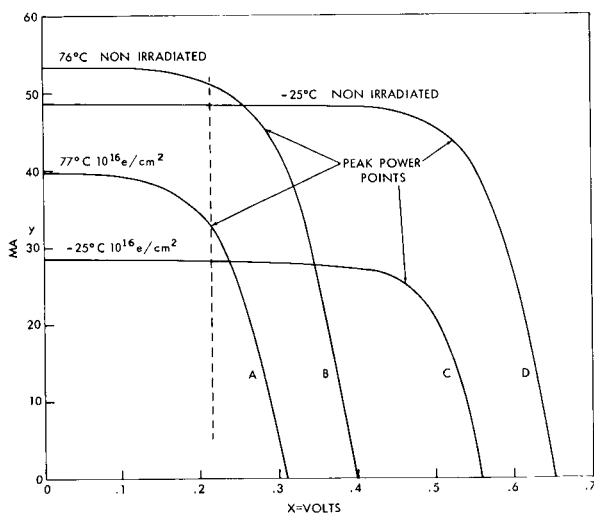


FIGURE 1.—Typical N/P solar cell I-E curves for various environmental conditions.

output. The maximum power generation capability for each cell is defined by the largest rectangle which can be inscribed under the I-E curve. It may be seen from figure 1 that:

1. The absolute values of maximum power varies quite drastically under differing conditions of temperature and irradiation.
2. The occurrence of maximum power varies quite drastically with voltage and current.

*Published as *Goddard Space Flight Center Document X-636-64-282*, October 1964.

At present the power system designer attempts to anticipate the operating condition in space and then chooses an optimum operational voltage or current at which to operate. Once chosen this value usually remains fixed throughout spacecraft life. This approach has the following disadvantages:

1. In a radiation environment initial power is sacrificed so that at the end of mission operation will be at maximum power.
2. Variations in environment from what is expected can be disastrous for the power supply.
3. Interplanetary space probes would be forced to sacrifice considerable power initially in order to compensate for the difference in solar intensity and temperature at some later time in the flight.

The solution to these problems is an electronic sensing system that will allow continuous operation at maximum power no matter what the operating environment. This report contains the description of such a system along with the results of initial tests on a breadboard model.

To determine the feasibility of this method a graphical analysis was made of solar cells under the following conditions:

1. Non irradiated at -25°C
2. Non irradiated at 76°C
3. Irradiated (10^{16} e/cm²) 77°C
4. Irradiated (10^{16} e/cm²) -25°C
5. Solar angles of 0, 40°, and 60°.

Figure 1 illustrates the possible variations in the E-I curves due to temperatures and space radiation environment. From figure 1 it becomes apparent that it is impossible to select any one voltage, or current for that matter, that will be at the peak power point under all conditions.

The problems involved in obtaining maximum power from the solar cells lie in the present method of charging the batteries. The most common method is to place the solar cells, battery pack and shunt regulator in parallel. The shunt regulator clamps the solar cell and battery voltages when they reach a predetermined level. Therefore, the solar array voltage is maintained at some predetermined maximum voltage.

Again referring to figure 1, it can be seen that if a certain maximum voltage is selected which will operate under all conditions, it would probably be in the vicinity of .2 volts per cell. If operation is selected for peak power from curve "A" of figure 1, then if we refer to figure 5, we can select the optimum voltage. Figure 5 is a reproduction of curve "A", figure 1, with a superimposed power curve. This power curve ("B" of figure 5) indicates maximum power at approximately .215 volt. If the load is operating at .215 volt, then optimum power will be delivered to the load. But now, if we refer back to curve

"C" of figure 1, we will find that at .215 volt less current is available and obviously less power. Figure 4 is a reproduction of 1-C along with the power curves. At .215 volt approximately 5.9 milliwatts of power would be delivered to the load, while actually 11.5 milliwatts are available.

Curves 1-B and 1-C are reproduced in figures 3 and 4 and the same conditions prevail. From the foregoing it becomes apparent that a method of determining the peak power point and operating in this region is of the utmost importance for utilizing maximum available energy.

The method to be discussed is based on operation at a fixed "slope" of the E-1 characteristic curve. Figures 2, 3, 4, and 5 are reproductions of each of the individual curves of figure 1 with a superimposed power curve. Also indicated are the peak power point, the 95 percent power points, and the tangent to the curve at these points. 95 percent was arbitrarily selected as the lowest desirable operating point. The tangent has the slope x/y , where volts is represented by x

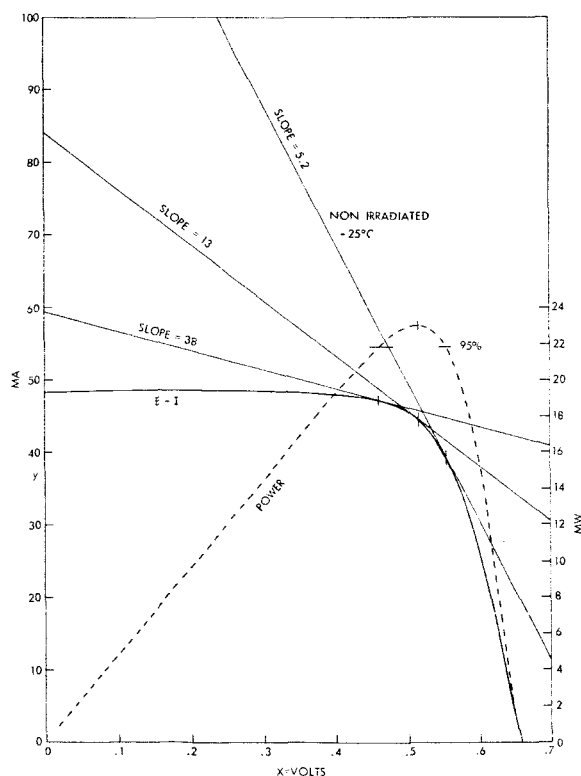


FIGURE 2.—Graphical determination of optimum slope for maximum power coupling to detector for a non-irradiated solar cell at -25°C .

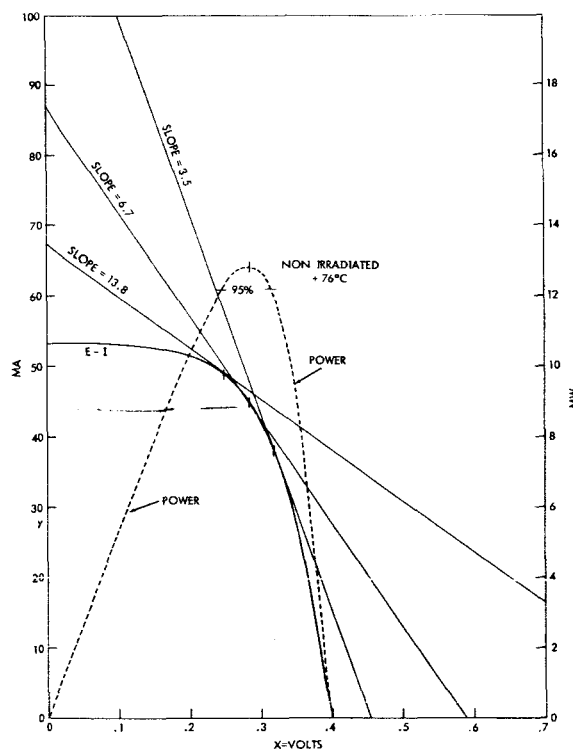


FIGURE 3.—Graphical determination of optimum slope for maximum power coupling to detector for a non-irradiated solar cell at 76°C .

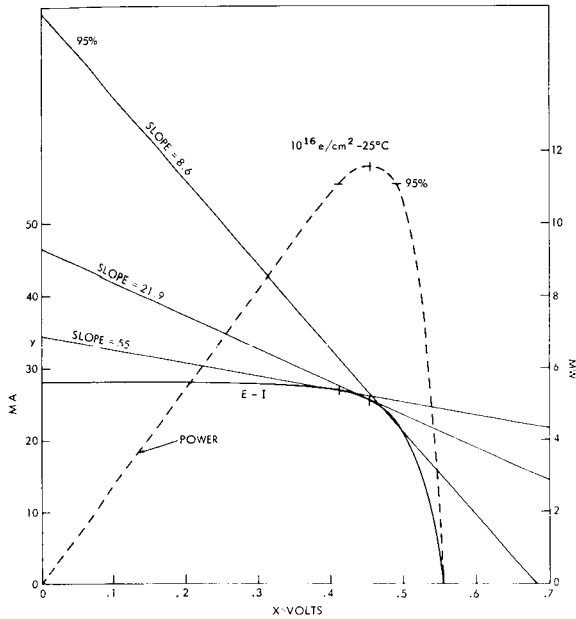


FIGURE 4.—Graphical determination of optimum slope for maximum power coupling to detector for an irradiated solar cell at -25°C .

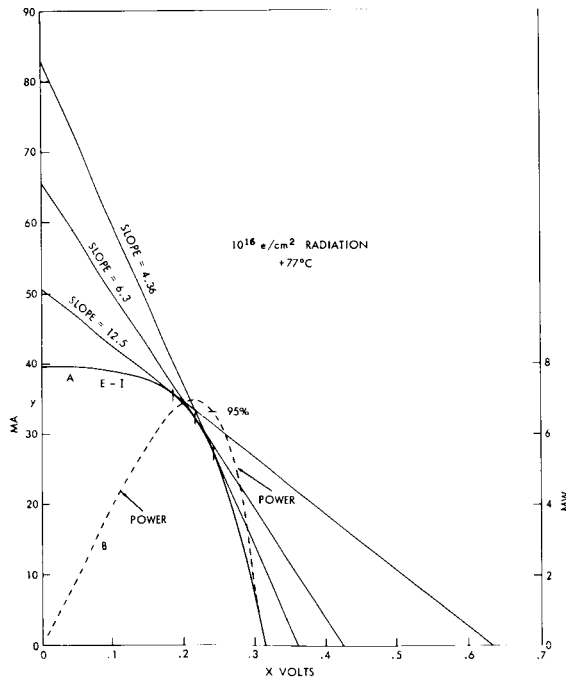


FIGURE 5.—Graphical determination of optimum slope for maximum power coupling to detector for an irradiated solar cell at 77°C .

and amperes is represented by y . In analyzing figures 2 through 5 we can see that there is a range of slopes which is within 95 percent of the peak power point under all conditions. But there is also the solar angle to consider. Figure 6

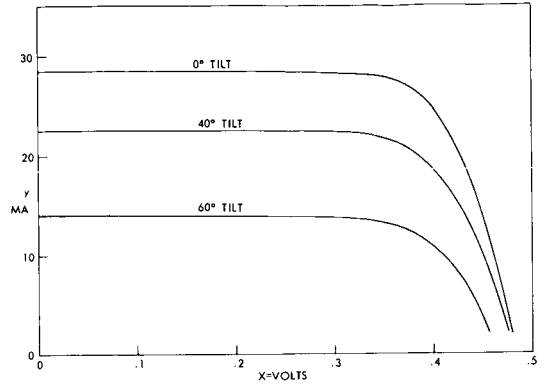


FIGURE 6.—Typical solar cell E-I curves vs angle of incidence.

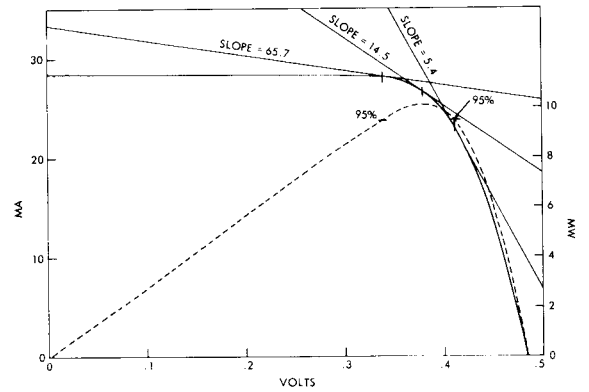


FIGURE 7.—Typical solar cell E-I curve at normal incidence.

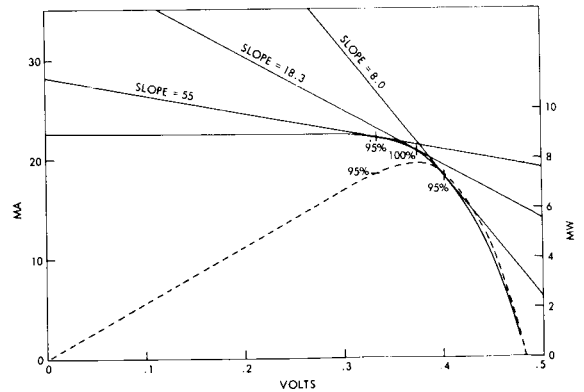


FIGURE 8.—Typical solar cell E-I curve at 40° tilt from normal incidence.

indicates the effects of solar angle from 0° to 60°. Note the change in scales. The original plot was 10 cells in series, to maintain uniformity, the x scale numbers have been altered by 1/10. Figures 7, 8, and 9 are reproductions of figure 6 along with

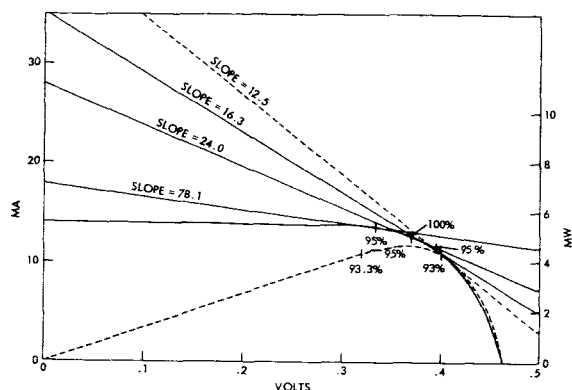


FIGURE 9.—Typical solar cell EI curve at 60° tilt from normal incidence.

the slopes and also a super-imposed power curve. Figure 10 illustrates the range of slopes in each of the preceding graphs. From figure 10, it becomes apparent that if the slopes were held

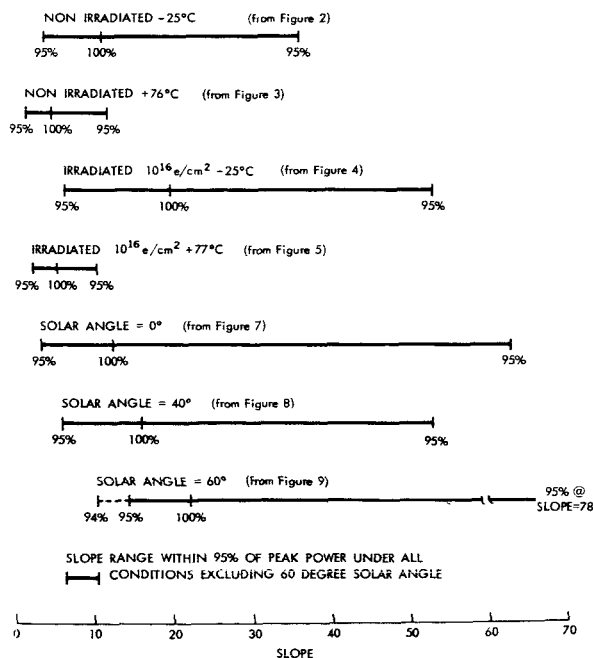


FIGURE 10.—Compiled slope variation of Figures 2 through 9.

between 8.5 and 12.5, operation would be 95 percent of the peak power point or better under all conditions except the 60° solar angle. A slope of 12.5 on the 60° solar angle curve, figure 9, represents approximately 93 percent of the available peak power. As the solar angle exceeds 60°, the deviation from peak power will increase. The solar angle then sets the limits of operation and should not exceed 60° for this set of conditions.

Since we are interested in the slope, then a method of scanning this slope is necessary. Figure 11 illustrates a block diagram of a slope detector.

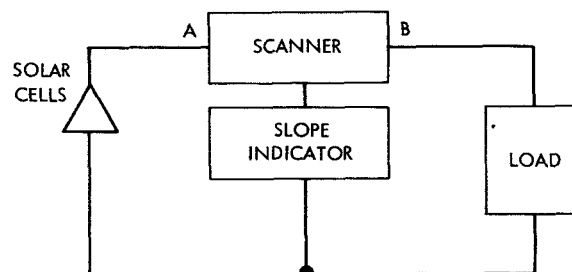


FIGURE 11.—Block diagram of slope detector.

The scanner applies an AC voltage signal across points A and B of a constant magnitude. This scanner must be insensitive to DC load currents and sensitive only to the slope x/y . A change of slope causes a change in reflected current to the slope indicator.

An operational circuit that will perform this function is illustrated in figure 12.

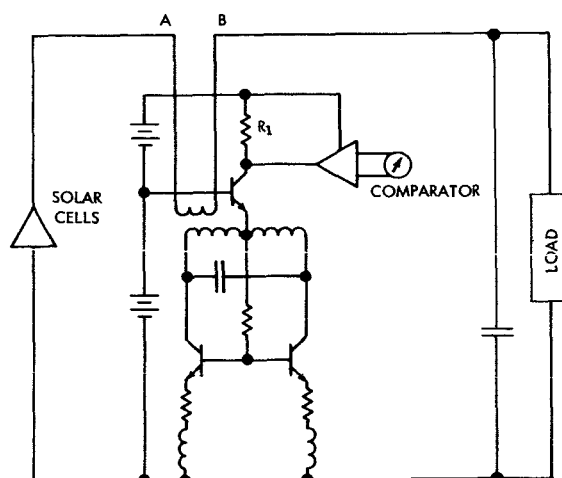


FIGURE 12.—Basic diagram of slope detector.

The voltage across $A-B$ will be constant and represent x . The AC loading effect will be inversely proportional to the slope x/y and reflect a current through the transformer and appear as a voltage across R_1 . This voltage is then compared to a reference voltage which represents the desired slope. The differential amplifier will balance at the desired slope and the meter will be either positive or negative above or below the desired slope. This voltage difference may be utilized to control the load or charging circuit.

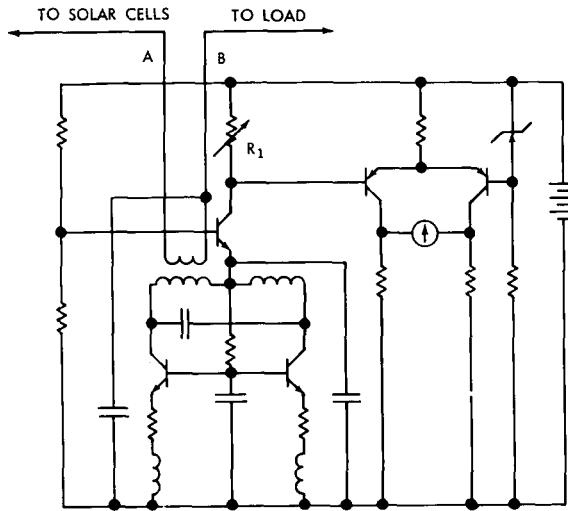


FIGURE 13.—Schematic of the breadboard model of the slope detector.

A breadboard model was constructed (refer to figure 13) and tested. The differential amplifier or comparator was connected to a 100-0-100 microampere meter. R_1 was made adjustable for various slope selections. $E-I$ plots using an x/y plotter and S51 solar paddles are shown in figures 14 through 16. The plots represent full light intensity, partially shaded solar cells, full paddle, $\frac{1}{2}$ paddle and $\frac{1}{4}$ paddle, as indicated.

In figure 14, $E-I$ traces of full paddle, $\frac{1}{2}$ paddle, and a shaded $\frac{1}{2}$ paddle are illustrated. The slope detector was set to slopes of 30, 70, 100, and 150. These various points are also indicated in figure 14. A slope of 70 was selected as being the nearest to the peak power point although not necessarily optimum. The slope detector was then set to a slope of 70 for figures 15 and 16. The arrows indicate the point selected by the

slope detector (meter balanced). The lines drawn tangent to the $E-I$ plots are for a calculated slope of 70. The calculated slope and the slope selected by the slope detector coincide very nicely. The curves of figure 15 may be extreme conditions since they represent a power excursion of almost 10 to 1. Although the solar paddles are likely to have shadowing conditions, it is very unlikely that they would be shadowed by 90 percent. Under lesser shadowing conditions the slope could be

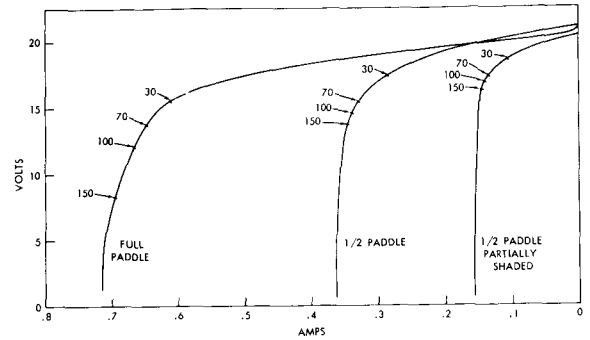


FIGURE 14.—Solar paddle $E-I$ curves, slope detector set to 30, 70, 100, and 150.

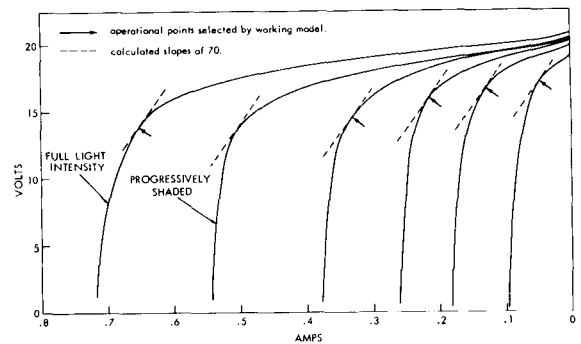


FIGURE 15.—Solar paddle $E-I$ curves, slope detector set to 70.

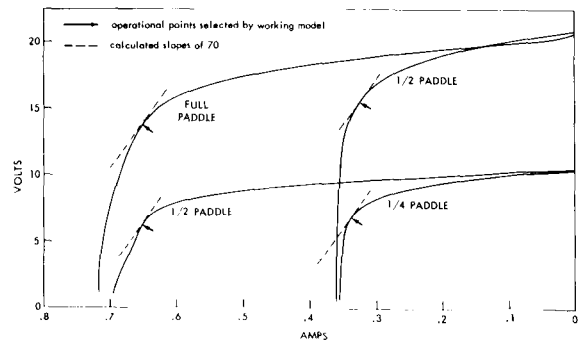


FIGURE 16.—Solar paddle $E-I$ curves, slope detector set to 70.

optimized for operation nearer to the peak power point.

Figure 17 illustrates the slope detector combined with a pulse width modulated regulator. This method of power detection combined with a pulse width modulated regulator would certainly increase the charging power to the satellite bat-

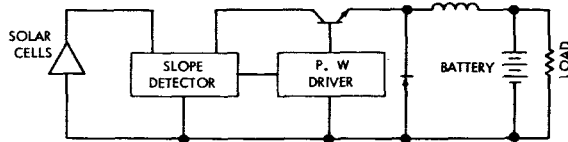


FIGURE 17.—A block diagram of the slope detector controlling a pulse width modulated regulator for maintaining optimum matching between the load and the solar paddles.

tery pack. The pulse width regulator would act as a matching transformer between the solar cells and the battery pack, while the slope detector would be the impedance sensing device and control circuit for maintaining this optimum match.

ACKNOWLEDGMENTS

Brian T. Cunningham and Luther W. Slifer, for the use of the solar laboratory facilities that these experiments could be conducted.

REFERENCE

BRIAN T. CUNNINGHAM, ROBERT L. SHARP, and LUTHER W. SLIFER, Jr., The Electrical Characteristics of Irradiated Silicon Solar Cells as a Function of Temperature, NASA-Goddard Space Flight Center, Greenbelt, Maryland, March 30, 1964.

PULSE-FREQUENCY-MODULATION TELEMETRY*

ROBERT W. ROCHELLE

This report is concerned with the heuristic development of the basic features of pulse-frequency modulation, an information encoding technique which has been used in a number of spacecraft. The primary advantages are its noise-immunity characteristics and its ease of generation. A description of the present method of forming and synchronization is presented in order to illustrate the convenient handling of both analog and digital data.

The theory of group-coded binary sequences is derived by using a special type of correlation matrix. By combining particular sequences from a number of these correlation matrices, a new matrix is generated which is identical to the correlation matrix of a set of pulse-frequency modulation words. Thus, it is shown that pulse-frequency modulation has the same communication efficiency as a comparable set of coded binary sequences with an equal number of quantized levels.

In the detection process for PFM signals, a set of contiguous unmatched filters is used to enhance the signal-to-noise ratio. The effects of Rayleigh noise on the output of these filters, the word-error probability is derived as a function of the energy per bit, noise power density, and degree of coding. The same development is given for the matched-filter case.

The analysis of the excitation of an unmatched filter to a step sinusoid is carried out to indicate the magnitude of the errors affecting the direct frequency measurement of the pulsed sine wave. The analysis also covers the action of a pulsed sine wave in which the frequency is changing as it passes through the filter.

Experimental results from the spectral analyses of satellite recordings perturbed with random noise are presented to illustrate the noise-immunity characteristic of pulse-frequency modulation.

INTRODUCTION

The problem of communicating over large distances with minimum power is particularly challenging in view of the advances in information theory during the last 2 decades. With this theory as a guide, a number of interesting types of communication systems can be postulated which, in the limit, approach the theoretical maximum communication efficiency. The encoding methods of some of these systems are exceedingly complex in their implementation. The ideal type of system retains a high communication efficiency with little loss in simplicity. Pulse-frequency modulation is an attempt to fulfill these two basic criteria.

Pulse-frequency modulation (PFM) has been successfully employed as the information encoding technique in a number of small (<200 lb) scientific satellites and space probes where the reduction of spacecraft power and weight was a prime consideration. The use of such systems leaves a

greater percentage of the spacecraft available for scientific instrumentation. Since little information has appeared in the literature on the subject, it is hoped that this work will provide a background for practical usage, as well as a theoretical derivation, of pulse-frequency modulation.

The Encoding Problem

Telemetry implies measurement at a distance. The concept of modern telemetry includes the gathering and encoding of information at a remote station, transmitting this encoded signal to the receiving station, decoding the signal at this station, and presenting the measurements in an acceptable form. The transmission path may involve propagation by radio, light, or even sound waves. Attenuation of the transmitted signal is usually a predominant factor because of the physical separation between the transmitting and receiving apparatus. If attenuation is low, the choices of the encoding and modulation methods are not critical, and methods which lead to simple and reliable systems are usually employed. With the introduction of attenuation in

*Published as *NASA Technical Report R-189*, January 1964. This work was also submitted to the University of Maryland as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

the transmission path, interference in various forms may perturb the signals and cause transmission errors. By proper encoding of the information, the effect of these noise perturbations can be minimized. The amount of improvement or immunization to noise that can be accomplished is governed by Shannon's channel capacity theorem (Reference 1):

$$C = B \log_2 \frac{P + N}{N} \quad (1)$$

where C is the channel capacity (bits per second), B is the channel bandwidth, P is the signal power, and N is the average noise power. As the signal power is decreased, the bandwidth of the encoded signal must be increased to maintain the same channel capacity and probability of error.

Various schemes can be devised to take advantage of Shannon's channel capacity theorem and thereby reduce the transmitter power for a given information rate. Kotel'nikov has described the general theory of high efficiency encoding (Reference 2). The problem involved is to approach the Shannon channel-capacity limit as closely as possible. But this must be consistent with the ability to implement the encoding scheme with equipment which is not of undue complexity. A practical example of this problem is the telemetering of scientific data from satellites and space probes. Since spacecraft power is limited, savings in transmitter power due to efficient coding can extend either the range or information rate or both. The improvement is greater in the small satellite class (<200 lb), where a larger percentage of the satellite weight is devoted to the transmitter power system.

History of Coded Telemetry

Coded telemetry began with the early work on frequency modulation. At that time frequency modulation was thought to require less bandwidth than amplitude modulation, since the frequency deviation could be made smaller than the modulating frequency. Carson disproved this theory in 1922 (Reference 3). In 1936 Armstrong demonstrated the major advantage of frequency modulation—the improvement of the signal-to-noise-ratio for large frequency deviations (Reference 4). The system traded bandwidth for signal-to-noise-ratio

improvement. A recognized deficiency was the sharp threshold or decrease in the signal-to-noise ratio at low signal levels.

Pulse-code modulation* developed much more slowly than frequency modulation (Reference 5). It was not until 1948 that Shannon predicted that pulse-code modulation would have error-reducing properties (Reference 1). In 1950 Hamming devised a practical scheme of coding to effect these error-reducing properties (Reference 6). He introduced the error-detecting parity bit for a binary sequence and added bits to these for error-correcting purposes. Since 1950, numerous papers have been written on the subject of coding theory. Viterbi utilized these theories in calculating error probability as a function of the signal energy per bit for various degrees of coding (Reference 7).

With the entry of the United States into space exploration, the need for a lightweight low-power telemetry system developed. Such a system (Reference 8), utilizing the noise-reducing properties of frequency modulation and the error-reducing properties of pulse-code modulation, was incorporated in the first Vanguard scientific space vehicle and ultimately achieved orbit in Vanguard III (1959 η 1). The telemetry encoder had a 48 channel capability, weighed 6 ounces unpotted, and required 12 milliwatts of power. Pulse-frequency modulation, as the system was called, was used again on the Ionosphere Direct Measurements Satellite, Explorer VIII (1960 ξ 1). In April 1961 the space probe Explorer X (1961 κ 1), which measured the interplanetary magnetic field, was sent to an altitude of 240,000 km and again pulse-frequency modulation provided the encoding technique. Since that time, four more scientific satellites have been orbited with pulse-frequency-modulation telemeters. Explorer XII (1961 ν 1) provided continuous measurements of the energetic particles in the Van Allen radiation belts out to 80,000 km; Ariel I (1962 σ 1) was an ionospheric satellite (a joint effort between the United States and the United Kingdom); Explorer XIV (1962 $\beta\gamma$ 1) was a follow-on to Explorer XII; and Explorer XV (1962 $\beta\lambda$ 1) is providing a study of the artificial radiation belt. Two satellites to be launched in the near future, the

*P. M. Rainey, U. S. patent 1,608,527, November 30, 1926, issued to Western Electric Co., Inc.

Interplanetary Monitoring Probe and the successor to Ariel I, will use pulse-frequency modulation as the encoding technique.

CHARACTERISTICS OF PULSE-FREQUENCY MODULATION

General Description

Consider a time function $f_1(t)$ which is band-limited between zero and $1/2T_0$ cps. The function may be completely described by a series of impulses of separation T_0 and area $f_1(nT_0)$, with $n = -\infty, \dots, -2, -1, 0, 1, 2, \dots, +\infty$. A sequence of these time samples may be encoded for transmission over a telemetry link in several ways. The method considered here is to encode the magnitude of the area of each impulse as the frequency of a pulsed subcarrier, the duration of the pulse being some fraction of the sample time T_0 .

A set of k analog time functions $f_m(t)$ may be multiplexed by sequentially sampling each function with spacings between samples of T_0/k . In order that the spectrum may be utilized more efficiently, each pulse length should be equal to the sampling time, T_0/k . In practice the pulse length has been set at half this value to allow for response-time limitations in the crystal filters used in the detection process. Thus, a train of sequential pulses, the frequency of each being proportional to the amplitude of a sample, comprises the basic configuration for pulse-frequency modulation.

The signals $f_m(t)$ may not necessarily be analog in nature. A sizable portion of the outputs of experiments aboard spacecraft occur as digital signals, and from a signal-to-noise-ratio viewpoint it is desirable to retain the digital character of the signal. The binary digits of the signal are combined and presented as a single digit of a higher order base. The present state of the art of PFM telemetry employs the encoding of three bits as one digit to the base eight; that is, three binary digits are encoded as one of eight frequencies. A special digital pulsed-subcarrier oscillator has been developed for this purpose; it is restricted to oscillation at only one of the eight possible frequencies, the frequency depending on the value of its three-bit input. In this manner the complexity of the switching circuitry in the encoder is

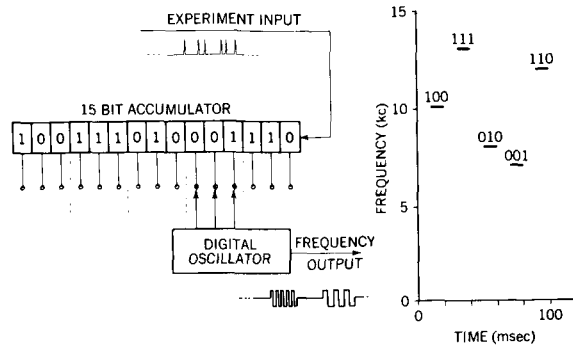


FIGURE 1.—Digital data readout system.

materially reduced, since readout of the accumulators and scalars is accomplished three bits at a time instead of serially. Figure 1 illustrates the manner of commutating the digital oscillator to scan the stored data in the accumulator. Pulses from the experiment are counted and stored as a binary number in the accumulator. The digital oscillator is commutated three bits at a time in five steps through the stored fifteen bits in the accumulator. The output of the oscillator, for this case (octal number 47216), is represented, as in Figure 1, by five serial pulses of discrete frequencies.

Since both analog and digital signals are encoded as frequencies of pulses, they may be intermixed in any order in multiplexing. Any one channel may be subcommutated to extend the number of signals $f_m(t)$ which may be encoded, with a corresponding reduction in the signal bandwidth.

The present format for PFM specifies that sixteen sequential channels comprise a telemetry frame. The first channel is devoted to synchronization, and the remaining fifteen are distributed between the analog and digital data to be telemetered. A group of sixteen sequential frames forms a telemetry sequence of 256 pulses (sixteen of these are devoted to synchronization). Figure 2 shows the frame and sequence formats.

Synchronization is assured in two ways. First, the energy in the sync pulse is increased by increasing its length 50 percent. Second, a unique frequency is utilized outside the data band for the sync pulse. To provide a means for identifying the subcommutated data, the frequency of the sync pulse in every other telemetry frame is stepped in sequence through eight frequencies. These frequencies lie in the data band and are the

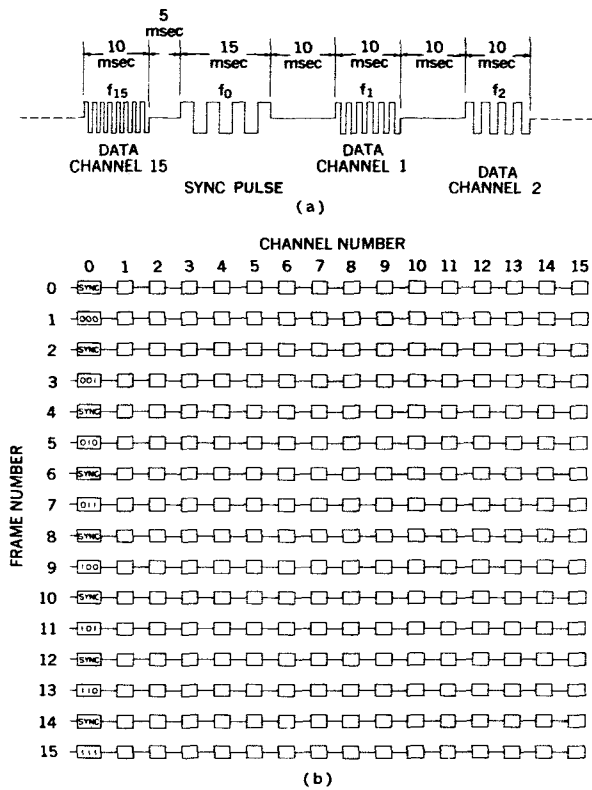


FIGURE 2.—Pulse-frequency-modulation format: (a) telemetry frame; (b) sequence.

same as the eight frequencies of the digital oscillator. When there is a poor signal-to-noise ratio, the energy at the sync frequency may be stored over a number of frames, and thus the sync signal-to-noise ratio is improved. This, of course, increases the acquisition time but it is a necessary compromise in regions of poor signal-to-noise ratios.

Design Considerations

In the design of any telemetry system, such features as bit rate, accuracy, precision, error rate, and bandwidth must be firmly specified. Proper choices for such factors are products of the application of information theory to the problem.

Let us investigate a few factors which are necessary in a PFM telemetry system. Let the maximum value of the analog time function $f_1(t)$ be defined as the full-scale value F_1 of the channel. If F_1 is divided into N equal parts, the magnitude F_1/N represents the precision of the system. This is the smallest discernible change between two samples of $f_1(t)$. The value of N is

determined by the requirements of the experiment being telemetered. For small scientific satellites a precision factor $1/N$ of 0.01, or 1 percent of the full-scale value, is normally considered adequate. For experiments, such as the flux-gate magnetometer, which depend on the measuring of the modulation of the magnetic field by the satellite's spin, precisions to 0.1 percent are desirable, although accuracy better than 1 percent is not necessary. For analog signals encoded with PFM, the precision is a function of the signal-to-noise ratio. At close range, the precision is an order of magnitude better than at maximum range. This feature, which will be discussed later in more detail, is equivalent to encoding with a variable bit-rate.

Since the frequency of the pulse contains the signal information, a measurement precision of 1 percent, for example, requires that the detection system be able to discern 100 different frequencies. The total frequency band B is therefore divided into 100 equal parts. The separation between frequencies, Δf_0 , is $B/100$. The power spectra of the pulses (Appendix A) have zeros of power density on each side of the center frequency at multiples of the reciprocal of T the pulse length. T is determined by the sampling rate necessary to telemeter the highest frequency of $f_1(t)$ at least better than at the Nyquist rate. The zeros of the power spectrum then should fall at multiples of $B/100$, or $\Delta f_0 = 1/T$. This gives the relation between the bandwidth, precision, and pulse

$$B = \frac{N}{T} \quad (2)$$

length, where $1/N$ is the precision constant.

The above analysis was made on the basis of the sampling time of a single channel $f_1(t)$ determining the pulse length T . Since time-division multiplexing is used, the pulse length must be shortened in proportion to the number of commutated channels. Therefore the bandwidth B will increase in proportion to the number of commutated channels.

In effect, the application of Equation 2 predicts that the sample will assume any one of N , in this case 100, discrete amplitudes. The power spectrum for any one of these amplitudes is centered in one of the 100 equal parts of width Δf_0 in the band B .

A first-order solution to the detection problem is to arrange 100 contiguous bandpass filters with bandwidth $\Delta f_0 = 1/T$ to cover the bandwidth B. A maximum-likelihood detector on the outputs of the filters can select the filter with the greatest output. The signal is thus quantized into one of 100 possible levels. The original amplitude of a sample would not necessarily cause the frequency to fall exactly in the center of any one of the 100 filters. The output would still be the greater in one filter than in an adjacent filter. As the amplitude of the sample is changed, the maximum-likelihood detector indicates that the adjacent filter has acquired the signal only when its response has an amplitude greater than the output of the original filter. With a good signal-to-noise ratio the filter output which is greatest can be gated into a discriminator in order that the frequency might be measured with better than 1 percent precision.

Rather than samplings only at times nT , the frequency of the pulse can be a continuous representation of the amplitude of the sample for the whole duration of the pulse, T. By using a discriminator on the output of the contiguous filters, not only can the average amplitude during the time T be determined but also the rate of change of amplitude. Having the latter information is equivalent to doubling the sampling rate (Reference 9). These topics will be discussed in detail later.

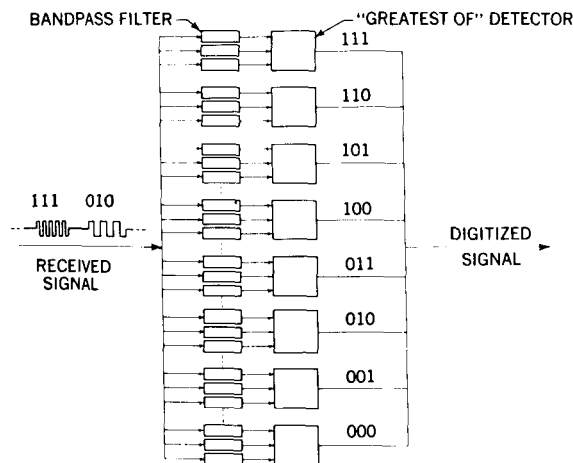


FIGURE 3.—Contiguous-filter bank in digital-oscillator mode.

For data encoded as digital information, the minimum number of filters required in the detection system is the same as the number of frequency levels encoded. Filters for these frequency levels as well as filters contiguous to these discrete frequencies are tied together in a "greatest of" configuration as in Figure 3. This allows some latitude in the stability of each discrete frequency-temperature drift or aging of the digital oscillator might cause a discrete frequency to fall outside its assigned filter. Thus, the "greatest of" detector selects the filter with the greatest signal and transmits to the output a digital number indicative only of the group of filters in which the signal fell.

CHARACTERISTICS OF CODED BINARY SEQUENCES

A somewhat different approach to the theory of coded binary sequences will be developed in this chapter in order to illustrate the position of pulse-frequency modulation in coded telemetry. The idea is to show that PFM is a special binary code taken from a group of many codes which can be made up of patterns of binary zeros and ones.

Pulse-code modulation (PCM) is ordinarily thought of as the representation of sequential samples of a signal by a binary code; however, the original definition of PCM included all codes: binary, ternary, quaternary, etc. Patterns of these code elements make up the quantized amplitude value of the sample. For two reasons the binary code is used almost exclusively: the advantage in the signal-to-noise-ratio relation, and the ease of generation. If the amplitude of a sample is to be encoded as an n -bit binary sequence, 2^n different sequences are available to quantize the amplitude. Shannon has shown by his second theorem that the probability of error in recognizing any of the transmitted 2^n sequences may be reduced by recoding the n -bit sequences or words into selected sequences of larger m -bit sequences (Reference 1); or, conversely, if only selected sequences of the 2^n available sequences are allowed, then the probability of error per bit is reduced. Coded n -bit binary sequences are defined as a set of M selected sequences in the $M < 2^n$. When $M = 2^n$ the set of sequences is said to be uncoded.

Group Coding

Advantages may be gained by allowing the transmission of only selected sequences in the available 2^n sequences of an n -bit encoded sample. This, of course, reduces the precision to which the samples may be quantized; however, the precision may be increased back to the original value by increasing n , since this increases the number of sequences from which a judicious selection may be made. Later we shall discuss how these selections are made. A signal-to-noise advantage is obtained if the selected sequences are detected not bit-by-bit but as an entire sequence or group. Use is made of the fact that signals add directly (since their phases are correlated) and noise adds as the square root of the sum of the squares.

The detection process is accomplished by cross-correlating each of the selected or allowed sequences with the transmitted signal; the allowed sequence which yields the highest cross-correlation coefficient when correlated with the signal is selected as the most probable representation of the signal.

Let $f_i(t)$ be one of a set of k transmitted sequences and $f_m(t)$ be one of the allowed sequences stored in k correlators. Both $f_i(t)$ and $f_m(t)$ are zero for $t > T$ and $t < 0$. The correlator performs the mathematical operation:

$$\frac{n}{T} \int_0^T f_i(t + \tau) f_m(t) dt = C_{im}(\tau), \quad (3)$$

where

n = number of bits in the word,

T = length of sequence,

τ = lag time,

$C_{im}(\tau)$ = unnormalized cross-correlation function.

For the matched or optimum condition the lag time should be zero. The correlator with the greatest value of $C_{im}(0)$ is selected as the one having the greatest probability of containing the signal.

A *zero* may be represented in the transmitted signal by a $+1$ volt and a *one* by a -1 volt. The selected sequence or stored waveform in the correlator $f_m(t)$ is normalized so that the units of the cross-correlation coefficient $C_{im}(0)$ will cor-

respond to the voltage measured at the output of the correlator. When $l=m$, the cross-correlation coefficient is n volts. Since the correlator with the highest coefficient is being selected, it is desirable for the coefficient of all the other correlators to be as low as possible, so that the largest possible distinction can be made between the correct and the incorrect values of the coefficient.

As an example, let us take all the available sequences of a four-bit word and construct a truth table which gives the values of the cross-correlation coefficients of Equation 3. Table 1 contains such a table; however, the order of the four-bit words was prearranged to bring out some salient features. For $l=m$ the correlation coefficient is $+4$ volts; when $f_i(t)$ equals the complement of $f_m(t)$, the coefficient becomes -4 volts. In Table 1(a) there are two groups of eight sequences each. In either group the cross-correlation coefficient is zero except when the stored waveform matches the signal or is its complement. If an attempt was made to correlate signals in group 1 with the stored waveforms of group 2, the cross-correlation coefficient would be ± 2 volts. Noise could easily perturb a $+2$ volt output into $+4$ volts, causing an error. Thus, if the transmitted sequences of a four-bit word are restricted entirely to either group 1 or group 2 and the stored waveforms in the correlators are of the same group, then an improvement in the signal-to-noise ratio can be realized. These code groups are of the Reed-Muller type (Reference 10).

Intuitively, if the signal-to-noise ratio were unity for a single bit of the four-bit word, the output signal-to-noise ratio would be $2/1$ (the bits of the signal add directly, but the noise adds as the square root of the sum of the squares). The signal-to-noise ratio has doubled and the information rate has reduced by only one-fourth. This is a good improvement; unfortunately, it cannot be realized. True, the signal-to-noise ratio out of the correlator is improved over the bit-by-bit detection, but the outputs of the correlators must be compared to determine the largest coefficient. It is probable that noise in another correlator will be greater than the coefficient in the correct correlator. This probability increases as the number of correlators increases. The situation is discussed in detail

Table 1

Correlation Coefficients: (a) Four-Bit Sequences; (b) One-Bit Sequences; (c) Two-Bit Sequences.

		SIGNAL - $f_i(t)$																
		GROUP 1								GROUP 2								
		0000	0011	0101	0110	1001	1010	1100	1111	0001	0010	0100	0111	1000	1011	1101	1110	
GROUP 1	CORRELATOR - $f_m(t)$	0000	+4	0	0	0	0	0	0	-4	+2	+2	+2	-2	+2	-2	-2	-2
	0011	0	+4	0	0	0	0	0	-4	0	+2	+2	-2	+2	-2	+2	-2	-2
	0101	0	0	+4	0	0	-4	0	0	0	+2	-2	+2	+2	-2	-2	+2	-2
	0110	0	0	0	+4	-4	0	0	0	0	-2	+2	+2	+2	-2	-2	-2	+2
	1001	0	0	0	-4	+4	0	0	0	0	+2	-2	-2	-2	+2	+2	+2	-2
	1010	0	0	-4	0	0	+4	0	0	0	-2	+2	-2	-2	+2	+2	-2	+2
	1100	0	-4	0	0	0	0	+4	0	0	-2	-2	+2	-2	+2	-2	+2	+2
	1111	-4	0	0	0	0	0	0	+4	0	-2	-2	-2	+2	-2	+2	+2	+2
GROUP 2	CORRELATOR - $f_m(t)$	0001	+2	+2	+2	-2	+2	-2	-2	-2	+4	0	0	0	0	0	0	-4
	0010	+2	+2	-2	+2	-2	-2	-2	-2	0	+4	0	0	0	0	0	-4	0
	0100	+2	-2	+2	+2	-2	-2	+2	-2	0	0	+4	0	0	-4	0	0	0
	0111	-2	+2	+2	+2	-2	-2	-2	+2	0	0	0	+4	-4	0	0	0	0
	1000	+2	-2	-2	-2	+2	+2	+2	-2	0	0	0	-4	+4	0	0	0	0
	1011	-2	+2	-2	-2	+2	+2	-2	+2	0	0	-4	0	0	+4	0	0	0
	1101	-2	-2	+2	-2	+2	-2	+2	+2	0	-4	0	0	0	0	0	+4	0
	1110	-2	-2	-2	+2	-2	+2	+2	+2	-4	0	0	0	0	0	0	0	+4

(a)

		$f_i(t)$	
		0	1
$f_m(t)$	0	+1	-1
	1	-1	+1

(b)

		$f_i(t)$			
		00	01	10	11
$f_m(t)$	00	+2	0	0	-2
	01	0	+2	-2	0
	10	0	-2	+2	0
	11	-2	0	0	+2

(c)

later. The group code method still provides a substantial improvement in the signal-to-noise ratio over bit-by-bit detection.

The number of correlators in group 1 can be reduced from eight to four if the -4 volt output also is utilized. Since each correlator responds only to its matched signal and its complement, a single correlator can be used to detect two sequences. The correlator with the greatest absolute magnitude is selected as the one having the best probability of containing the signal. The sign of the coefficient determines whether the signal was in the first or second four sequences of a group. A code of this type is said to be bi-orthogonal since negative as well as positive cross-correlation coefficients are utilized. When only the positive outputs are used the code is orthogonal.

In the preceding example the prearranged order

of the four-bit sequences was most important. A few fundamental rules have been developed to generate codes with bit lengths of $n=2^k$, where $k=0, 1, 2, 3, \dots$. These rules are amenable to computer programming but are somewhat different from those found in the literature. Consider first the correlation table for a one-bit word as in Table 1(b). Only one correlator is required if the bi-orthogonal properties of the signal are utilized. This table and the correlation table for the two-bit sequences of Table 1(c) can be written by inspection. The first word of the four-bit table is generated by combining the first word, 00, of the two-bit table with itself to form 0000. The next word is made by combining the first word, 00, with its complement, 11, to form 0011. The same process is carried out with the second word of the two-bit table to form the next two words. The complements of these four words

are now taken in reverse order; this generates the code of group 1. The group 2 code is started by combining the first word of the two-bit table with the second. The next word is obtained by combining the first word of the two-bit table with the complement of the second. For the third and fourth words the two-bit second word is combined with the first word and then with its complement. The next four words are the complements of the newly generated words in reverse order.

The eight-bit correlation table has 256 sequences which form a matrix of 65,536 correlation coefficients. A digital computer program for gen-

erating the 8-bit table was written by utilizing the four-bit table and the rules for forming the table from the preceding paragraph. The program was written in FAP (Fortran Assembly Program) so that the bit patterns of the signals $f_i(t)$ and stored waveforms $f_m(t)$ could be cyclically varied. That is, the first bit of all the eight-bit sequences was moved to the last-bit position and a new set of correlation coefficients computed.

The program is arranged so that a print-out contains the correlation between the first sixteen signals and the first sixteen stored waveforms. Table 2 is a print-out of the first page of the code

Table 2
Correlation Coefficients for Bi-Orthogonal Code Group.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
	0	C	0	C	0	C	C	0	1	1	1	1	1	1	1	1	
	0	0	0	0	1	1	1	1	1	0	C	0	1	1	1	1	
	0	0	1	1	0	C	1	1	C	1	1	1	C	0	1	1	
	0	0	1	1	1	1	C	C	1	1	C	C	0	0	1	1	
	0	1	0	1	C	1	C	1	C	1	C	1	0	1	0	1	
	0	1	0	1	1	C	1	C	1	C	1	0	C	1	0	1	
	C	1	1	C	C	1	1	C	1	C	1	1	1	0	C	1	
	0	1	1	0	1	0	C	1	C	1	1	C	1	0	0	1	
1	00000000	8	C	C	C	0	0	C	C	C	0	0	0	0	0	-8	
2	00001111	0	8	0	0	0	0	C	0	1	0	C	0	C	0	-8	0
3	00110011	0	C	8	C	C	0	C	0	C	C	C	0	-8	0	C	
4	00111100	C	0	C	8	C	0	C	0	C	0	0	-8	0	C	0	
5	01010101	0	C	C	C	8	0	C	0	C	0	0	-8	C	0	0	
6	01011010	0	0	0	C	0	8	C	0	C	0	-8	C	C	0	0	
7	01100110	0	C	0	C	0	C	8	C	C	-8	C	C	0	0	0	
8	01101001	0	0	C	C	0	C	C	8	-8	0	C	0	0	0	C	
9	10010110	0	C	C	0	C	0	C	-8	8	0	C	0	0	0	C	
10	10011001	0	C	C	C	C	C	-8	0	C	8	0	0	0	C	0	
11	10100101	0	C	0	C	C	-8	C	0	C	C	8	C	0	0	0	
12	10101010	0	0	C	C	-8	C	C	0	C	C	8	C	0	0	0	
13	11000011	0	0	C	-8	C	C	0	C	C	C	C	8	0	0	C	
14	11001100	0	0	-8	0	C	C	C	0	C	C	C	C	8	0	C	
15	11100000	0	-8	0	C	C	C	C	0	C	C	C	0	C	8	0	
16	11111111	-8	C	C	C	C	C	C	0	C	C	0	C	0	0	8	

Table 3
Cross-Correlation Coefficients for Eight-Bit Sequences.

		33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
		0	0	0	0	C	0	C	0	1	1	1	1	1	1	1	1
		0	0	0	0	1	1	1	1	C	0	C	C	1	1	1	1
		0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
		0	0	1	1	1	1	0	0	1	1	C	0	0	0	1	1
		0	1	0	1	C	1	0	1	C	1	C	1	0	1	C	1
		1	0	1	0	C	1	C	1	C	1	C	1	1	0	1	0
		0	1	1	0	C	1	1	0	1	C	C	1	1	0	0	1
		1	0	0	1	C	1	1	0	1	C	0	1	0	1	1	0
113	00001111	6	2	2	-2	-2	2	2	-2	2	-2	-2	2	2	-2	-2	-6
114	00C01000	2	6	-2	2	2	-2	-2	2	-2	2	2	-2	-2	2	-6	-2
115	00110001	2	-2	2	6	2	-2	2	-2	2	-2	2	-2	-6	-2	2	-2
116	00111110	-2	2	6	2	-2	2	-2	2	-2	2	-2	2	-2	-6	-2	2
117	01010010	-2	2	2	-2	6	2	2	-2	2	-2	-2	-6	2	-2	-2	2
118	01011101	2	-2	-2	2	2	6	-2	2	-2	2	-2	-2	-2	2	2	-2
119	01100100	2	-2	2	-2	2	-2	2	6	-6	-2	2	-2	2	-2	2	-2
120	01101011	-2	2	-2	2	-2	2	2	-2	-6	-2	2	2	-2	2	-2	2
121	10010100	2	-2	2	-2	2	-2	-2	-2	2	6	2	-2	2	-2	2	-2
122	10011011	-2	2	-2	2	-2	2	-2	-6	2	2	-2	2	-2	2	-2	2
123	10100010	-2	2	2	-2	-2	-6	2	-2	2	-2	6	2	2	-2	-2	2
124	10101101	2	-2	-2	2	-6	-2	-2	2	-2	2	2	6	-2	2	2	-2
125	11000001	2	-2	-6	-2	2	-2	2	-2	2	-2	2	-2	2	6	2	-2
126	11001110	-2	2	-2	-6	-2	2	-2	2	-2	2	-2	2	6	2	-2	2
127	11110111	-2	-6	2	-2	-2	2	2	-2	2	-2	-2	2	2	-2	6	2
128	11111000	-6	-2	-2	2	2	-2	-2	2	-2	2	2	-2	-2	2	2	6

book for the program. On the next page are the first sixteen signals correlated with the second sixteen correlators, and so on. There are sixteen code groups of the type in Table 2. Any one of these sixteen code groups could be used to encode four bits of data for a signal-to-noise-ratio improvement over a standard four-bit code. Table 3 illustrates the correlation between groups.

It is interesting to note that a cyclic variation of the four-bit sequences does not generate any new codes. It only rearranges the order in the two code groups. In the eight-bit case, however,

new codes are generated; in fact, a tremendous number of new codes can be made not only by cyclically varying the bit locations of the signal and stored waveforms but also by interchanging the order of the bits in all possible combinations.

The process may be carried to the sixteen-bit case, for which there are over 4×10^9 cross-correlation coefficients. Sanders has described the Digilock coded binary PCM telemeter, which encodes sixteen-bit words to represent five-bit samples (Reference 11). This system has been flown on an Air Force Blue Scout rocket.

a table. Each frequency contains an integral number of cycles because the pulse length, derived from Equation 3, is constant. The table is strictly an orthogonal code. If the sequences starting with a *one* instead of a *zero* were also used, the code would become bi-orthogonal, since negative values of the cross-correlation coefficient would result. Finally, if the frequencies of PFM are restricted to a set containing an integral number of cycles, then it appears that PFM is certainly a form of coded binary PCM. Of course, it is a special group code, chosen primarily because of its ease of generation.

NOISE ANALYSIS

In the previous section pulse-frequency modulation was exemplified by a special set of binary coded sequences which were selected because of their ease of generation. Of primary interest are the characteristics of this code set in the presence of noise. A number of methods are available to aid in recovering the signal set from noise. The optimum method is to use matched filters with maximum-likelihood detection. The number of filters employed is a function of the desired efficiency of encoding; n -bit precision for a single

word requires 2^{n-1} filters. The matched-filter technique has been well reported (References 12 and 13). The idea is to force the power spectrum of the impulsive response of the filter to be identical to the power spectrum of the signal. Although the use of matched filters is most desirable, there are numerous situations in which the use of unmatched filters can markedly reduce hardware and implementation complexity without too great a degradation in the signal-to-noise ratio. The noise analyses differ, depending on whether the filter in the detection system is matched or unmatched.

Unmatched Filter

The use of unmatched filters for PFM signals in which the number of quantized levels exceeds eight ($n=3$) can simplify the operation of the detection system, compared with the use of a matched-filter set. The unmatched-filter set would consist of N contiguous bandpass filters, where N is the number of quantized levels in the signal.

Let us examine one of these filters in detail. Assume the filter to have the following power spectrum:

$$\left. \begin{aligned} K(\omega) &= 0 && \text{for } \omega < \left(\omega_0 - \frac{\Delta\omega_0}{2} \right), \\ K(\omega) &= 1 && \text{for } \left(\omega_0 - \frac{\Delta\omega_0}{2} \right) < \omega < \left(\omega_0 + \frac{\Delta\omega_0}{2} \right), \\ K(\omega) &= 0 && \text{for } \omega > \left(\omega_0 + \frac{\Delta\omega_0}{2} \right), \end{aligned} \right\} \quad (4)$$

where $\Delta\omega_0$ is the bandwidth of the filter and ω_0 is the center frequency of the filter.

White additive Gaussian noise with a power density of N_0 watts/cps will have a variance of σ^2 watts in the output when passed through this filter. The noise output of a narrow-band filter may be represented by using inphase and quadrature components:

$$u(t) = x(t) \cos \omega_0 t + y(t) \sin \omega_0 t \quad (5)$$

where $x(t)$ and $y(t)$ are random variables.

Neither the $x(t)$ nor $y(t)$ frequency spectrum extends up to the center frequency ω_0 . The mean-square-deviation or variance σ^2 is the sum of the variances of the inphase and quadrature components, $\sigma^2 = \sigma_I^2 + \sigma_Q^2$ (the covariance between the inphase and quadrature components is zero); also $\sigma^2 = \sigma_x^2/2 + \sigma_y^2/2$, where σ_x^2 and σ_y^2 are the variances of $x(t)$ and $y(t)$.

It would be rare indeed for a filter to have infinite slope on its band edges. This postulates a filter having an infinite number of sections with infinite delay time; however, a simple RLC

single-pole filter serves as the first approximation to this ideal filter. The following analysis will involve only the ideal filter.

Ideal Filter

Consider a single ideal filter of bandwidth $\Delta\omega$ and center frequency ω_0 . In the presence of N_0 watts/cps of noise on the input, the output power measured in a matched 1 ohm load would be $N_0\Delta\omega/2\pi$ or σ^2 watts. If a step sinusoidal voltage of amplitude C_0 and frequency ω_0 is added to the input, the correct prediction of whether or not the signal is present after a period T has elapsed is a matter of statistics. The sinusoidal signal will be termed a "word," and the probability of incorrectly assuming the presence of a signal will be termed the "word-error probability."

Although the noise output of the filter cannot be predicted, certain probabilities may be assigned to the noise characteristics in order to evaluate the long-term effect of the presence of the noise. Some of these characteristics are the first and second moments of the noise, the probability of the rate of zero crossings, and the envelope probability density function. For the unmatched-filter case this report will be concerned mainly with the probability density function of the envelope amplitude of the noise output, since an uncorrelated detector produces a voltage proportional to the envelope of its input.

The amplitude of the noise in the output of the unmatched filter, like that in the input, has a Gaussian probability density function. Positive and negative amplitudes are equiprobable, zero being the most probable value. The autocorrelation function $\phi_{11}(\tau)$ of the input noise extending to the upper frequency limit ω_c is

$$\phi_{11}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T u(t) u(t+\tau) dt \quad (6)$$

From theorem 202 in Reference 14

$$\phi_{11}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{|F(\omega)|^2}{T} e^{i\omega\tau} d\omega \quad (7)$$

where

$$F(\omega) = \int_{-\infty}^{\infty} u(t) e^{-i\omega t} dt$$

Then

$$\phi_{11}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{P(\omega)}{2} e^{i\omega\tau} d\omega \quad (8)$$

since

$$\lim_{T \rightarrow \infty} \frac{|F(\omega)|^2}{T} = \frac{P(\omega)}{2}$$

where $P(\omega)$ is the power spectrum of $u(t)$. But $P(\omega) = N_0$ watts/cps and

$$\begin{aligned} \phi_{11}(\tau) &= \frac{N_0}{2} \int_{-\omega_c}^{\omega_c} e^{i\omega\tau} \frac{d\omega}{2\pi} \\ &= \frac{N_0 \omega_c}{2\pi} \frac{\sin \omega_c \tau}{\omega_c \tau} \end{aligned} \quad (9)$$

The autocorrelation function is periodic as τ is increased, the correlation damping out approximately as a $1/\tau$ function. As ω_c approaches infinity $\phi_{11}(\tau)$ becomes an impulse function centered at $\tau=0$:

$$\begin{aligned} \lim_{\omega_c \rightarrow \infty} \phi_{11}(\tau) &= \lim_{\omega_c \rightarrow \infty} \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} \frac{N_0}{2} e^{i\omega\tau} d\omega \\ &= \frac{N_0}{2} \delta(\tau) \end{aligned} \quad (10)$$

where

$$\int_{-\infty}^{\infty} \delta(\tau) d\tau = 1$$

For the output of the bandpass filter ($\omega_c \gg \omega_0$) the autocorrelation function for Gaussian noise becomes (Appendix B)

$$\phi_{11}(\tau) = N_0 \frac{\Delta\omega}{2\pi} \frac{\sin \frac{\Delta\omega\tau}{2}}{\frac{\Delta\omega\tau}{2}} \cos \omega_0 \tau \quad (11)$$

The function is plotted in Figure 4. The output of the filter tends to exhibit periodic correlation until the lag time $\tau_0 = 2\pi/\Delta\omega$ is reached. Let $n(t)$ be the noise voltage output and e_d the detector voltage output. When the output of the

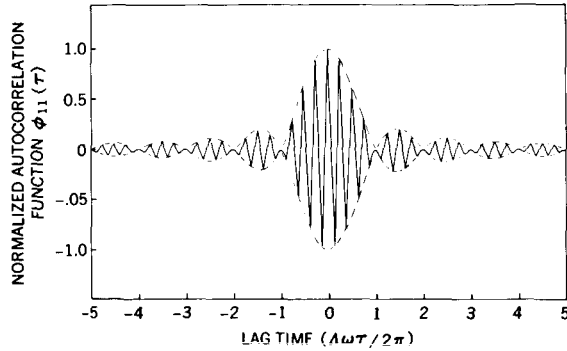


FIGURE 4.—Autocorrelation function for Gaussian noise ($\omega_0 = 4\Delta\omega$).

filter is passed through a simple detector with the characteristics

$$\left. \begin{aligned} e_d &= -n(t) && \text{for } n(t) \leq 0, \\ e_d &= +n(t) && \text{for } n(t) \geq 0, \end{aligned} \right\} \quad (12)$$

full-wave rectification of the output occurs. Smoothing of this rectified output yields $\sqrt{|x(t)|^2 + |y(t)|^2}$, the amplitude of the envelope of $u(t)$ (Equation 5).

Of great interest is the joint probability density function for the amplitudes of $x(t)$ and $y(t)$. Since $x(t)$ and $y(t)$ are equiprobable, have a normal distribution, and are statistically independent, they may be conveniently plotted on

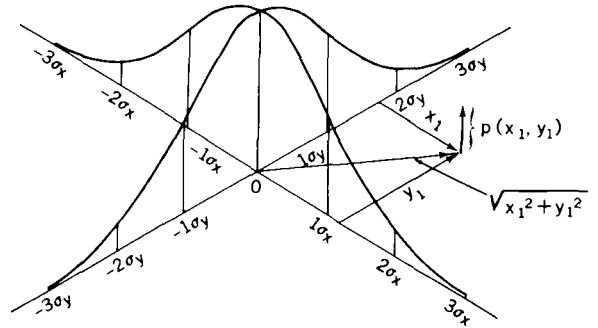


FIGURE 5.—Joint probability density function for Gaussian noise.

orthogonal coordinates with the z axis representing the joint probability, as in Figure 5. The amplitude at the point x_1, y_1 represents the probability density of the joint occurrence of the amplitude x_1 at the same time as the quadrature amplitude y_1 . The distance from the center to this point is $\sqrt{x_1^2 + y_1^2}$; however, other values of x and y with the same square-sum value lie on the circle with radius $\sqrt{x_1^2 + y_1^2}$. This circle is the locus of points of constant probability density for the amplitude of the envelope of the noise. To show this the variables must be changed from rectangular coordinates x and y to polar coordinates r and θ .

The probability density for the joint occurrence of x_1 and y_1 is:

$$\begin{aligned} p(x_1, y_1) &= \frac{1}{\sqrt{2\pi} \sigma_x} e^{-x_1^2/2\sigma_x^2} \frac{1}{\sqrt{2\pi} \sigma_y} e^{-y_1^2/2\sigma_y^2} \\ &= \frac{1}{2\pi\sigma_x \sigma_y} e^{-\frac{1}{2}\left(x_1^2/\sigma_x^2 + y_1^2/\sigma_y^2\right)}. \end{aligned} \quad (13)$$

Now $x=r \cos \theta$ and $y=r \sin \theta$. The Jacobian is

$$\begin{aligned} \left| t_k^k \right| &= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} \end{vmatrix} \\ &= \begin{vmatrix} \cos \theta & \sin \theta \\ -r \sin \theta & r \cos \theta \end{vmatrix} \\ &= r. \end{aligned}$$

The joint probability for r and θ is then

$$\begin{aligned}
 p(r, \theta) &= \left| t_{\mathbf{r}}^k \right| p(x, y) \\
 &= \frac{r}{2\pi\sigma_x \sigma_y} e^{-\frac{1}{2} \left(r^2 \cos^2 \theta / \sigma_x^2 + r^2 \sin^2 \theta / \sigma_y^2 \right)} .
 \end{aligned}
 \tag{14}$$

In the present case $\sigma_x = \sigma_y$, so

$$p(r, \theta) = \frac{r}{2\pi\sigma_x^2} e^{-r^2/2\sigma_x^2} . \tag{15}$$

The probability for any value of θ is

$$\begin{aligned}
 p(r) &= \int_0^{2\pi} p(r, \theta) d\theta \\
 &= \frac{r}{\sigma_x^2} e^{-r^2/2\sigma_x^2} .
 \end{aligned}
 \tag{16}$$

This is the Rayleigh distribution which describes the probability of obtaining a specified amplitude of the envelope of bandwidth-limited random noise.

Rayleigh noise is characterized by a probability which is zero for zero envelope amplitude, a

maximum for amplitude equal to the deviation, and low for large amplitudes.

Word-Error Probability for the Matched Filter

The addition of a sinusoidal signal to the noise causes a shift in the envelope probability density function, as shown by Rice (Reference 15). Adding a sine wave carrier of amplitude A_0 displaces the probability density function a distance of A_0 units along the x axis. In the limit, as the variance of the noise approaches zero the probability density function for the envelope approaches an impulse at $x = A_0$. The volume under this impulse is unity, so the probability of an amplitude other than A_0 is zero. Adding a cosine of amplitude B_0 shifts the center to $\sqrt{A_0^2 + B_0^2}$ at an angle $\theta_0 = \tan^{-1}(B_0/A_0)$ to the x axis. Figure 6 depicts this general case. The

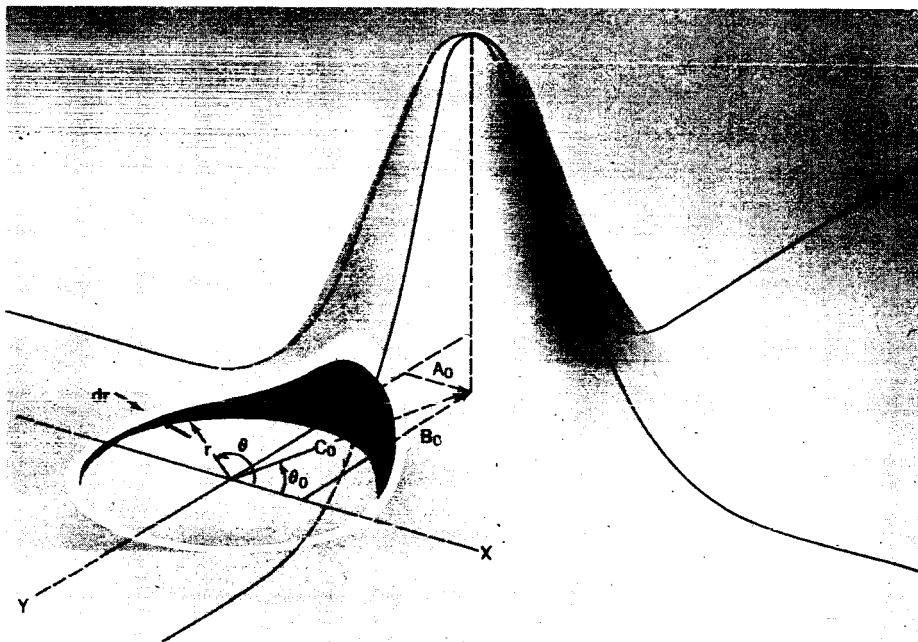


FIGURE 6.—Joint probability density function for signal plus Gaussian noise.

volume above any area of the x-y plane gives the probability of the envelope having an amplitude falling in the region of that x-y area.

However, we are mainly interested in constant envelope amplitudes with unknown phases since they characterize a typical PFM signal. The

probability of an amplitude lying between r and $r+dr$ (when $dr = \sqrt{dx^2 + dy^2}$) is the volume of the ring of radius r and width dr ; the height is given by the joint probability density between x and y . The volume in the clear segment of Figure 6 gives this probability:

$$p(r) = \int_0^{2\pi} \frac{1}{\sqrt{2\pi} \sigma_x} e^{-(r \cos \theta - A_0)^2 / 2\sigma_x^2} \frac{1}{\sqrt{2\pi} \sigma_y} e^{-(r \sin \theta - B_0)^2 / 2\sigma_y^2} r dr d\theta \quad (17)$$

Rice has compiled the probability curves for various carrier amplitudes (Reference 15).

The probability that an envelope amplitude will be equal to or less than r is given by the

volume inside the circle of radius r_0 :

$$P(r < r_0) = \int_0^{r_0} \int_0^{2\pi} \frac{1}{\sqrt{2\pi} \sigma_x} e^{-(r \cos \theta - A_0)^2 / 2\sigma_x^2} \frac{1}{\sqrt{2\pi} \sigma_y} e^{-(r \sin \theta - B_0)^2 / 2\sigma_y^2} r d\theta dr$$

When $\sigma_x = \sigma_y$,

$$\begin{aligned} P(r < r_0) &= \int_0^{r_0} \int_0^{2\pi} \frac{r}{2\pi\sigma_x^2} e^{-[r^2(\cos^2 \theta + \sin^2 \theta) - 2r(A \cos \theta + B \sin \theta) + A_0^2 + B_0^2] / 2\sigma_x^2} d\theta dr \\ &= \frac{1}{\sigma_x^2} \int_0^{r_0} r e^{-(r^2 + A_0^2 + B_0^2) / 2\sigma_x^2} \frac{1}{2\pi} \int_0^{2\pi} e^{r(A_0 \cos \theta + B_0 \sin \theta) / \sigma_x^2} d\theta dr \end{aligned} \quad (18)$$

Let $A_0 = C_0 \cos \theta_0$ and $B_0 = C_0 \sin \theta_0$. Then

$$P(r < r_0) = \frac{1}{\sigma_x^2} \int_0^{r_0} r e^{-(r^2 + C_0^2) / 2\sigma_x^2} \frac{1}{2\pi} \int_0^{2\pi} e^{r C_0 \cos(\theta - \theta_0) / \sigma_x^2} d\theta dr \quad (19)$$

The second integral in the preceding equation is the modified Bessel function $I_0(r)$.

We have examined the action of an individual ideal bandpass filter and, given signal-with-noise input, have determined the probability of a particular output. Now let us look at the action of signal and noise as an input to a bank of parallel bandpass filters (the band edges of the filters are touching). With N filters, a frequency range of $N\Delta\omega/2\pi$ is covered. It should be remembered that the primary advantage of PFM is that each PFM word is orthogonal to any other word in that PFM set. This gives a uniqueness to the power spectrum of each word, resulting in a bandwidth equal to the products of the number of words in the PFM set and the bandwidth of the power spectrum of a word.

Consider a PFM word and white additive Gaussian noise as an input to the bank of filters.

The sinusoidal frequency of the word lies in the center of one of the N bandpass filters; the probability density function for that filter will appear as in Figure 7(a). The remaining $N-1$ filters will have a probability density function as given in Figure 7(b). Ultimately we are looking for a number which represents the ratio of the number of times that noise in one or more of the $N-1$ filters is greater than the signal-plus-noise in the correct filter. This figure is defined as the word-error probability.

It is interesting to consider some of the aspects of those things which influence the word-error probability. First, with N filters in parallel and the outputs connected in a "greatest of" configuration, the probability that at least one of the N filters will exceed a certain amplitude r_0 is almost N times the probability that a single filter output will exceed that amplitude. For

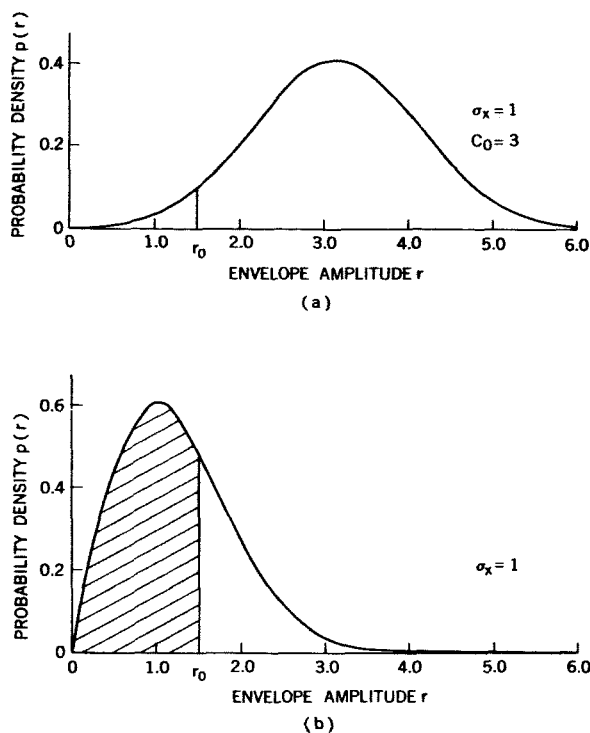


FIGURE 7.—Probability density functions: (a) signal plus Rayleigh noise; (b) Rayleigh noise.

example, for a bank of 100 filters with a threshold set on the "greatest of" output of 3 times the average noise voltage of a single filter, the number of times that this level is exceeded by any filter is less than 100 times the average rate a single filter will exceed that amplitude. The reason the rate is not exactly 100 times greater is the finite probability that two (or more) of the filters may simultaneously exceed the threshold. In Figure 7(b) the shaded area represents the probability that the envelope amplitude of one filter is less than the threshold amplitude r_0 . The probability that all the filters, except the one with the signal, have amplitudes less than r_0 is $P(r < r_0)^{N-1}$. The

probability that at least one of the amplitudes is greater than r_0 then becomes $1 - [P(r < r_0)]^{N-1}$. By increasing the signal an amount Δr_0 , the error rate becomes the same for $N-1$ filters as for a single filter. From this increase of r_0 by the amount Δr_0 , the following equality can be set up:

$$1 - P(r < r_0) = 1 - [P(r < r_0 + \Delta r_0)]^{N-1} \quad (20)$$

The fact that the signal amplitude would also be increased by an amount Δr_0 is more than compensated for by the increase of the information rate by $\log_2 100$ over that of a single filter. In order to keep the information rate constant, the bandwidth of each of the N filters can be reduced by the factor $1/\log_2 100$, which substantially decreases the signal power needed to maintain the same word-error probability at the same information rate.

The calculation of the word-error-probability curves for a set of 2^n uncorrelated filter outputs may proceed generally as above; i.e., the probability that an output is correct is calculated and then subtracted from unity for the probability of error. The probability of one of the $N-1$ filters having an amplitude equal to or less than the envelope amplitude r_0 of the filter containing the signal is given by

$$P(r < r_0) = \frac{1}{\sigma_x^2} \int_0^{r_0} r e^{-r^2/2\sigma_x^2} dr \quad (21)$$

However, the probability that noise amplitude in any one of the $N-1$ filters does not exceed r_0 is the probability distribution function raised to the $N-1$ power. The probability that the signal-plus-noise in the correct filter has the amplitude r_0 is

$$P(r_0) = \frac{r_0}{\sigma_x^2} e^{-(r_0^2 + C_0^2)/2\sigma_x^2} \frac{1}{2\pi} \int_0^{2\pi} e^{r_0 C_0 \cos(\theta - \theta_0)/\sigma_x^2} d\theta \quad (22)$$

Then, the joint probability that none of the $N-1$ filters has noise exceeding r_0 and that the signal-

plus-noise in the correct filter has the amplitude r_0 is

$$P_c(N) = \left(\frac{1}{\sigma_x^2} \int_0^{r_0} r e^{-r^2/2\sigma_x^2} dr \right)^{N-1} \frac{r_0}{\sigma_x^2} e^{-(r_0^2 + C_0^2)/2\sigma_x^2} \frac{1}{2\pi} \int_0^{2\pi} e^{r_0 C_0 \cos(\theta - \theta_0)/\sigma_x^2} d\theta dr_0 \quad (23)$$

For the total probability that the amplitude of the correct filter is always greater than any amplitude in the $N-1$ filters, the integration is

$$P_e(N) = 1 - \int_0^\infty \left(\frac{1}{\sigma_x^2} \int_0^{r_0} r e^{-r^2/2\sigma_x^2} dr \right)^{N-1} \frac{r_0}{\sigma_x^2} e^{-(r_0^2 + c_0^2)/2\sigma_x^2} \frac{1}{2\pi} \int_0^{2\pi} e^{r_0 c_0 \cos(\theta - \theta_0)/\sigma_x^2} d\theta dr_0. \quad (24)$$

This equation was programmed on an IBM 7090 digital computer for various numbers of filters; the resulting curves are shown in Figure 8. The value of n in the figure is equivalent to $\log_2 N$ in Equation 24. This notation was used so that direct comparisons could be made with Viterbi's curves for "greatest of" detectors using matched filters (Reference 7).

Matched Filter

It was stated for the case of the unmatched filter that a bandpass filter is the first approximation to a matched filter for a pulsed sine wave, $C_0 \sin \omega_0 t [u(t) - u(t - T_0)]$. A matched filter is one which yields the maximum signal-to-noise ratio at a time $t = T_0$ for a given signal at the input accompanied with additive white Gaussian noise. Several excellent treatments are available in the literature for the proof of this (Reference 16). The transfer function of the matched filter is proportional to the complex conjugate of the spectrum of the signal, $S^*(\omega)$:

$$k(\omega) = e^{-i\omega T_0} S^*(\omega).$$

The factor $e^{-i\omega T_0}$ is the complex shifting function and corresponds to the delay time of the filter.

Multiple Matched Filter

Essentially, a matched filter performs the process of cross-correlating the signal-plus-noise with a stored matching signal. This process is actually autocorrelation of the signal with the stored image of the signal in the matched filter. The cross-correlation function,

$$\frac{1}{T_0} \int_0^{T_0} f_l(t) f_m(t + \tau) dt,$$

is a maximum for $l=m$ when $\tau=0$. If other

signals are added to the set such that

$$\frac{1}{T_0} \int_0^{T_0} f_l(t) f_m(t) dt = 0 \quad \text{for } l \neq m. \quad (25)$$

then any one filter in the set of matching filters will respond to only one signal in the set; i.e.,

$$\frac{1}{T_0} \int_0^{T_0} f_l(t) f_m(t) dt = 0 \quad \text{for } l \neq m,$$

and

$$\frac{1}{T_0} \int_0^{T_0} f_l(t) f_m(t) dt = C_{lm} \quad \text{for } l = m.$$

By observing the outputs of the filters when $\tau=0$ (maximum autocorrelation), a probability can be established that the matched filter with the greatest output contains the signal. The error probability is, then, 1 minus the probability of being correct.

The waveform for the set of signals $f_l(t)$ has not been given. It is possible to use any orthonormal set, such as Legendre polynomials, Bessel functions, elliptic functions, or sine functions. PFM is based on the use of a set of sine functions having orthonormal properties. The time-limited sine wave of Figure 9(a) is multiplied by its matching-filter stored image of Figure 9(b). Figure 9(c) gives the product and Figure 9(d) the integral of the product as a function of time for $\tau=0$. The autocorrelation function is shown in Figure 9(e). If the frequency is increased so that the signal has an extra half of a sine wave in the same pulse length T_0 , then the signal is orthogonal to the filter, i.e., no match. Figure 10 shows the new signal of higher frequency, the correlating waveform (stored image in the filter), the product, the integral, and the general cross-correlation function.

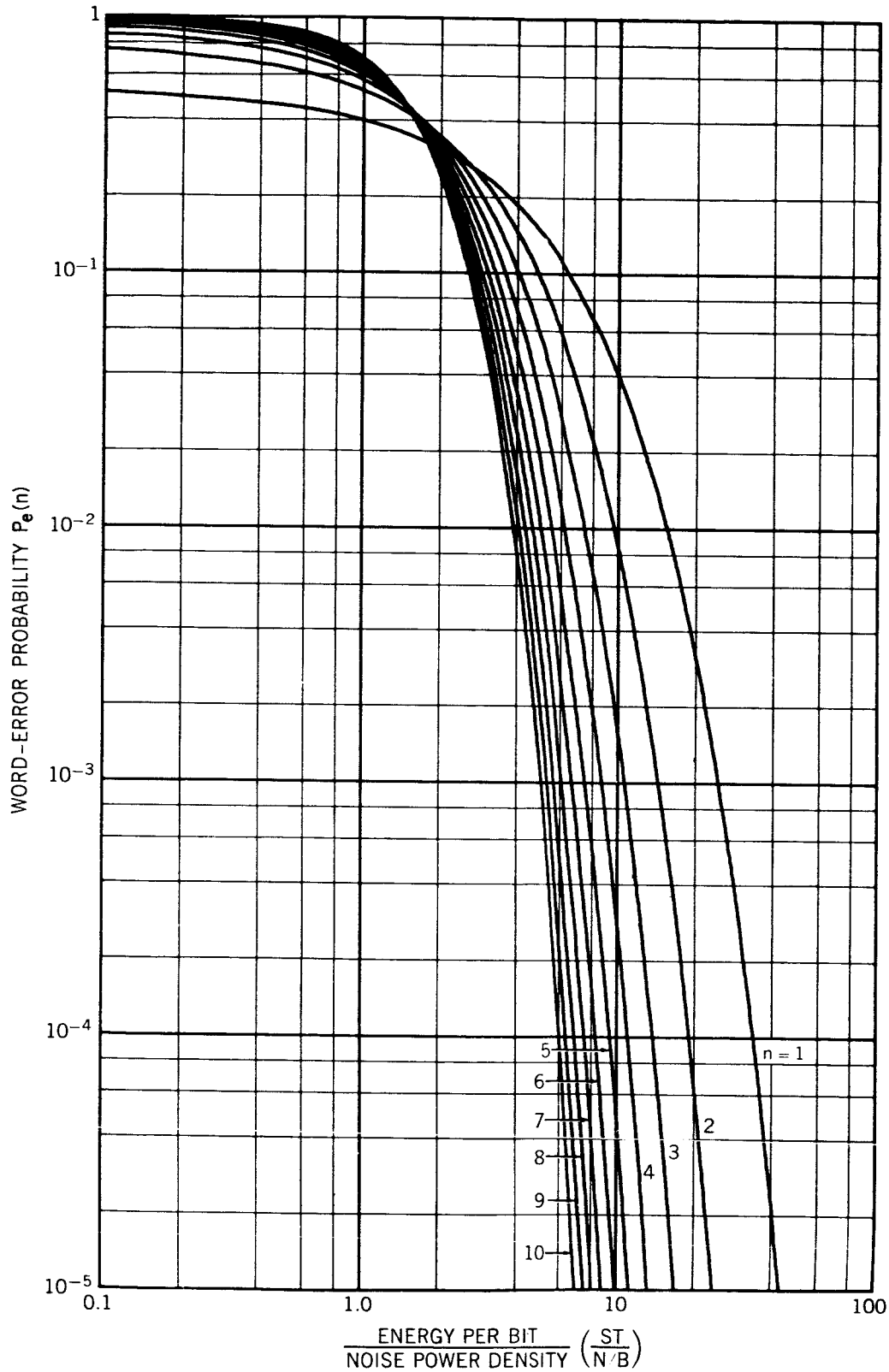


FIGURE 8.—Word-error probability curves for Rayleigh noise.

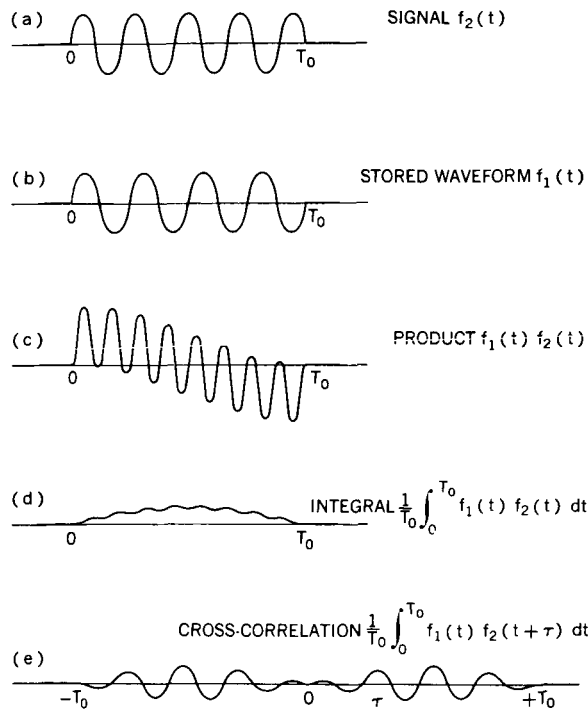


FIGURE 9.—Matched-filter detection: (a) time-limited sine wave; (b) matching-filter stored image; (c) product; (d) autocorrelation function for $\tau=0$; (e) autocorrelation function for values of τ .

Of interest is a plot of the cross-correlation function for $\theta=0$ as a function of frequency:

$$\begin{aligned}
 \phi_{mn}(0) &= \frac{1}{T_0} \int_0^{T_0} \sin \omega_m t \sin \omega_n t dt \\
 &= \frac{1}{2T_0} \left[\frac{\sin(\omega_m - \omega_n) t}{\omega_m - \omega_n} - \frac{\sin(\omega_m + \omega_n) t}{\omega_m + \omega_n} \right]_0^{T_0} \\
 &= \frac{\sin(\omega_m - \omega_n) T_0}{2(\omega_m - \omega_n) T_0} - \frac{\sin(\omega_m + \omega_n) T_0}{2(\omega_m + \omega_n) T_0} .
 \end{aligned} \tag{26}$$

The contribution of the second term will be small, so the covariance or cross-correlation for $\theta=0$ is a $\sin x/x$ function. The eigenvalues of this func-

The eigenvalues occur when

$$\frac{1}{2} \frac{\sin(\omega_m - \omega_n) T_0}{(\omega_m - \omega_n) T_0} - \frac{1}{2} \frac{\sin(\omega_m + \omega_n) T_0}{(\omega_m + \omega_n) T_0} = 0 . \tag{27}$$

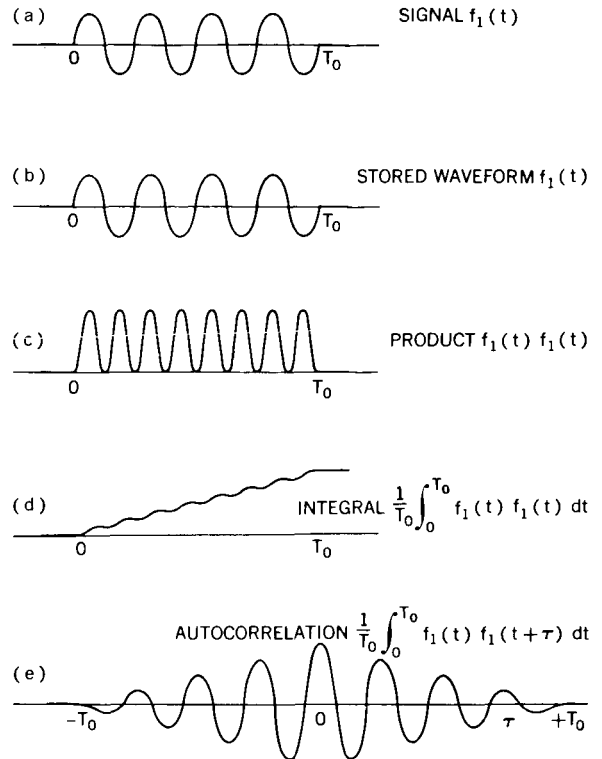


FIGURE 10.—Matched-filter detection: (a) time-limited sine wave at higher frequency than stored image; (b) stored image; (c) product; (d) cross-correlation function for $\tau=0$; (e) cross-correlation function for values of τ .

tion are obviously the allowed frequencies of the orthonormal set, since there is no cross-correlation at these points.

Both of these terms are zero if

$$(\omega_m - \omega_n)T = k\pi \quad \text{for } k = 1, 2, \dots,$$

and

$$(\omega_m + \omega_n)T = l\pi \quad \text{for } l = 1, 2, \dots,$$

Defining $\omega_n = \omega_m + \Delta\omega$ and substituting $\Delta\omega T_0 = k\pi$ and $(2\omega_m + \Delta\omega)T = l\pi$, we find that, if $k=1$, then $\Delta\omega = \pi/T_0$ and

$$\Delta f = \frac{1}{2T_0}, \tag{28}$$

subject to the restriction that $\omega_m = l\pi/2T$.

The smallest allowable difference in frequency between any two members of a set of orthonormal frequencies is the reciprocal of twice the word length. It should be noted that the limit to the number of frequencies or eigenvalues that can be used in a set is usually dictated only by the limit on the complexity that can be tolerated.

Figure 11 is a plot of the cross-correlation function $\phi(0)$ as a function of ω_n for a fixed value of ω_m . The dashed curve is the amplitude spectrum of the pulsed sine wave. A practical example is a pulsed sine wave of 10 msec duration which has zeros of its cross-correlation with other

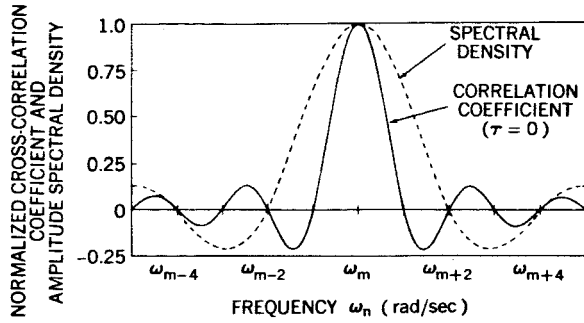


FIGURE 11.—Cross-correlation function and power spectrum of pulsed sine wave $f(t) = \sin(2\pi f_1 t) \cdot \{u[t + (T_0/2)] - [t - (T_0/2)]\}$.

frequencies in multiples of 50 cps; however, the zeros of the power spectrum of the pulsed sine wave fall in multiples of 100 cps away from the center frequency.

Amplitude spectral density curves for the general case when $f(t) = \sin(2\pi f_1 t + \theta) \cdot [u(t - \tau_1) - u(t - \tau_2)]$ are plotted in Figures 12 and 13. Figure 13 is varied in steps of 30 degrees for a total 180 degree change. The redistribution of densities can be seen.

Word-Error Probability for the Matched Filter

In calculating the word-error probability for the case of the matched filters the same procedure

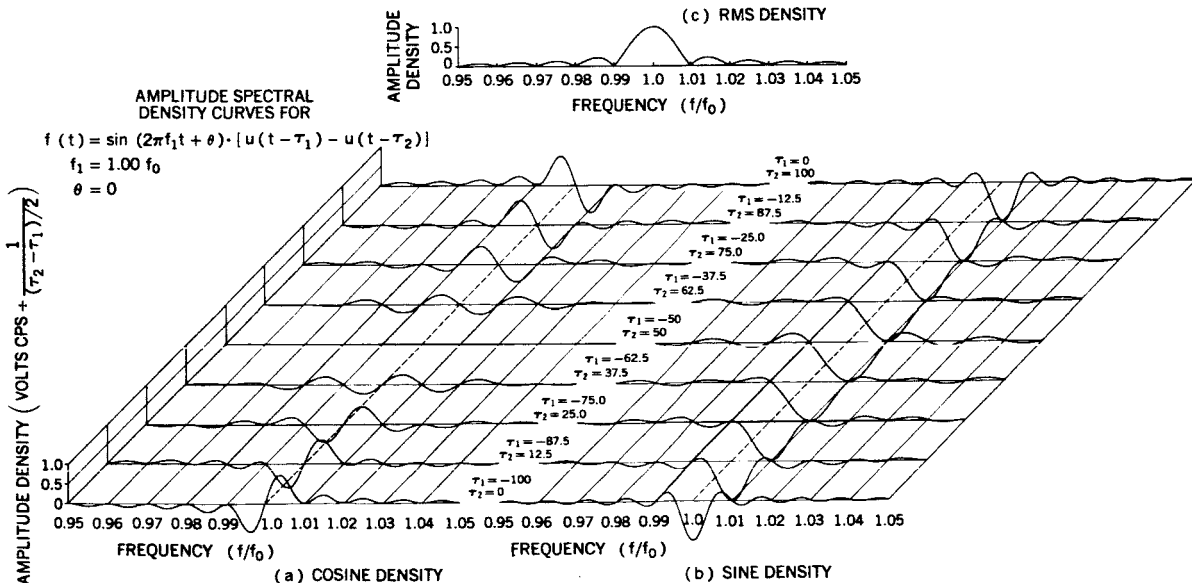


FIGURE 12.—Amplitude spectral density curves for pulsed sine waves as a function of time shifting.

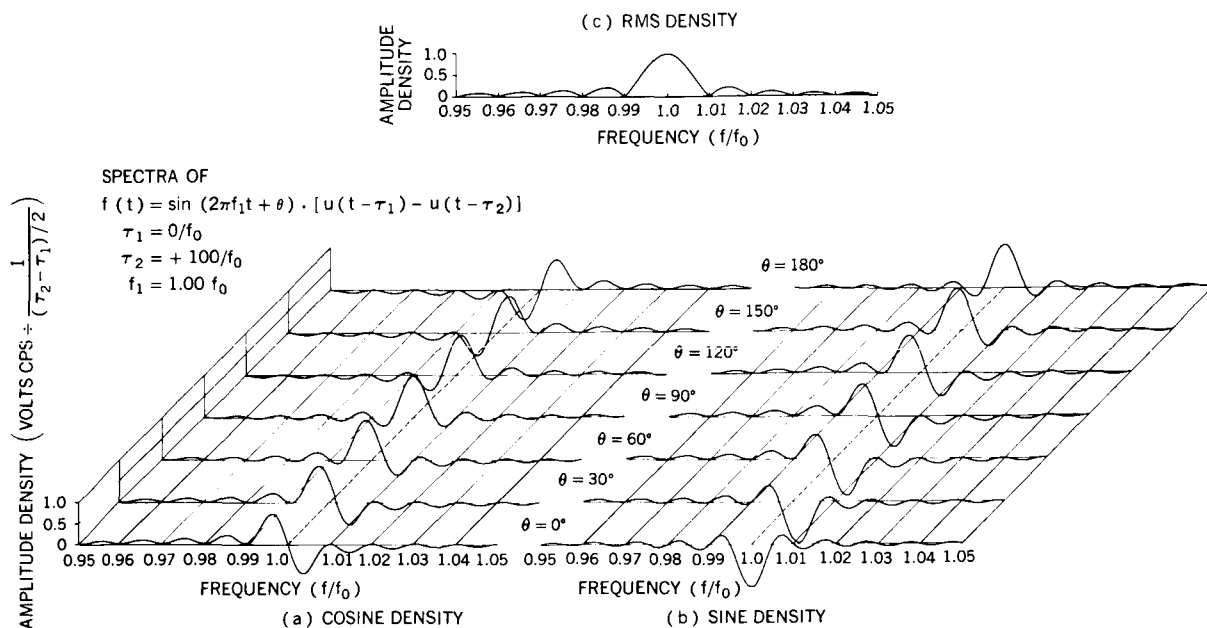


FIGURE 13.—Amplitude spectral density curves for pulsed sine waves as a function of phase shifting.

is followed as for the unmatched filters, except that the character of the noise is changed. This noise has a Gaussian probability density [Figure 14(a)]. The probability density function for a signal of amplitude C_0 and noise is shown in Figure 14(b). The probability that there is no signal with an amplitude greater than x_0 in a

has an amplitude greater than x_0 is

$$P_c(N) = \left(\frac{1}{\sqrt{2\pi} \sigma_m} \int_{-\infty}^{x_0} e^{-x^2/2\sigma_m^2} dx \right)^{N-1} \frac{1}{\sqrt{2\pi} \sigma_m} e^{-(x_0 - C_0)^2/2\sigma_m^2} dx_0, \quad (29)$$

where σ_m^2 = variance of noise at output of matched filter, and C_0 = amplitude of signal at matched filter output. The probability of being correct

$$P_e(N) = 1 - \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi} \sigma_m} \int_{-\infty}^{x_0} e^{-x^2/2\sigma_m^2} dx \right)^{N-1} \frac{1}{\sqrt{2\pi} \sigma_m} e^{-(x_0 - C_0)^2/2\sigma_m^2} dx_0. \quad (30)$$

These curves have been programmed for computation on an IBM 7090 digital computer, and the results are plotted in Figure 15 for different values of n , the \log_2 of the number of filters used.

CONTIGUOUS-FILTER DETECTION SYSTEM

A set of contiguous filters with a maximum-likelihood detector can enhance the signal-to-

filter is given by the integral of the Gaussian probability density function from $-\infty$ to x_0 . Raising this probability distribution function to the $N-1$ power gives the probability that is no signal greater than x_0 in any of the $N-1$ filters. The probability that there is a signal at x_0 in the correct filter and that none of the other filters

for any amplitude x_0 is this integral over all values of x_0 , from $-\infty$ to $+\infty$. The probability of an error is 1 minus the probability of being correct or

noise ratio of a PFM-encoded signal. The maximum-likelihood detector samples all filters and decides which filter contains the signal. This filter is connected to the output, so the signal-to-noise ratio is determined by the signal and noise in the one filter only. The output of the selected filter may be used in the following ways: (1) as an indicator that the signal was

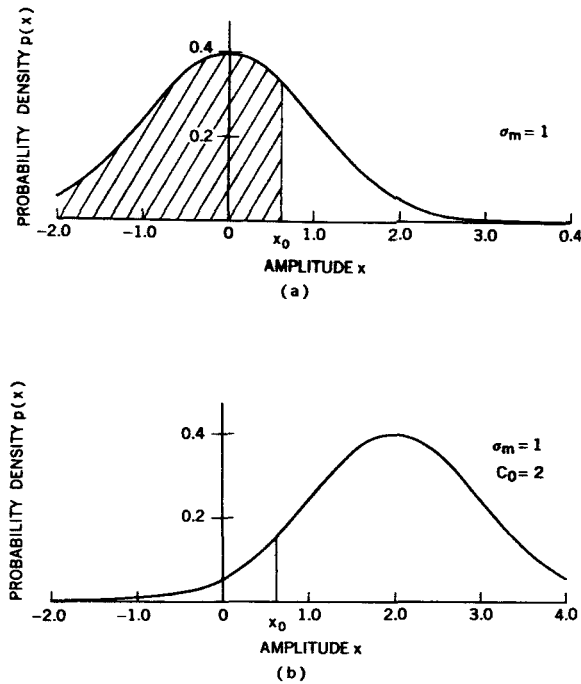


FIGURE 14.—Probability density functions: (a) Gaussian noise; (b) signal plus Gaussian noise.

present; with N filters, 2^n levels of the signal can be detected, or the PFM word gives a precision of n bits; (2) to measure the frequency on the output of the selected filter in order to achieve a precision better than one part in 2^n . The latter technique may be applied only to an unmatched-filter set.

Analysis of Unmatched Filters

Earlier the noise analysis of a bandpass filter excited with random noise was given, and the

$$e_0(t) = \frac{\Delta\omega_0 C_1}{(\Delta\omega_0^2 + 4\Delta\omega_1^2)^{1/2}} \left(1 + e^{-\Delta\omega_0 t} - 2e^{-\Delta\omega_0 t/2} \cos \Delta\omega_1 t \right)^{1/2} \sin(\omega_1 t + \phi_2 + \phi_3), \quad (33)$$

where

$$\phi_2 = \tan^{-1} \frac{2\Delta\omega_1}{\Delta\omega_0},$$

$$\phi_3 = \tan^{-1} \left(\frac{e^{-\Delta\omega_0 t/2} \sin \Delta\omega_1 t}{e^{-\Delta\omega_0 t/2} \cos \Delta\omega_1 t - 1} \right).$$

Equation 33 is derived in Appendix C.

It is instructive to examine the terms of this equation for transient effects during a frequency

error probability was calculated. It was assumed that the output of the filter contained no transients from the step input signal at the time of sampling, that the frequency of the input signal remained constant, and that an ideal filter was used. These assumptions will be considered in more detail now.

Response to Sinusoidal Step

The single-pole bandpass filter is the simplest and the most economic filter than can be used to make a bank of contiguous filters. Consider the filter circuit in Figure 16. The transfer function is given by:

$$\frac{E_0(s)}{E_{in}(s)} = \frac{\Delta\omega_0 s}{s^2 + \Delta\omega_0 s + \omega_0^2}, \quad (31)$$

where $\Delta\omega_0 = 1/RC$ and $\omega_0 = 1/\sqrt{LC}$. The location of the singularities of the expression are plotted in Figure 16. We are interested in the output of the filter when the input is excited with a step sinusoid,

$$f(t) = C_1 \sin \omega_1 t \cdot u(t). \quad (32)$$

A convenient substitution is to let $\omega_0 = \omega_1 + \Delta\omega_1$ so that Equation 31 is a function of the excitation frequency and the difference between the excitation frequency and the resonant frequency of the filter. The equation for the response of a single-pole filter excited with a sinusoidal input of amplitude C_1 and frequency ω_1 is:

measurement. One method of frequency measurement is to determine the period of each cycle and find a k period average. The averaging is started at least 1 time constant, $1/\Delta\omega_0$ seconds, away from the leading edge of the pulse and may be continued until the end of the pulse. The value ϕ_2 will have no effect upon the measurement since it represents a constant phase shift of the zero crossings; however, ϕ_3 definitely contributes to the frequency measurement error if $e^{-\Delta\omega_0 t/2}$ is not damped appreciably. Figure 17 is a plot of $\phi_2 + \phi_3$ against time, for values of $\Delta\omega_1$ ranging

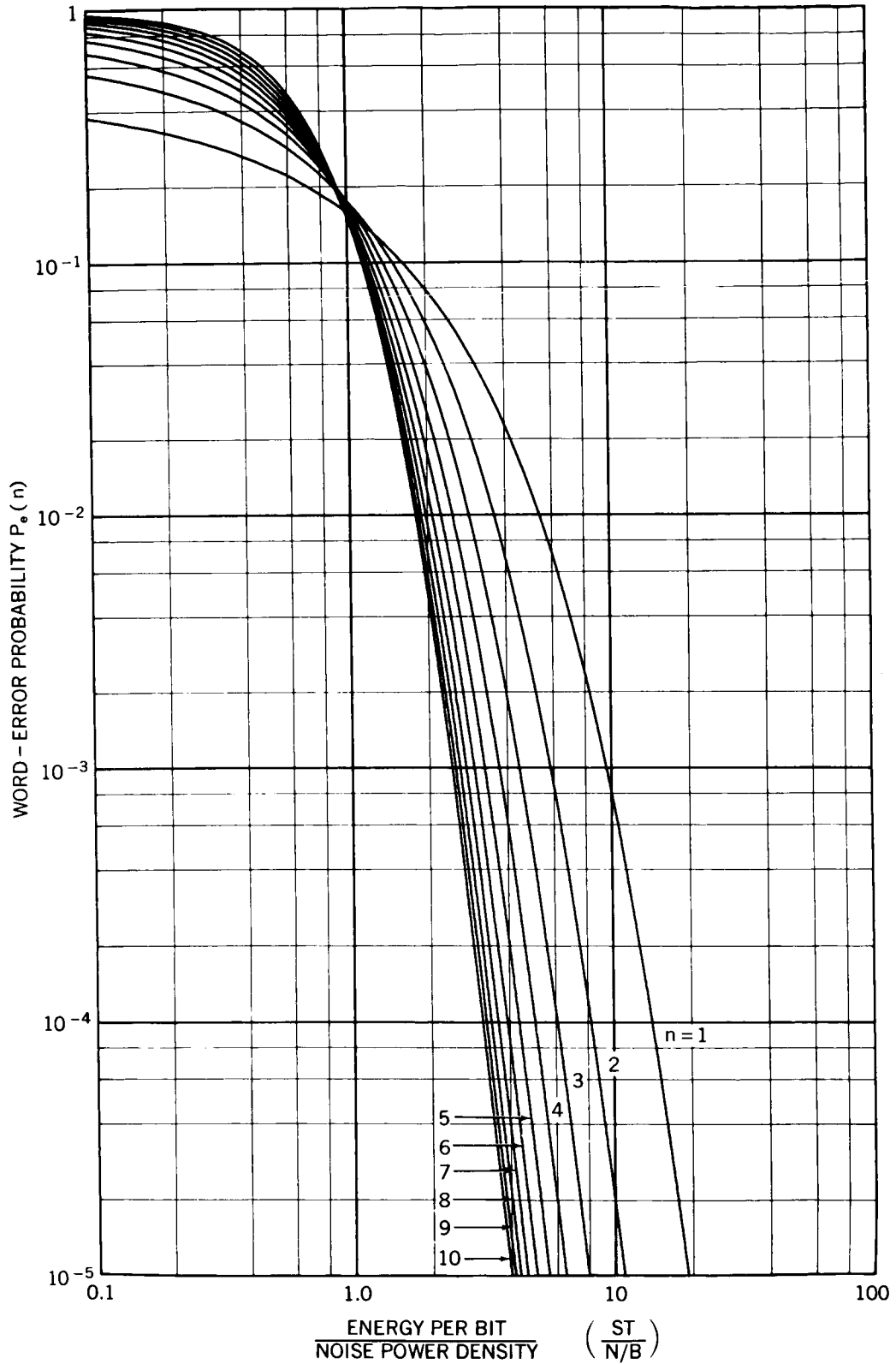


FIGURE 15.—Word-error probability curves for Gaussian noise.

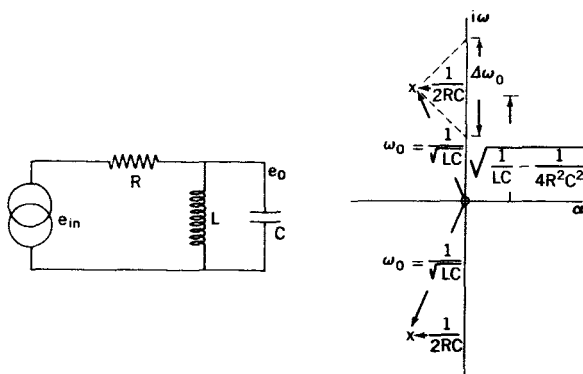


FIGURE 16.—Filter element circuit.

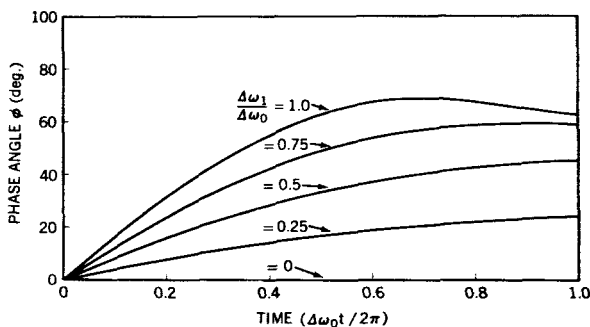


FIGURE 17.—Phase response step sinusoid for a single-pole filter.

from the frequency of the center of the adjacent filter to the frequency of the center of the filter under investigation. From these curves and the start and stop times for the frequency measure-

convolution of the excitation function with the impulsive response of the single-pole filter is

$$\begin{aligned}
 h(t) = & \Delta\omega_0 A_1 e^{-\Delta\omega_0 t/2} \left[\cos\left(\omega_0 t + \tan^{-1} \frac{\Delta\omega_0}{2\omega_0}\right) \cdot \right. \\
 & \cdot \int_0^t e^{\Delta\omega_0 \tau/2} \cos \omega_0 \tau \cdot \sin\left(\omega_1 \tau - \frac{\omega_d}{\omega_s} \cos \omega_s \tau\right) d\tau \\
 & \left. + \sin\left(\omega_0 t + \tan^{-1} \frac{\Delta\omega_0}{2\omega_0}\right) \cdot \int_0^t e^{\Delta\omega_0 \tau/2} \sin \omega_0 \tau \cdot \sin\left(\omega_1 \tau - \frac{\omega_d}{\omega_s} \cos \omega_s \tau\right) d\tau \right]. \quad (34)
 \end{aligned}$$

where ω_s is the peak frequency deviation and ω_d is the modulation frequency.

ment, the error in frequency can be calculated (Figure 18).

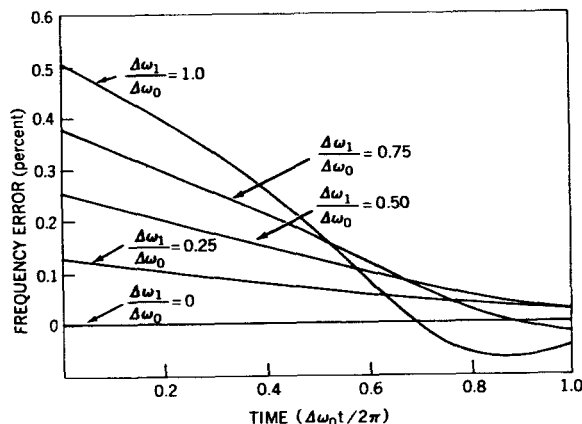


FIGURE 18.—Frequency error from response of a single-pole filter to a step sinusoid.

Response to a Gliding Tone

The case for the PFM excitation frequency not centered in the filter has been examined. Of considerable interest in the case for the signal which is changing frequency while it is in the filter. This frequency change is very evident when sample-and-hold techniques are not used in the analog oscillators. A spin-stabilized spacecraft can frequency modulate any vector-type measurement at the spin frequency.

The solution for the continuous gliding tone has been given by Barber and Ursell (Reference 17). In Appendix D the equation for a step-function changing frequency is derived. The

If the rate of change of frequency is too great, the filter does not have sufficient time to respond,

and the output is attenuated. Equation 34 has been programmed on an IBM 7090 computer to determine the attenuation for various rates of frequency change. The parameters for the Explorer X space probe telemeter were used to

compute an amplitude-frequency-time plot. A three-dimensional model was constructed to show the interesting results of the beating of the transient response and the gliding excitation function; a photograph of the model is shown in Figure 19.

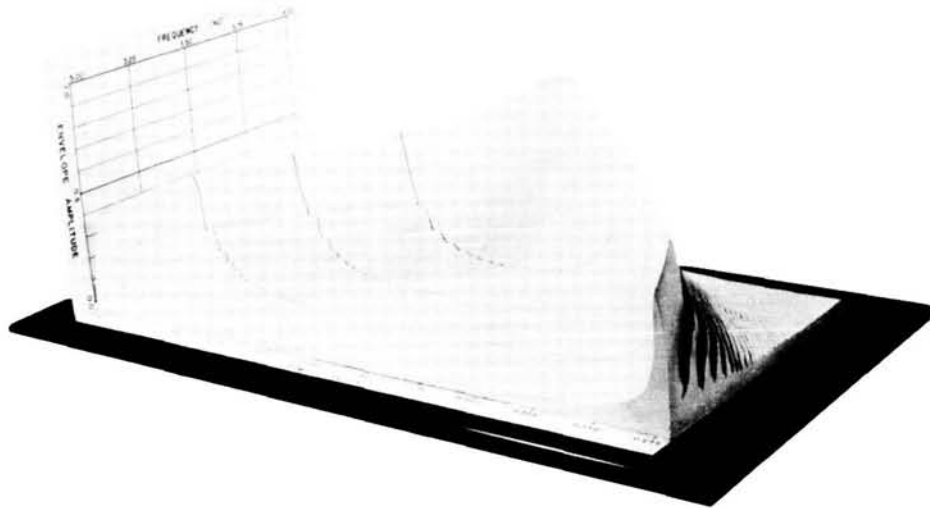


FIGURE 19.—Amplitude-frequency-time function of the response of a single-pole filter to a gliding tone.

In this figure time is the x-axis coordinate, and the scale along the x-axis depicts half of a revolution of the spacecraft. The y-axis is in units of the center frequency of a filter $\Delta\omega_0$ radians per second in width. The z-axis is the amplitude of the envelope of the output of a filter. Parameters for Explorer X used in the calculation were: a bandwidth of 25 cps for a filter, 100 filters covering the frequency range from 1.75 to 4.25 kc; and a sinusoidal varying frequency centered at 3 kc, with a maximum deviation of 1 kc for a spin frequency $\omega_a/2\pi$ of 1.8 cps. The responses of 17 filters distributed between 3 and 4 kc were calculated and plotted. These were used as templates in the construction of the model. Any cross-section parallel to the time axis gives the response of the filter of that frequency to the gliding tone (Figure 20).

Matched-Filter Techniques

Matched filters used with a maximum-likelihood-detector system provide the optimum in

signal detection. These techniques can be applied to PFM-encoded signals if phase coherence is maintained at the beginning of each pulse.

It can be noted in Figure 10 that the signal and stored image are phase coherent at the beginning of the pulse. For perfect orthogonality, only those frequencies of the signal are allowed which cause an integral number of half cycles in the pulse. Since the phase of analog PFM signals can have any value and the signals are therefore not always orthogonal, it is not possible to use the matched-filter techniques directly. A compromise which allows the use of these techniques is to force the phase of the pulsed subcarrier oscillator to start always in the same phase at *a priori* known times. The phase at the end of the pulse is, of course, unknown. But, if the frequency separation of the stored image is governed by the relation $\Delta\omega = \pi/T_0$ (Equation 28), where $\Delta\omega$ is the angular frequency separation and T_0 is the length of the pulse, then the correlation coefficient for one of the N correlators will always be at least

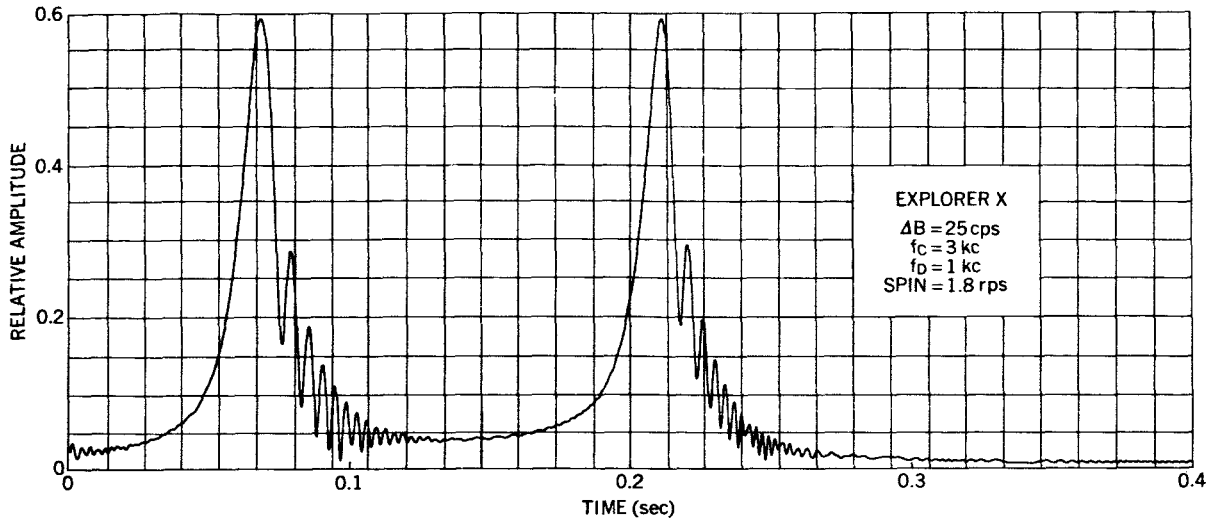


FIGURE 20.—Response of one single-pole filter to a gliding tone.

$1/\sqrt{2}$ of the maximum. However, when this occurs, it will occur in two correlators simultaneously. This area has not, as yet, been investigated thoroughly, but the technique should be applicable to the telemetry signals received from Ariel I. A tape recorder aboard the satellite played back signals of stored PFM analog and digital oscillator pulses. Because of the ratio of record to playback speeds, the frequencies of the oscillators were scaled down by a factor of 48 to 1. This forced each pulse to start in the same phase and should allow application of the matched-filter technique to data already received.

EXPERIMENTAL RESULTS

Spectral Analysis

The operation of a set of contiguous filters combined with a maximum-likelihood detector can become quite complicated, especially when the signal is derived from magnetic tape recordings of actual satellite passes. Since output is influenced by many variables, evaluation in conditions of poor signal-to-noise ratios is difficult. To obtain a more qualitative estimate of the value of the contiguous-filter system, an experimental system was set up so that measurements under actual noisy conditions could be made.

The equipment consisted of a tape scanner which performs the same task as an endless loop tape recorder. The tape scanner has a seven-track reproducing head mounted on a revolving

1-foot-diameter disk. The tape is in contact with half of the circumference of the disk so that, with the tape motionless and the disk rotating, the head sweeps along the tape for half of each revolution of the disk. The peripheral speed of the disk is 30 ips. Thus, it is possible to scan one portion of a magnetic tape recording many times and observe the output on an oscilloscope.

By inserting a wave analyzer between the tape scanner and the oscilloscope, the spectral content of the signal for any pulse duration T and bandwidth $\Delta\omega_0$ can be obtained. A Quan-Tech Model 303 Wave and Noise Spectrum Analyzer set at a bandwidth of 100 cps was used to obtain the envelope of the signal and noise in the 100 cps bandwidth over the frequency range of 3.6 to 16.4 kc. An important advantage of the Model 303 Analyzer is the flat-topped filter characteristic, which is a Butterworth approximation. It provides, for this purpose, a good approximation to an ideal filter.

By very slowly changing the center frequency of the analyzer as a section of tape is scanned, a spectral analysis of the signal can be made. The output of the analyzer is used to intensity-modulate the beam of the oscilloscope as the sweep is being triggered at the beginning of each revolution of the scanner. The analyzer has a dc output which is proportional to the position of the frequency dial; this voltage is used on the y-axis of the oscilloscope in the same manner as the

vertical deflection on a television raster. By taking a time exposure of the oscilloscope face, a high quality spectral analysis of the signal is produced. The x-axis represents time, the y-axis is a function of the frequency, and the intensity represents the energy in a filter at a particular time.

An interesting aspect to the problem developed as this equipment was being checked. Explorer XII, the Energetic Particles Satellite, suddenly ceased transmitting after approximately 2000 hours of continuous operation. The signal was last heard over Woomera, Australia, at a range of

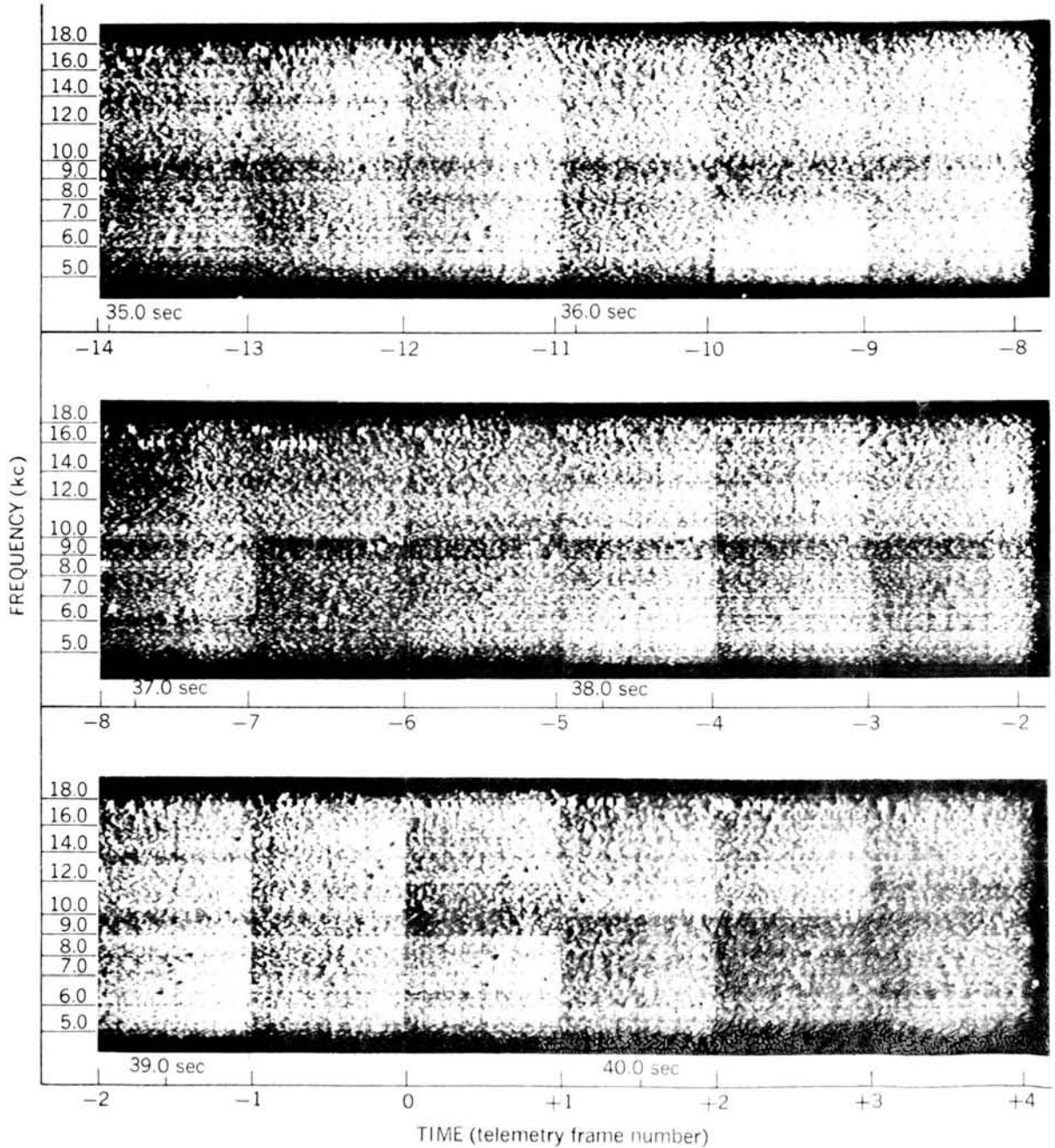


FIGURE 21.—Spectral analysis of PFM telemetry signals.

65,000 km. It was highly desirable to examine the magnetic tape recordings at the time of failure to see if any clues could be obtained on the possible cause of failure. Some of the equipment aboard Explorer XII was identical to that to be used in the next scheduled spacecraft

so that, if the trouble area were located, corrective measures could be taken before flight.

A copy of the Woomera tape was played on the tape scanner into the wave analyzer. Close examination of the output at the point where the signal apparently disappeared revealed that the

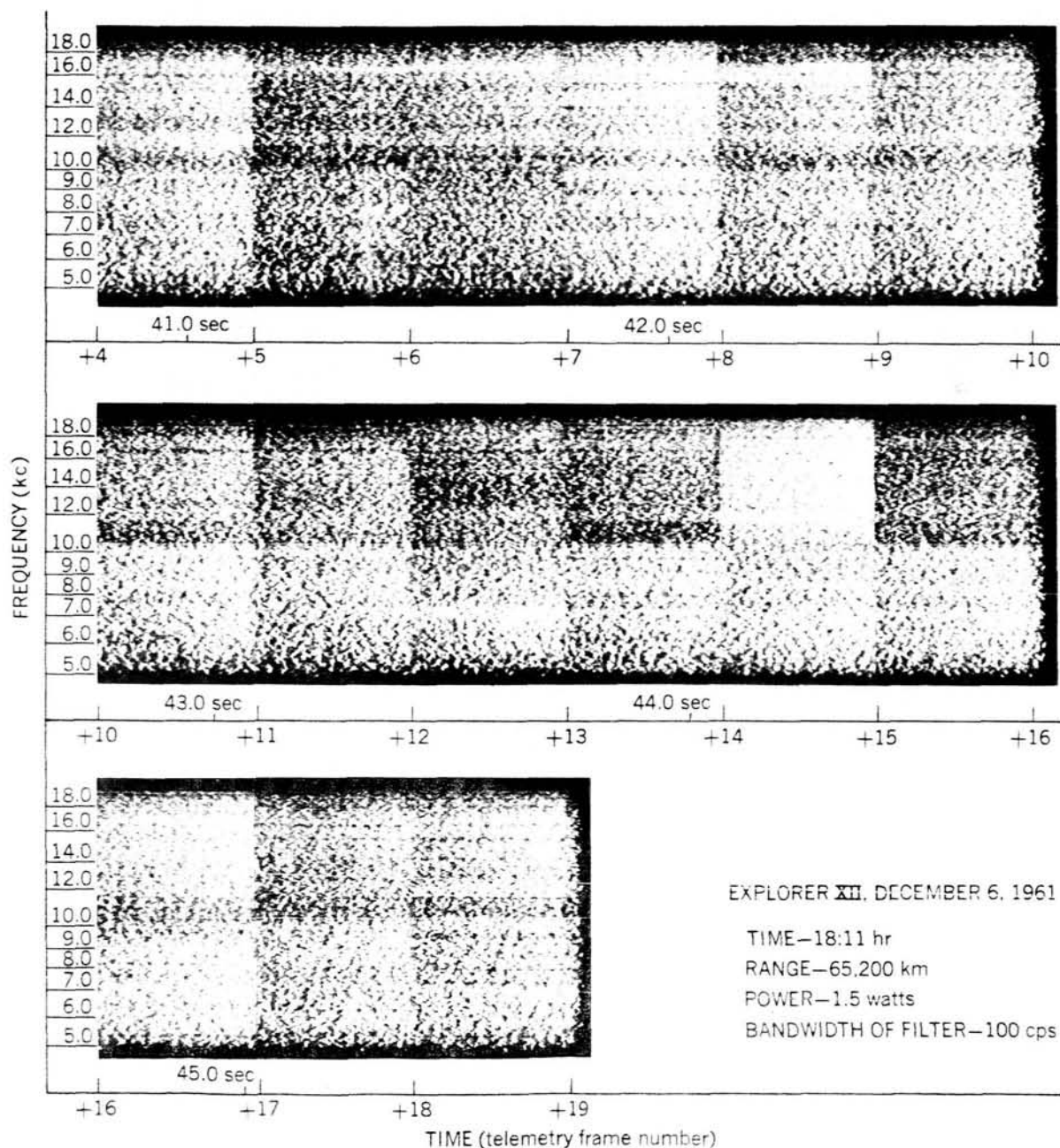


FIGURE 21 (continued).—Spectral analysis of PFM telemetry signals.

signal was still present although attenuated by about 10 to 15 db. The signal remained at this level for 2 seconds and then disappeared completely. During these last 2 seconds, seven frames of the telemetry were received. Although the signal-to-noise ratio was poor during this interval, the majority of the channels could be read. The tracking filter used at the ground station was swinging in and out of lock because of the large amount of noise in the feedback loop. This caused a random obliteration of some channels.

To demonstrate the noise-immunity characteristics of PFM, the complete failure sequence is shown in Figure 21. Two spectral analysis photographs were taken of each telemetry frame; one covered the frequency range from 4.5 to 10 kc and the other from 10.0 to 18.0 kc. Composite photographs were then made up from the twelve photographs of each six frames. A set of these cover the time span from fourteen frames before the signal disappearance to nineteen frames afterward. In order to maintain high resolution along the frequency axis, the frequency dial was motor driven at a rate sufficient to yield at least 300 horizontal lines per photograph; this gives an effective 600 lines in the vertical dimension.

The failure sequence started between the ninth and the tenth channel pulses in telemetry frame zero. The increase in noise is immediately apparent since the tracking filter at the ground receiving station increased its gain to compensate for the loss in signal; however, evidence of signal energy in channels one through six is very definite in telemetry frame one. Because of the redundant nature of some of the data, the frequencies of these channels could be predicted a few frames in advance. With this *a priori* information, it was easy to establish the existence of the signal for the 2-second period after the signal disappeared. During this time all channels appeared to be giving their normal indications. This cleared the majority of the spacecraft subsystems from suspect and left only the transmitter and transmitter power converter to be investigated further.

The instantaneous decrease in output power of approximately 15 db could be attributed to either a failure in the transmitter output stage or a loss of power from the transmitter converter. There

was enough interelectrode capacitance in the transmitting tube to allow a feed-through of energy from the transmitter oscillator-buffer stage to the antenna. Because of the catastrophic failure at the end of the 2-second attenuation period, it is probable that the failure occurred in the transmitter converter. Subsequent flights were scheduled to use all-transistorized transmitters; therefore the potential trouble sources would be eliminated.

This method of spectral analysis is not extremely effective for presenting noisy data. The number of intensity levels that the human eye can distinguish is fairly limited. To aid in increasing the contrast between signal and noise, the output of the wave analyzer was passed through an RC network which added a small amount of the derivative of the input to its output. The objective was to give the picture more of a three-dimensional look by casting "shadows" on one side of each PFM pulse. The effect is very prominent on the original photographs, although some of the effect is lost in the reproduction of these figures. The block diagram of the equipment used in making the photographs is shown in Figure 22, which also includes the circuit diagram of the network giving the three-dimensional effect.

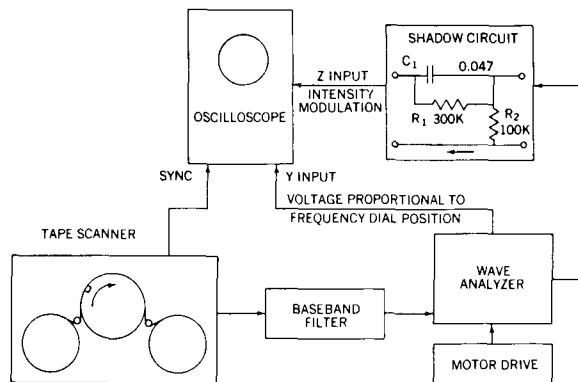


FIGURE 22.—Block diagram of spectral analysis equipment.

This same technique of spectral analysis was applied to the recording of the amplitude modulation output of the tracking filter; the results are shown in Figure 23. Before the failure there was a strong 50 cps component, the phase of which was used to provide channel synchronization. The third and fifth harmonics of this frequency

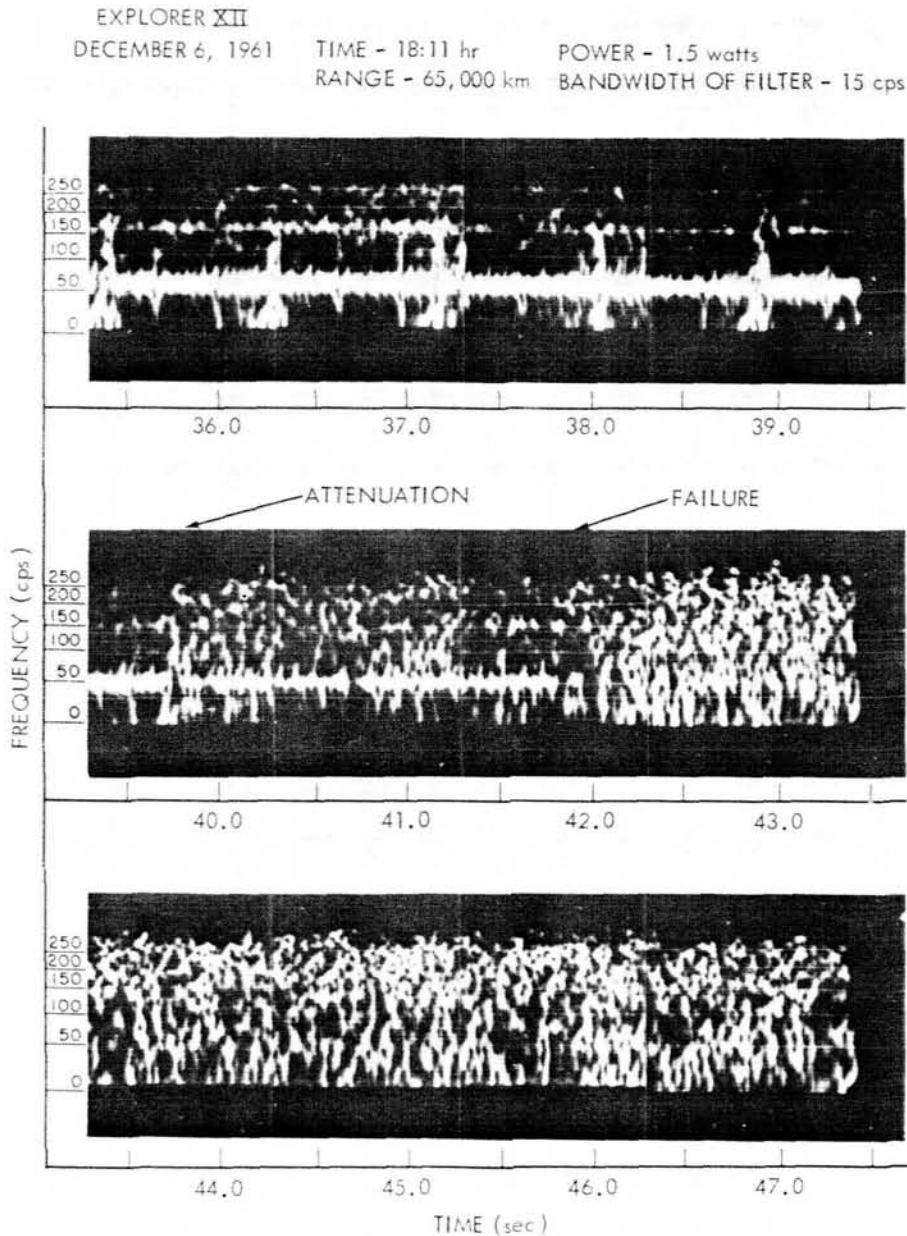


FIGURE 23.—Spectral analysis of telemetry receiver AM output during failure sequence.

can also be seen. The series of impulses spaced slightly less than 1 second apart are a result of the shadowing of the solar cells by their supporting paddles. This introduced a transient into the power system and caused a slight pulling of the transmitter frequency. At 39.75 seconds the intensity decreases; the signal completely dis-

appears at 41.8 seconds. These composite photographs were taken without the RC network since only continuous signals were being observed.

Signal-to-Noise-Ratio Comparisons

The Woomera tape provided an excellent opportunity to compare theoretical calculations

of the signal-to-noise ratio and the error probability with actual measurements. The calculations are made as follows:

1. Determine the received signal power:

The measured prelaunch transmitter power was 1.5 watts. The sideband power for a ± 57 degree phase-modulated signal is twice the carrier power or 1.0 watt ($P_t = +30.0$ dbm). The transmitting antenna gain includes the dipole gain, wiring harness loss, and circular polarization loss ($G_t = -4$ db). The receiving antenna gain is taken from quoted values of gain for an array of 16 crossed-yagi antennas ($G_r = +21$ db). Attenuation due to the 63,200 km path loss is

$$\left(\frac{\lambda}{4\pi r}\right)^2 = \left(\frac{300}{136 \times 4\pi \times 63.2 \times 10^6}\right)^2$$

$$= -171.1 \text{ db}.$$

Substitution of these factors in the received power equation yields:

$$W_r = P_t G_t G_r \left(\frac{\lambda}{4\pi r}\right)^2$$

$$= +30.0 - 4 + 21 - 171.1$$

$$= -124.5 \text{ dbm}.$$

2. Determine the received noise power

The galactic noise temperature (Reference 18) at 136 Mc in the plane of the ecliptic has an average value of about 600°K ; however, there is a hot spot of about 2000°K looking toward the center of the galaxy. The receiver noise temperature is quoted at 3 db. This corresponds to a receiver noise temperature of 290°K . The noise temperature due to the earth and atmospheric noise as sensed by the antenna side lobes is 55°K . Since synchronous detection was performed on the carrier by the station tracking filters, the noise bandwidth can be taken to be the width of the individual filter in the contiguous-filter bank employed in data reduction. This bandwidth is

100 cps. Therefore the received noise power is

$$P_n = kTB \text{ watts}$$

$$= 1.38 \times 10^{-23} \times 945 \times 100$$

$$= 1.305 \times 10^{-15} \text{ mw}$$

$$= 148.9 \text{ dbm},$$

where

$$k = 1.38 \times 10^{-23} \text{ watt-sec/deg},$$

$$T = 600^\circ + 290^\circ + 55^\circ = 945^\circ\text{K}.$$

$$B = 100 \text{ cps}.$$

3. Determine the signal-to-noise ratio

The calculated signal-to-noise ratio is

$$\frac{W_r}{P_n} = -124.5 - (-148.9)$$

$$= 24.4 \text{ db}.$$

A reasonably accurate signal-to-noise-ratio measurement can be taken from Figure 24. The x-axis encompasses one telemetry frame, and each trace in the y direction represents a change of 50 cps in the filter center frequency. The signal amplitude can be measured fairly accurately, but a good qualitative measurement of the noise is more difficult. The voltage signal-to-noise ratio is estimated to fall between 10 and 15 to 1 or between 20 and 23.5 db. This compares very favorably with the calculated value of 24 db; however, it should not be too close, since we have calculated the input signal-to-noise ratio and assumed that it is the same as the output ratio. But the actual measurement is, on the other hand, purely the output signal-to-noise ratio.

A more meaningful measurement is the word-error probability. An examination of the data print-out records taken on a typical operational data-reduction line, which uses a bank of contiguous filters, indicated that the word-error probability just prior to the failure sequence was 10^{-2} .

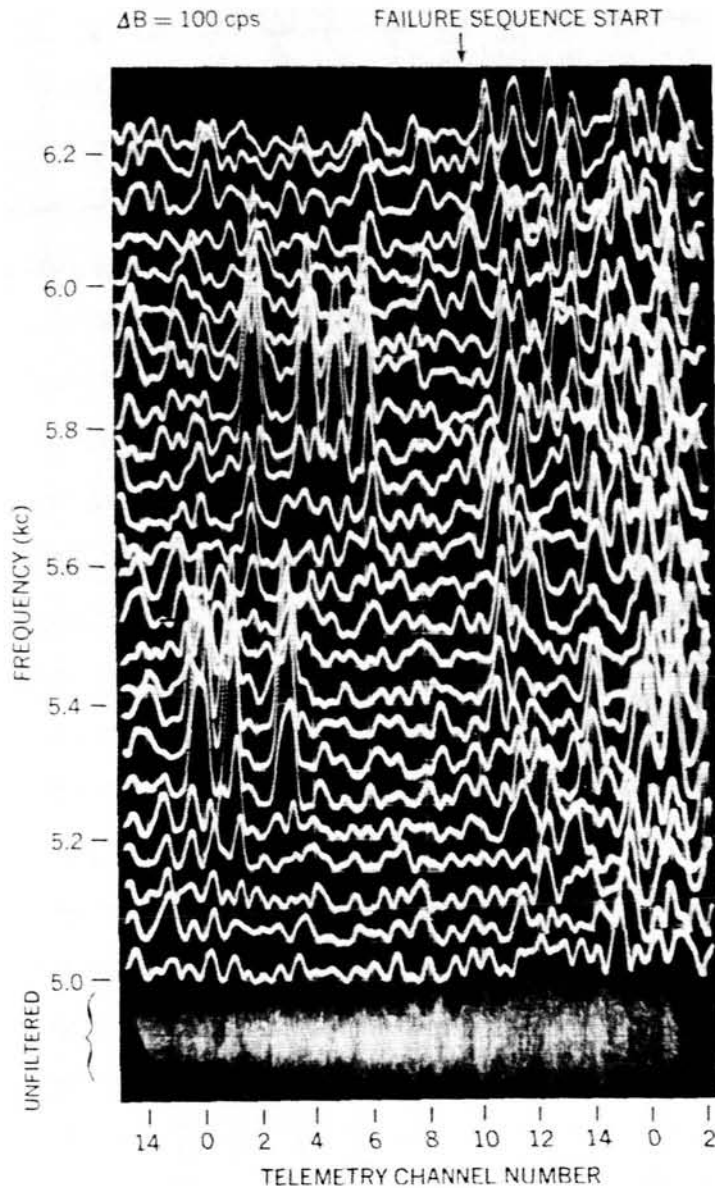


FIGURE 24.—Spectral analysis of telemetry frame of Explorer XII at time of failure.

If the word-error probability curves for Rayleigh noise of Figure 8 are used, the error probability for 24 db, after adjustment for the number of bits in the word and by using the total variance, is still greater than 10^{-5} . The operational contiguous filters have a threshold detector which must be exceeded before data are printed out. The actual errors in the operational data reflect the effects of the threshold detector and are

not representative of theoretical filters operating with maximum-likelihood detectors.

CONCLUSION

The object of this report has been twofold: to present some of the characteristics of pulse-frequency modulation; and to show that pulse-frequency modulation, with certain restrictions, is a special PCM group code of coded binary

sequences having orthonormal characteristics.

A number of successful spacecraft have used pulse-frequency modulation as the encoding technique. In the decoding of their analog data, a set of contiguous filters was used to extract the magnitude of the original signal. Since the individual filters were not flat within their passbands, degradation of the signal-to-noise ratio occurred if the encoded signal did not fall exactly in the center of a filter. Also, for response-time considerations, the filter bandwidth was equated to the reciprocal of the pulse length. By employing cross-correlation techniques using maximum-likelihood detectors, the resolution of the measurement can be doubled for the same pulse length without increasing the baseband width. This emphasizes an outstanding characteristic of pulse-frequency modulation: Quantizing errors for analog signals are not forced into the data at the source; the quantization is accomplished during the reduction of the data. The signals are encoded in a manner which approaches the maximum communication efficiency; the amount of enhancement of the signal-to-noise ratio depends on the sophistication of the decoding process.

The method of presenting the cross-correlation coefficients (when $\tau=0$) of binary sequences in a matrix array aids in visualizing the communication efficiency that can be expected for any particular grouping of the sequences. The method ties together the efficiency of group and parity codes, and the same techniques can be applied to other codes, such as the Bose-Chaudhuri code, which have more members per set but are less efficient (Reference 19). The correlation tables for any of these coding schemes can be generated conveniently by means of a digital computer.

Error-probability curves have been derived for unmatched-filter set, with maximum-likelihood detection, perturbed by Rayleigh noise. The matched-filter case has also been given so that direct comparisons may be made. With large signal-to-noise ratios, the probability of a word error as given by the Rayleigh probability density function approaches the error probability for the matched filter given by the normal probability density function. At lower signal-to-noise ratios the matched filter is appreciably better than the unmatched; however, at these low values, the

high error probability usually renders the data useless.

The output of a single-pole bandpass filter in a set of contiguous unmatched filters, can be fed into a discriminator to determine the frequency more accurately. The transient terms generated by the excitation of the filter with a step-function sinusoid introduce a small error in the frequency measurement of the pulse. The error is a function of the displacement of the pulse frequency from the center frequency of the filter, although its magnitude may be neglected in most cases.

Calculations were presented for the response of a single-pole filter to a sinusoidally varying frequency. The beating of the transient response of the filter with the excitation frequency was illustrated in the model of the amplitude-frequency-time relation with parameters for a typical spacecraft.

Nothing has been said about the methods of implementation of the encoding process. These methods are now fairly well established as practical engineering techniques. Instead, the emphasis has been on the versatility of pulse-frequency modulation as an encoding technique. It is hoped that this work will stimulate effort to provide as much versatility in signal decoding as now exists in the encoding process.

ACKNOWLEDGMENTS

It is with sincere appreciation that I acknowledge the contributions of my associates at Goddard Space Flight Center who have aided in the successful application of pulse-frequency modulation to satellite communication. I would like to thank Mr. P. T. Cole, who designed and developed the tape scanner used to extract signals buried in noise. I would also like to mention Mr. David H. Schaefer for his efforts on the digital pulsed subcarrier oscillator; Mr. Roland L. Van Allen for his original work on the analog pulsed subcarrier oscillator; and Mr. Clarence B. House for his work on special devices which enhance the applicability of PFM; Mr. J. H. Perry for his very careful efforts in the layout and construction of the first printed circuit modules for the Vanguard III Satellite; Mr. Hosea D. White for his tireless effort in reducing PFM circuitry ideas to practical application; and Mr. Curtis M. Stout for his outstanding contributions

to the solution of PFM data-reduction problems. I am also grateful to Dr. Norman F. Ness who introduced me to the concept of the large scale digital computer as an everyday engineering tool.

REFERENCES

1. SHANNON, C. E., "A Mathematical Theory of Communication," *Bell System Tech. J.* 27(3):379-423, July 1948; 27(4):623-656, October 1948.
2. KOTEL'NIKOV, V. A., "The Theory of Optimum Noise Immunity," New York: McGraw-Hill, 1959, pp. 46-48.
3. CARSON, J. R., "Notes on the Theory of Modulation," *Proc. IRE* 10(1):57-64, February 1922.
4. ARMSTRONG, E. H., "A Method of Reducing Disturbances in Radio Signaling by a System of Frequency Modulation," *Proc. IRE* 24(5):689-740, May 1936.
5. GOODALL, W. M., "Telephony by Pulse Code Modulation," *Bell System Tech. J.* 26(3):395-409, July 1947.
6. HAMMING, R. W., "Error Detecting and Error Correcting Codes," *Bell System Tech. J.* 29(2):147-160, April 1950.
7. VITERBI, A. J., "On Coded Phase-Coherent Communications," *IRE Trans. on Space Electronics and Telemetry* SET-7(1):3-14, March 1961.
8. ROCHELLE, R. W., "Earth Satellite Telemetry Coding System," *Elec. Eng.* 76(12):1062-1065, December 1957.
9. LINDEN, D. A., and ABRAMSON, N. M., "A Generalization of the Sampling Theorem," *Info. and Control* 3(1):26-31, March 1960.
10. REED, I. S., "A Class of Multiple-Error-Correcting Codes and the Decoding Scheme," *Trans. IRE Prof. Group on Information Theory* PGIT-4:38-49, September 1954.
11. SANDERS, R. W., "Digilock Telemetry System," in: *Proc. Nat. Sympos. on Space Electronics and Telemetry, San Francisco, September 1959*, New York: IRE Prof. Group on Space Electronics and Telemetry, 1959, Paper 6.3.
12. NORTH, D. O., "An Analysis of the Factors Which Determine Signal/Noise Discrimination in Pulsed Carrier Systems," Radio Corp. Amer. Lab. Rept. PTR-6C, June 25, 1943.
13. HELSTROM, C. W., "Statistical Theory of Signal Detection," New York: Pergamon Press, 1960, pp. 91-95.
14. CAMPBELL, G. A., and FOSTER, R. M., "Fourier Integrals for Practical Applications," Princeton, N. J.: Van Nostrand, 1948.
15. RICE, S. O., "Mathematical Analysis of Random Noise," *Bell System Tech. J.* 23(3):282-332, July 1944; 24(1):46-156, January 1945.
16. WAINSTEIN, L. A., and ZUBAKOV, V. D., "Extraction of Signals from Noise," Englewood Cliffs, N.J.: Prentice-Hall, 1962, pp. 86-91.
17. BARBER, N. F., and URSELL, F., "The Response of a Resonant System to a Gliding Tone," *Phil. Mag.* 39(292):345-361, May 1948.
18. KO, H. C., "The Distribution of Cosmic Radio Background Radiation," *Proc. IRE* 46(1):208-215, January 1958.
19. BOSE, R. C., and RAY-CHAUDHURI, D. K., "On a Class of Error Correcting Binary Group Codes," *Info. and Control* 3(1):68-79, March 1960.

APPENDIX A

AMPLITUDE DENSITY SPECTRUM OF A PULSED SINE WAVE

Let $f(t)$ be a pulsed nonrecurring sine wave of frequency ω_1 , phase angle θ , and pulse length $\tau_2 - \tau_1$ (Figure A1). From the Fourier transform

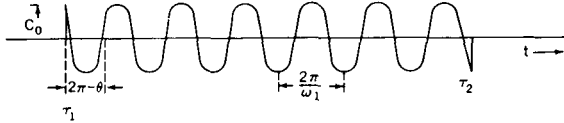


FIGURE A1.—Pulsed sine wave.

of $f(t)$ the amplitude spectrum of a pulsed sine wave is

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$$

$$= \int_{\tau_1}^{\tau_2} C_0 \sin(\omega_1 t + \theta) e^{-i\omega t} dt. \quad (A1)$$

By integrating and substituting limits,

$$F(\omega) = -\frac{C_0}{2} \left\{ \frac{\cos [(\omega_1 - \omega)\tau_2 + \theta] - \cos [(\omega_1 - \omega)\tau_1 + \theta]}{\omega_1 - \omega} \right.$$

$$+ \frac{\cos [(\omega_1 + \omega)\tau_2 + \theta] - \cos [(\omega_1 + \omega)\tau_1 + \theta]}{\omega_1 + \omega}$$

$$+ i \frac{\sin [(\omega_1 - \omega)\tau_2 + \theta] - \sin [(\omega_1 - \omega)\tau_1 + \theta]}{\omega_1 - \omega}$$

$$\left. - i \frac{\sin [(\omega_1 + \omega)\tau_2 + \theta] - \sin [(\omega_1 + \omega)\tau_1 + \theta]}{\omega_1 + \omega} \right\} \quad (A2)$$

This represents the general expression for the amplitude density spectrum of a pulsed sine wave of arbitrary phase and duration. The classic

example is the case of a sine wave turned on at $-\tau/2$ and off at $+\tau/2$ with $\theta=0$. The expression for this reduces to:

$$F(\omega) = \frac{C_0 \tau}{2} \left[\frac{\sin(\omega_1 - \omega) \frac{\tau}{2}}{(\omega_1 - \omega) \frac{\tau}{2}} - \frac{\sin(\omega_1 + \omega) \frac{\tau}{2}}{(\omega_1 + \omega) \frac{\tau}{2}} \right]. \quad (A3)$$

APPENDIX B

AUTOCORRELATION FUNCTION FOR GAUSSIAN NOISE

With white additive Gaussian noise as the input of bandwidth $\Delta\omega$ radians/sec, the autocorrelation to an ideal filter centered at ω_0 radians/sec and of the output becomes:

$$\phi_{11}(\tau) = \frac{1}{2\pi} \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} \frac{N}{2} e^{i\omega\tau} d\omega + \frac{1}{2\pi} \int_{-(\omega_0 + \frac{\Delta\omega}{2})}^{-(\omega_0 - \frac{\Delta\omega}{2})} \frac{N}{2} e^{i\omega\tau} d\omega, \quad (\text{B1})$$

where N is the noise power density in watts per cps. Integrating and combining yields

$$\phi_{11}(\tau) = \frac{N}{2\pi} \left[\frac{\sin\left(\omega_0 + \frac{\Delta\omega}{2}\right)\tau}{\tau} - \frac{\sin\left(\omega_0 - \frac{\Delta\omega}{2}\right)\tau}{\tau} \right]. \quad (\text{B2})$$

But

$$\sin(A + B) - \sin(A - B) = 2 \cos A \sin B.$$

Therefore

$$\phi_{11}(\tau) = \frac{N \Delta\omega}{2\pi} \frac{\sin \frac{\Delta\omega \tau}{2}}{\frac{\Delta\omega \tau}{2}} \cos \omega_0 \tau. \quad (\text{B3})$$

APPENDIX C

RESPONSE OF A SINGLE-POLE FILTER TO A STEP SINUSOID

The transfer function (Laplace transform) of the circuit in Figure 16 is:

$$\begin{aligned} \frac{E_o(s)}{E_{in}(s)} &= \frac{\frac{1}{R}}{Cs + \frac{1}{R} + \frac{1}{Ls}} \\ &= \frac{\frac{s}{RC}}{s^2 + \frac{1}{RC}s + \frac{1}{LC}} \end{aligned} \quad (C1)$$

Substituting $\Delta\omega_0 = 1/RC$ and $\omega_0 = 1/\sqrt{LC}$ yields

$$\frac{E_o(s)}{E_{in}(s)} = \frac{\Delta\omega_0 s}{s^2 + \Delta\omega_0 s + \omega_0^2} \quad (C2)$$

This expression is in terms of the center frequency of the filter ω_0 and the bandwidth $\Delta\omega_0$. If the

filter is excited with a step sinusoid $f(t) = A_1 \sin \omega_1 t \cdot u(t)$, where $u(t)$ is the unit step function, then

$$F(s) = \frac{A_1 \omega_1}{s^2 + \omega_1^2} \quad (C3)$$

Substituting gives

$$E_o(s) = \frac{\Delta\omega_0 s A_1 \omega_1}{(s^2 + \Delta\omega_0 s + \omega_0^2)(s^2 + \omega_1^2)} \quad (C4)$$

This is of the form

$$\frac{Ks}{(s^2 + as + b)(s^2 + c)}$$

where $a = \Delta\omega_0$, $b = \omega_0^2$, $c = \omega_1^2$, $K = \Delta\omega_0 \omega_1 A_1$.

By performing a partial fraction expansion and a few manipulations,

$$E_o(s) = \frac{\Delta\omega_0 \omega_1 A_1}{(\omega_1^2 - \omega_0^2) + \Delta\omega_0^2 \omega_1^2} \left[\frac{(\omega_1^2 - \omega_0^2)s - \Delta\omega_0 \omega_0^2}{s^2 + \Delta\omega_0 s + \omega_0^2} + \frac{(\omega_0^2 - \omega_1^2)s + \Delta\omega_0 \omega_1^2}{s^2 + \omega_1^2} \right] \quad (C5)$$

Taking the inverse transform gives

$$\begin{aligned} e_o(t) &= \frac{\Delta\omega_0 \omega_1 A_1}{(\omega_1^2 - \omega_0^2)^2 + \Delta\omega_0^2 \omega_1^2} \left\{ e^{-\Delta\omega_0 t/2} \left[(\omega_1^2 - \omega_0^2)^2 + \Delta\omega_0^2 (\omega_1^2 + \omega_0^2)^2 \right]^{1/2} \right. \\ &\quad \left. \cdot \sin \left(-\omega_0 \sqrt{1 - \left(\frac{\Delta\omega_0}{2\omega_0} \right)^2} t + \phi_A \right) + \left[(\omega_0^2 - \omega_1^2)^2 + \Delta\omega_0^2 \omega_1^2 \right]^{1/2} \sin(\omega_1 t + \phi_B) \right\} \quad (C6) \end{aligned}$$

where

$$\phi_A = \tan^{-1} \frac{(\omega_1^2 - \omega_0^2) 2\omega_0 \left[1 - \left(\frac{\Delta\omega_0}{2\omega_0} \right)^2 \right]^{1/2}}{\Delta\omega_0 (\omega_1^2 + \omega_0^2)}$$

$$\phi_B = \tan^{-1} \frac{\omega_0^2 - \omega_1^2}{\Delta\omega_0 \omega_1}$$

The response is now expressed as the sum of two sinusoids, one being the transient term and the other the steady state term. It is more advantageous to express the response as a single sinusoid. A modulating factor which is the

envelope of the response will appear. Let $\omega_0 = \omega_1 + \Delta\omega_1$, where $\Delta\omega_1$ is the difference between the driven frequency and the resonant frequency. Making this substitution and neglecting second order terms yields

$$e_0(t) = \frac{\Delta\omega_0 A_1}{(\Delta\omega_0^2 + 4\Delta\omega_1^2)^{1/2}} \left[-e^{-\Delta\omega_0 t/2} \sin(\omega_1 t + \Delta\omega_1 t - \phi_1) + \sin(\omega_1 t + \phi_2) \right], \quad (C7)$$

where

$$\phi_1 = \tan^{-1} \left(-\frac{2\Delta\omega_1}{\Delta\omega_0} \right),$$

$$\phi_2 = \tan^{-1} \left(+\frac{2\Delta\omega_1}{\Delta\omega_0} \right).$$

Combining the two sinusoids yields

$$e_0(t) = \frac{\Delta\omega_0 A_1}{(\Delta\omega_0^2 + 4\Delta\omega_1^2)^{1/2}} \left(1 + e^{-\Delta\omega_0 t} - 2e^{-\Delta\omega_0 t/2} \cos \Delta\omega_1 t \right)^{1/2} \sin(\omega_1 t + \phi_2 + \phi_3), \quad (C8)$$

where

$$\phi_3 = \tan^{-1} \left(\frac{e^{-\Delta\omega_0 t/2} \sin \Delta\omega_1 t}{e^{-\Delta\omega_0 t/2} \cos \Delta\omega_1 t + 1} \right).$$

APPENDIX D

RESPONSE OF A SINGLE-POLE FILTER TO A CHANGING FREQUENCY

From Appendix C the transfer function $G(s)$ for the filter in Figure 16 is

$$G(s) = \frac{\Delta\omega_0 s}{s^2 + \Delta\omega_0 s + \omega_0^2}, \quad (D1)$$

where $\Delta\omega_0$ is the filter bandwidth and ω_0 is the filter center frequency. The response of the circuit to an impulse is

$$\begin{aligned} h(t) &= L^{-1} \frac{\Delta\omega_0 s}{s^2 + \Delta\omega_0 s + \omega_0^2} \delta(t) \\ &= \Delta\omega_0 e^{-\Delta\omega_0 t/2} \left[\cos \omega_0 \sqrt{1 - \left(\frac{\Delta\omega_0}{2\omega_0}\right)^2} t - \frac{\Delta\omega_0}{\omega_0 \sqrt{1 - \left(\frac{\Delta\omega_0}{2\omega_0}\right)^2}} \sin \omega_0 \sqrt{1 - \left(\frac{\Delta\omega_0}{2\omega_0}\right)^2} t \right]. \end{aligned} \quad (D2)$$

For a reasonably high Q circuit

$$\sqrt{1 - \left(\frac{\Delta\omega_0}{2\omega_0}\right)^2} \approx 1$$

and

$$h(t) = \Delta\omega_0 e^{-\Delta\omega_0 t/2} \left(\cos \omega_0 t - \frac{\Delta\omega_0}{2} \sin \omega_0 t \right) \quad (D3)$$

$$= \Delta\omega_0 e^{-\Delta\omega_0 t/2} \left(1 + \frac{\Delta\omega_0}{2\omega_0} \right)^{1/2} \cos \left(\omega_0 t + \tan^{-1} \frac{\Delta\omega_0}{2\omega_0} \right)$$

$$\approx \Delta\omega_0 e^{-\Delta\omega_0 t/2} \cos \left(\omega_0 t + \tan^{-1} \frac{\Delta\omega_0}{2\omega_0} \right). \quad (D4)$$

The driving function for the filter is a step sinusoid with varying frequency. The rate of change of phase is given by

$$\frac{d\phi(t)}{dt} = \omega_1 + \omega_d \sin \omega_d t, \quad (D5)$$

where ω_1 is the center frequency of driving function, ω_d is the peak frequency deviation, and ω_n is the modulation frequency. Integration yields

$$\phi(t) = \omega_1 t - \frac{\omega_d}{\omega_n} \cos \omega_n t + C_1 \quad (D6)$$

With this expression for the phase of the driving function and $C_1=0$,

$$f(t) = A_1 \sin\left(\omega_1 t - \frac{\omega_d}{\omega_n} \cos \omega_n t\right) \cdot u(t) \quad (D7)$$

The response can be expressed as the convolution of the driving function with the response of the system to be an impulse:

$$h(t) = \int_0^t f(\tau) g(t - \tau) d\tau \quad (D8)$$

And, by substitution,

$$\begin{aligned} h(t) &= \int_0^t A_1 \sin\left(\omega_1 \tau - \frac{\omega_d}{\omega_n} \cos \omega_n \tau\right) \cdot \Delta\omega_0 e^{-\Delta\omega_0(t-\tau)/2} \cos\left[\omega_0(t-\tau) + \tan^{-1} \frac{\Delta\omega_0}{2\omega_0}\right] d\tau \\ &= \Delta\omega_0 A_1 e^{-\Delta\omega_0 t/2} \left[\cos\left(\omega_0 t + \tan^{-1} \frac{\Delta\omega_0}{2\omega_0}\right) \int_0^t e^{\Delta\omega_0 \tau/2} \cos \omega_0 \tau \cdot \sin\left(\omega_1 \tau - \frac{\omega_d}{\omega_n} \cos \omega_n \tau\right) d\tau \right. \\ &\quad \left. + \sin\left(\omega_0 t + \tan^{-1} \frac{\Delta\omega_0}{2\omega_0}\right) \int_0^t e^{\Delta\omega_0 \tau/2} \sin \omega_0 \tau \cdot \sin\left(\omega_1 \tau - \frac{\omega_d}{\omega_n} \cos \omega_n \tau\right) d\tau \right] \quad (D9) \end{aligned}$$

A SOLID STATE SATELLITE SEPARATION SEQUENCE TIMER*

JUSTIN C. SCHAFFERT AND THOMAS D. CLEM

A solid state satellite separation sequence timer is described which is designed to provide for the payload separation sequence at appropriate intervals. Two methods of firing explosive charges are discussed. The versatility of a solid state system is noted as well as its qualification for adverse environments. Examples of the systems used in the UK-1/S-51 (Ariel I) and UK-2/S-52 satellites are given.

INTRODUCTION

Important to the success of a satellite is the proper sequence of events between final stage burnout and payload separation. When orbit entry has been achieved, satellite despin operations must be initiated; the various booms, experiments and solar paddles must be extended; and the satellite must be separated from the last stage of the vehicle.

Mechanical timers driven by clocks or motors have traditionally been used for this purpose. But in several instances their reliability has been in question, especially at the extreme vibration levels experienced on launch vehicles. Thus there arises the need for a more reliable timer to perform the timing functions and provide appropriate outputs to fire explosive charges used to initiate events. The timer developed to fill this need is a solid state system using basic timing circuits that are simple in operation and have a flight history of high reliability. Emphasis was placed on a system that lends itself to easy time adjustments for last minute changes or use in other applications. The final design incorporates possibilities for safe redundant tie-ins between parallel systems.

FUNCTIONAL REQUIREMENTS

Typical functions of a satellite separation system are:

1. Timer Initiation;
2. Deployment of despin mechanism;

3. Erection of inertia booms;
4. Erection of solar paddles; and
5. Payload separation from last stage.

The sequence timers described here are designed to provide outputs anywhere within the following time intervals by means of adjusting a single resistor per interval: The long interval is 0 to 2000 sec; the short interval; 0 to 90 sec.

The design of the separation timer is such that each interval timer is independent and may be connected in series or parallel with any number of other interval timers in any order. Therefore, the order of the intervals and the total number of intervals is dependent only upon the application requirements.

ACCURACY

The accuracy of the timing intervals depends on the degree of compensation used. Off the shelf components with no compensation will yield ± 10 percent or better accuracy over the temperature range of -20°C to $+70^{\circ}\text{C}$. Selection of low-temperature coefficient, low-leakage timing capacitors will yield ± 5 percent in this temperature range. Accuracy of ± 2 percent may be obtained by combinations of bias and timing-current compensation.

POWER REQUIREMENTS

A complete system containing one long-interval timer and three short-interval timers in series, exclusive of the power required to fire the explosive charges, requires approximately 50 mw of

*Published as NASA Technical Note D-2319, July 1964.

power at 12v dc during the timing intervals. Less than 1/4 mw is required before and after the timing interval. The lower limit of operating voltage is 9v as dictated by the decade divider. The upper limit is generally governed by the voltage limit of the capacitors used. In this design 20v capacitors were used at a maximum operating voltage of 18v.

ENVIRONMENTAL LIMITS

The timer operates reliably under the following test conditions:

- Temperature: -35° to $+80^{\circ}\text{C}$ (at atmospheric pressure)
- Thermal Vacuum: -15° to $+60^{\circ}\text{C}$ (at 10^{-6} mm Hg)
- Vibration: 10 to 5000 cps (at maximum acceleration of 54g).

PRINCIPLES OF OPERATION

Long Interval Timing

A block diagram of the timing circuit used for intervals up to 2000 sec is shown in Figure 1. A start pulse or gate turns on the silicon controlled rectifier (SCR) gate. This connects power to a

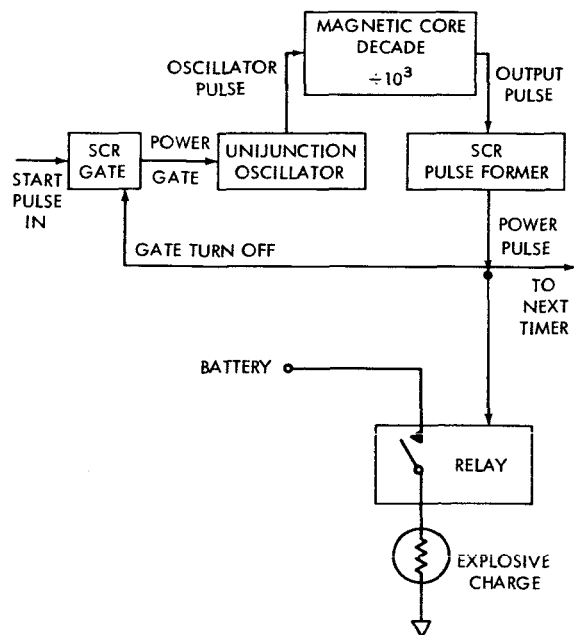


FIGURE 1.—Long interval timer block diagram (0-2000 sec).

unijunction-transistor relaxation oscillator. Three incremental magnetic core decade counter-stages are used to provide a pulse-frequency division of 1000 to 1. The output pulse from the counter fires a SCR. The resulting power pulse is used to:

1. Operate the firing relay for the explosive charges;
2. Start the next timing circuit; and
3. Turn off the timing circuit.

Timing turnoff is accomplished as shown in the schematic diagram (Figure 2). The timer output pulse is fed through R_1 and C_1 to the base of Q_1 . This causes Q_1 to conduct, providing connection of gate and cathode of the input SCR (Q_2), thus removing its drive. Q_2 is therefore turned off, removing power from the unijunction oscillator.

Short Interval Timing

Figure 3 shows a block diagram of the timing circuit used for intervals up to 90 seconds. This diagram is different from the long interval diagram in that a unijunction transistor used as a one-shot multivibrator replaces the oscillator and counter combination. The time interval is adjusted by choice of the R_1 - C_1 combination in the emitter circuit of the unijunction transistor (Q_3) (Figure 4). The time interval may be safely varied up to 90 sec using a $120\mu\text{f}$ wet-slug tantalum capacitor and to 75 sec using a $180\mu\text{f}$ tantalum foil capacitor. The difference in the maximum intervals is due to the inherent difference in leakage of the two types of capacitors. The lower leakage of the wet-slug type allows a higher resistance value for R_1 , and thus a longer maximum time interval.

Relay Method of Firing Explosive Charges

The time constant of the power pulse to the relay is determined, at a given system voltage, by capacitor C_2 at the anode of the output SCR Q_4 (Figures 2 and 4), and the resistance of the relay coil. This time constant is chosen to provide a switch closure of at least 80 msec which is about 10 times the normal firing time of the explosive charge used. The method of firing the explosive charge through a momentary switch closure is chosen to prevent shorting of the system battery in the event the conductive path through the charge should fail to open when it is fired.

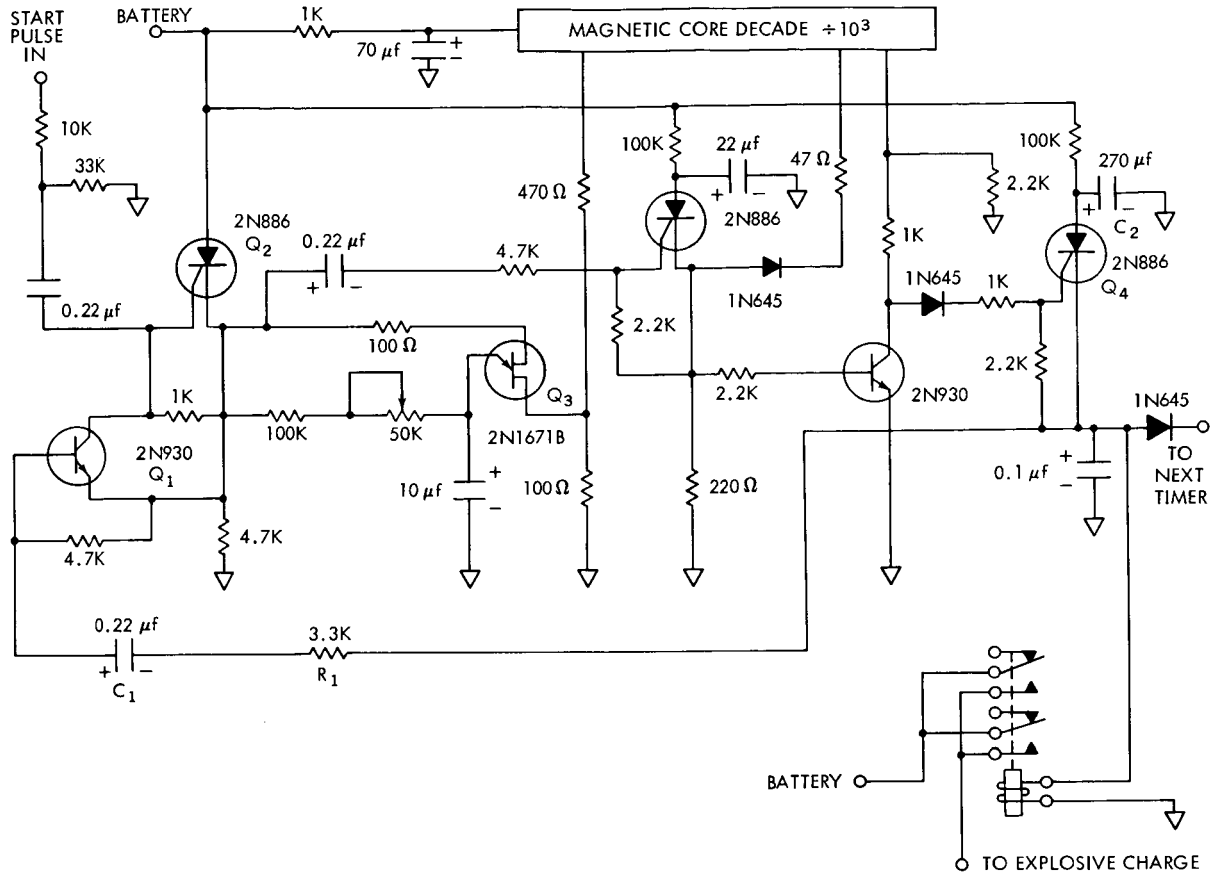


FIGURE 2.—Long interval timer schematic diagram (0-2000 sec.)

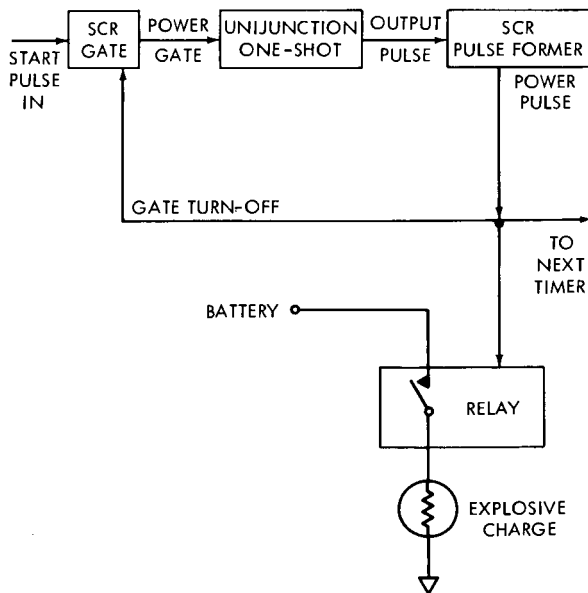


FIGURE 3.—Short interval timer block diagram (0-90 sec.)

The relays used have a normal contact rating of 2 amps at 28v dc. However, tests conducted during the timer development show them capable of an average of about 20,000 operations at a 10:1 duty cycle at 12v and the 25 amp maximum current expected through the explosive charges.

Solid State Method of Firing Explosive Charges

The explosive charge may also be fired directly through a SCR (Figure 5). The 2N1882 may be fired by another SCR, or it may be fired directly from the unijunction-transistor output pulse. The 2N1882 in turn fires the high current 2N682. The 4.7K resistor to ground prevents the normal transistor leakage current from giving the false indication that the SCR is on at the time the explosive charge is connected. This precaution is required by the normal prelaunch procedure for pyrotechnic installation.

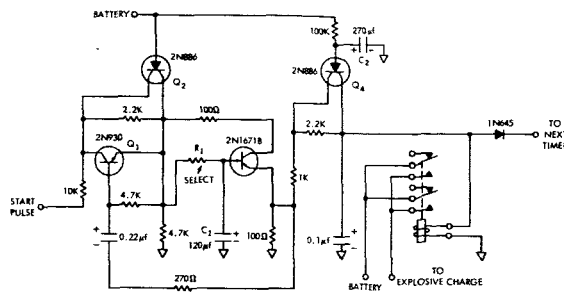


FIGURE 4.—Short interval timer schematic diagram (0-90 sec).

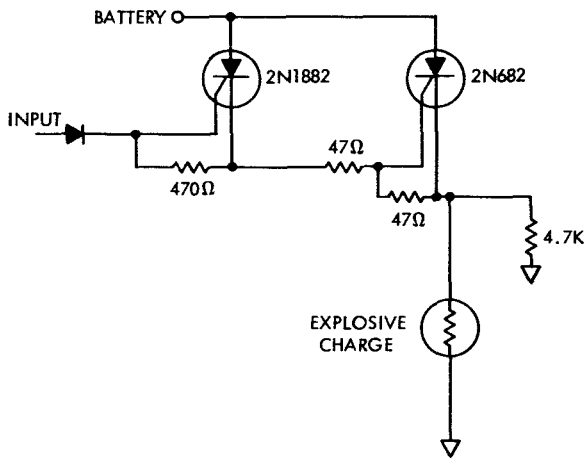


FIGURE 5.—Explosive charge firing circuit.

The method of using SCR's has the advantage of keeping the system completely solid state, but cannot be used where the possibility of a shorted explosive charge exists unless some method of current limiting is employed. Current limiting may be achieved by using fuses or low value resistors.

This solid state method was successfully used on the UK-1/S-51 satellite (Ariel I, 1962 o2).

Redundancy

Due to the independence of the timing interval and explosive-charge firing circuits, and the fact that the oscillator for each interval is normally off, redundant tie-in may be easily and safely accomplished. This is one of the major advantages of the solid state timer over the often used mechanical timers where time interval error generally cannot be corrected but will accumulate during successive intervals. It is also generally not feasible to provide any redundancy in a mechanical timer except at switch contacts.

Figure 6 shows some possibilities for redundant tie-in between two solid state systems. The line shown as R1 accomplishes initiation of both timers by a single start pulse. Lines R2 and R3 in combination would ensure explosive-charge firing and start pulses to both of the next timing circuits in case of failure of one timing interval. It should be noted that in this type of circuit, the most likely failure mode is an extended interval rather than a shortened interval. Line R4 shows another possibility for ensuring pulses to both explosive-charge firing circuits with outputs from either interval timer. Line R5 would start the following timers in case of failure of either explosive charge firing circuit.

The choice as to what amount of redundancy, if any, is needed can only be decided by considering the particular application and the reliability of the timer initiation mechanism, explosive charges, and associated wiring.

TIMER USE IN UK-2

The complete schematic diagram of the timer used in UK-2/S-52 is shown in Figure 7. In this application, the timer and explosive charges use the same battery. The capacitor C₁ and diode D₁ provide decoupling when the battery voltage drops from a nominal 18v to about 4v at the time of firing each explosive charge.

The wiring of the complete separation system used in UK-2/S-52 is shown in Figure 8. In this application, two sets of batteries, timers, and explosive charges are used. The only tie-in between units is at the explosive charges. No tie-in is made between timers. The UK-2/S-52 timer is shown in Figure 9.

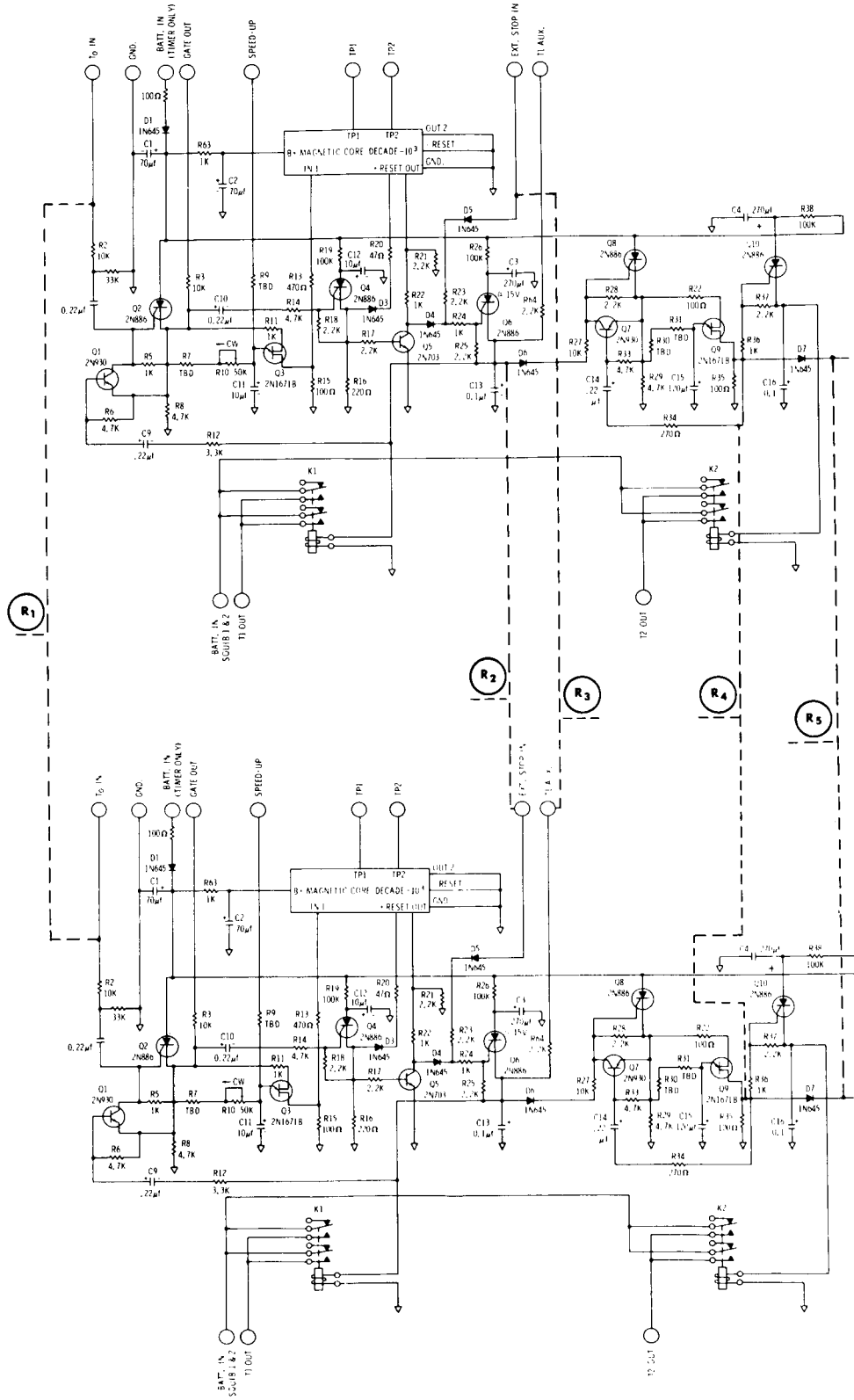


FIGURE 6.—Possible redundant tie-in.

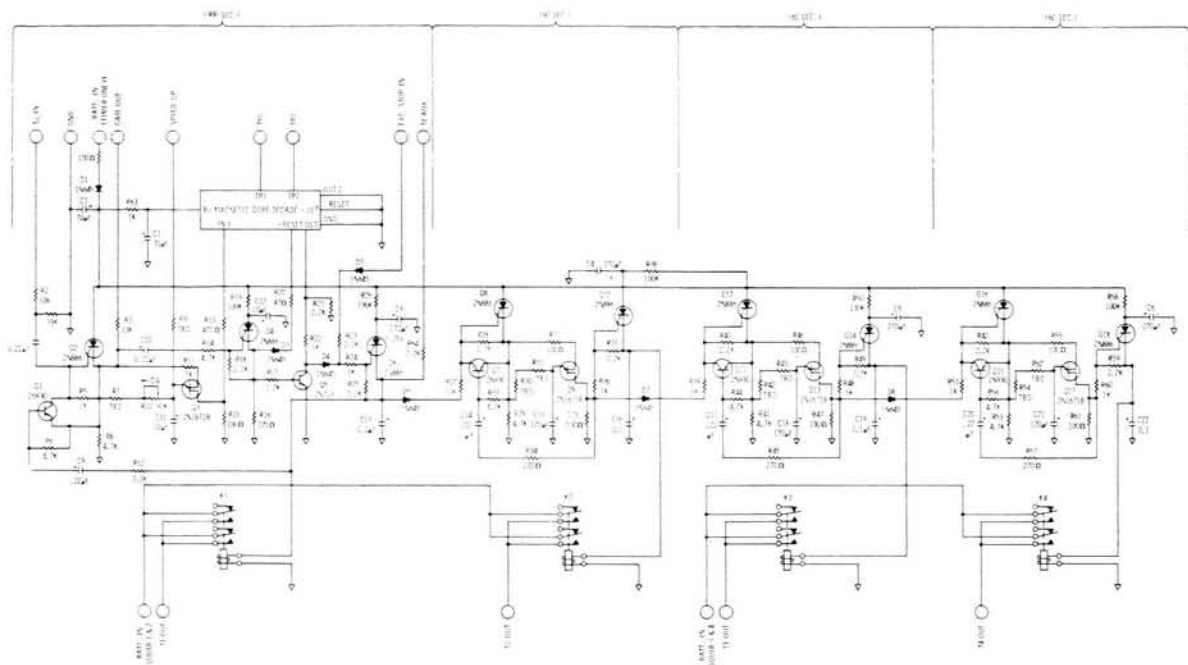


FIGURE 7.—Complete UK-2/S-52 application schematic.

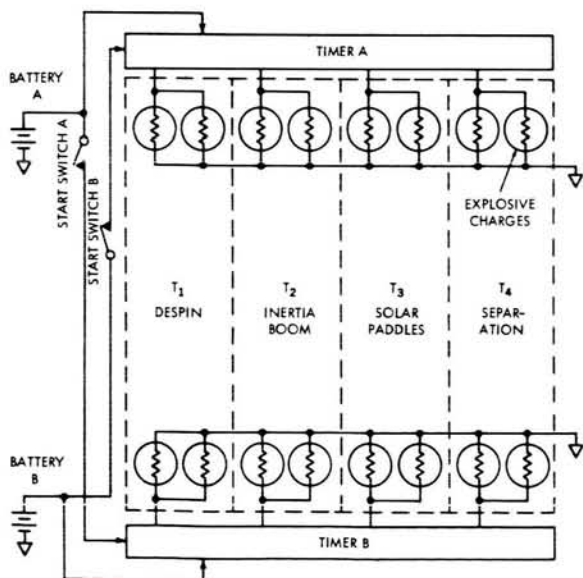


FIGURE 8.—UK-2/S-52 separation system wiring.

CONCLUSION

The methods developed for timing and explosive-charge firing have been thoroughly tested in the lab and proven in flight. The developed variations on the basic design will allow easy applica-

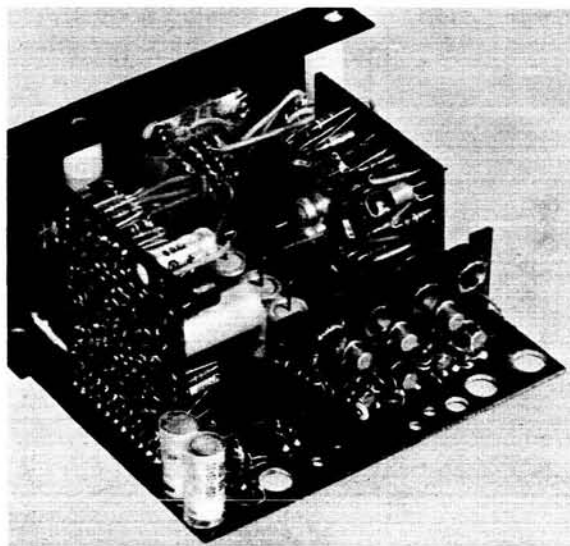


FIGURE 9.—Photo of UK-2/S-52 separation sequence timer.

tion in future programs. The number of events controlled is merely a matter of adding or deleting identical stages. Needs for various time intervals are met simply by the adjustment of a R-C time constant.

TRANSISTORIZED VHF TRANSMITTER DESIGN FOR SPACECRAFT APPLICATIONS*

CHARLES R. SOMERLOCK

Little has been published about VHF transmitter design employing transistors. Because of the fact that these devices are operating so far down on the frequency response curve, detailed circuit analysis is very difficult. However, with the use of liberal approximations and some experimental results, a rough design can be made. A design of a satellite transmitter is made as an example of the procedure. Such a transmitter demonstrates advanced practice in the use of radio frequency transistors as well as an application, which is itself of interest. As verification of the method, the sample transmitter was built and tested.

INTRODUCTION

In this age of transistor electronics, when solid state circuitry is replacing the more familiar vacuum tubes in many fields, there appears to be a great lack of informative literature about VHF transmitter design employing transistors. Of special interest is the application to missile and space vehicle telemetry, since the requirements of light weight, low power drain, and rugged environmental capability make these devices ideal for this purpose.

This paper, therefore, presents a sample design for a VHF telemetry transmitter using transistors throughout. Various problems and pitfalls peculiar to the use of transistors and to this application are discussed, and suggestions are made to assist the inexperienced designer.

PHILOSOPHY OF CIRCUIT DESIGN

Oscillators

At the heart of every transmitter is an oscillator that generates the RF energy. Because of the frequency stability requirements of most transmitters, this oscillator is usually crystal-controlled. Use of a quartz crystal for frequency control does not necessarily guarantee good stability, however. Where the transmitter parameters can affect the frequency of oscillation, the frequency stability of the oscillator with temperature may be several

times worse than that of the crystal itself. This is because the transistor parameters themselves are very temperature-sensitive. Accordingly, the transmitter oscillator circuit should be designed so that the frequency of oscillation is dependent solely on the quartz crystal. To help reduce the effect of the transistor, a transistor should be chosen with the highest possible cutoff frequency, so that its phase shift at the operating frequency is negligible.

For similar reasons, and to prevent the possibility of incidental frequency modulation from occurring, the oscillator frequency-determining elements should be isolated as well as possible from the load impedance. The impedance of the load on the oscillator is sure to vary over the modulation cycle and over the temperature range, so care must be taken that it does not affect the oscillator frequency. When a phase-lock receiving system of narrow bandwidth is used, incidental frequency shifts of only a few cycles can cause distortion and difficulty in maintaining lock.

Figures 1, 2, and 3 compare the frequency versus temperature curves of certain oscillators to those of their quartz crystals alone. It is clear that the oscillator circuit used in Figures 1 and 2 allows elements other than the crystal to affect the frequency. This circuit, a series mode type that is widely used in VHF devices, is shown in in Figures 4(a) and 4(b).

On the other hand, examine the circuit plotted in Figure 3, a modification of the familiar Pierce oscillator [see Figure 4(c)]. Within the limits of

*Published as *NASA Technical Note D-2104*, May 1964. Originally this work was submitted to the University of Maryland as partial fulfillment of the requirements for the degree of Master of Science, 1963, under the title "A Transistorized VHF Transmitter."

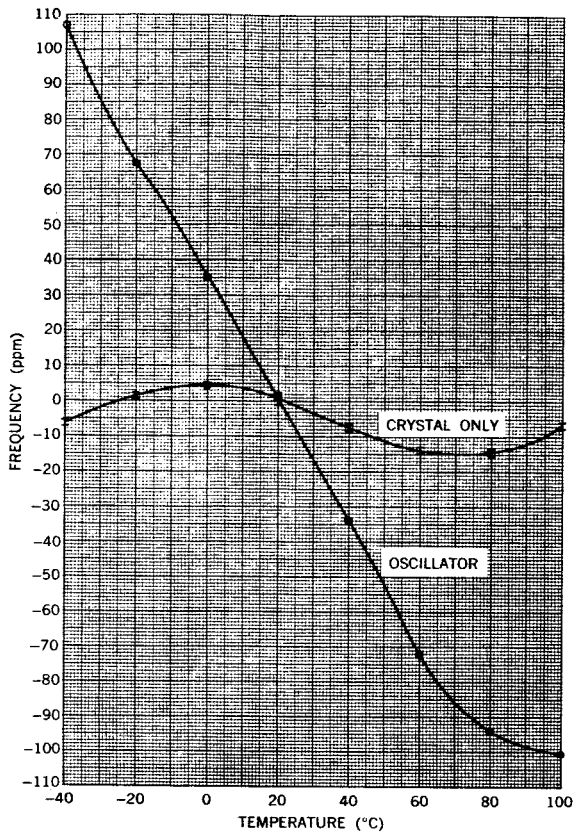


FIGURE 1.—Frequency drift of a 17 Mc series mode oscillator [see Figure 4(a)].

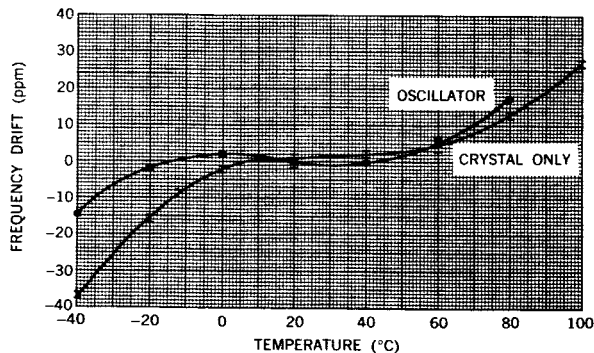


FIGURE 2.—Frequency drift of a 68 Mc series mode oscillator [see Figure 4(b)].

measurement error, its frequency is independent of all circuit elements except the crystal. This excellent characteristic is due partly to the fact that the crystal is the only resonant element in the whole oscillator and partly to the negative feedback that stabilizes the gain of the transistor.

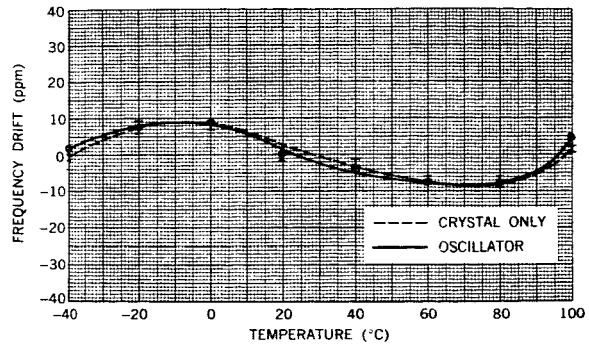


FIGURE 3.—Frequency drift of a 17 Mc modified Pierce oscillator [see Figure 4(c)].

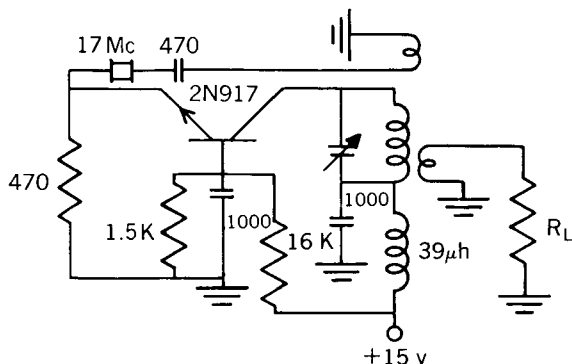
Also, since the oscillator is basically a Colpitts configuration, it retains all the advantages of that well-known circuit.

In a similar way, Figure 5 shows the effect that changes in load impedance have on these same oscillators. Again, the Pierce configuration shows excellent independence of loading while the series mode oscillator shows up poorly in comparison. The excellent stability of this oscillator under load is due to the large negative feedback provided by the emitter resistor. This feedback has the effect of lowering the output impedance of the circuit, making it insensitive to all but extremely low values of load resistance. The modified Pierce oscillator, while not the only type of oscillator that is usable, demonstrates the sort of performance that is recommended.

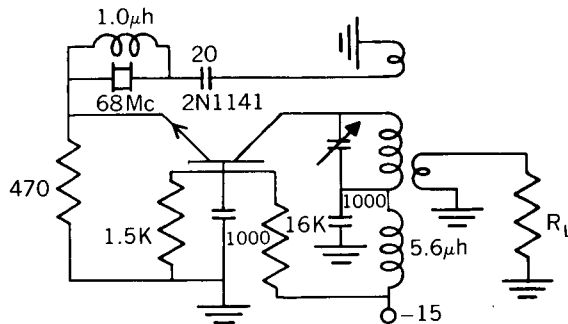
Along with good frequency stability, some degree of amplitude stability is usually required; but, as long as the output of the oscillator does not vary more than about 2 db, no problems will generally occur. This condition usually is easily met with class C oscillators, since their output amplitude is set by some limiting action in the circuit and plenty of excess gain is available to sustain oscillations.

Amplifiers

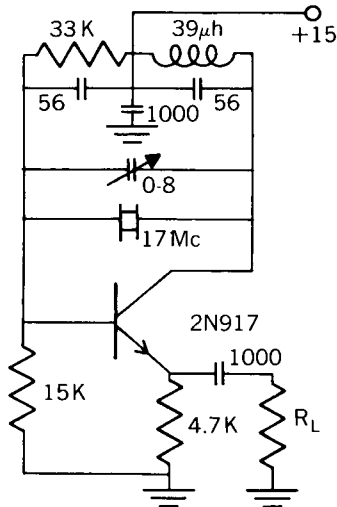
The high efficiency requirement of a transmitter amplifier stage dictates the use of class C operation in all but the lowest level stages. Class C circuits can be self-biased by placing a resistor in the dc path to ground from either the base or the emitter, where the signal current flowing will develop a reverse bias. Typical bias levels are usually less than 1 volt.



(a) 17 Mc series resonant oscillator



(b) 68 Mc series resonant oscillator



(c) 17 Mc modified pierce oscillator

FIGURE 4.—Oscillator circuits: (a) 17 Mc series resonant oscillator; (b) 68 Mc series resonant oscillator; (c) 17 Mc modified Pierce oscillator.

Because of the relatively low corner frequency of the gain versus frequency curve, the gain of even the best VHF power transistors available is

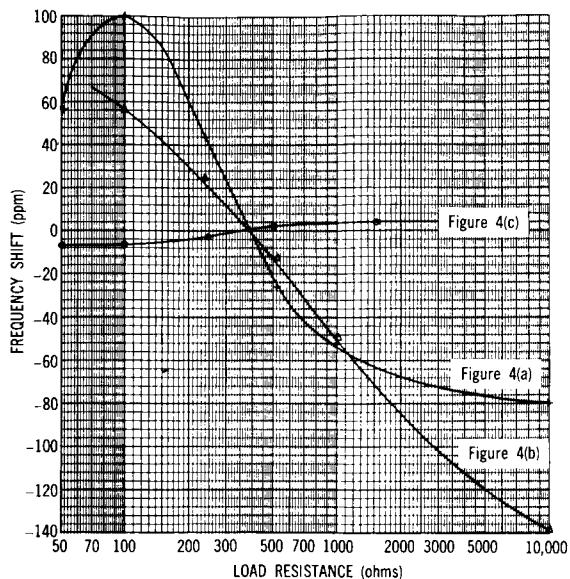


FIGURE 5.—Effect of oscillator load resistance.

only 6 to 10 db at 150 Mc. Typical collector efficiencies obtainable at this frequency are approximately 40 to 60 percent, including tank circuit efficiency. Calculating the expected power gain from the transistor collector curves as is done with vacuum tubes is an impossible task, since the operating point is so far down on the frequency response curve. Conventional small signal analysis is not adequate, since transmitter amplifier stages are large signal devices. Performance parameters, measured at the frequency of interest and supplied by the manufacturer, must be used along with liberal approximations and—many times—experimental results. An additional complication is the wide variation of the transistor parameters allowed between units of the same type by current practice in transistor manufacture. Because of these difficulties, detailed time-consuming analysis of transistor class C power amplifiers is not profitable. The liberal use of approximations and assumptions based on experience and experiment is justified and will yield results that are just as reliable as the more detailed ones based on invalid processes.

The design of an amplifier stage begins with the choice of the supply voltage for the transistor to be used. For maximum efficiency and gain, the supply voltage should be as high as practicable without exceeding the collector breakdown

voltage of the transistor. Since the peak-to-peak signal voltage is approximately twice the supply voltage, the supply is limited to:

$$V_{cc} \approx \frac{1}{2} V_{ce}, \text{ or } \frac{1}{2} V_{cb},$$

depending on the configuration. (All symbols are defined in Appendix A.) The grounded base circuit has demonstrated greater stability than the grounded emitter type and has only a negligibly smaller power gain.

Once the voltage and the required output power are known, the plate load resistance can be found:

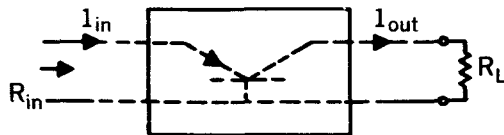
$$R_L = \frac{V_c^2}{2P_o} \approx \frac{V_{cc}^2}{2P_o},$$

from which the tank circuit parameters follow:

$$X_c = X_L = \frac{R_L}{Q}.$$

Optimum Q for a class C amplifier is typically 10 to 20, representing a balance between selectivity and tank circuit efficiency.

A relation between input resistance and power gain can be developed as follows:



$$I_{in} = \frac{I_{out}}{G_i},$$

$$P_{in} = \frac{P_o}{G_p},$$

$$R_{in} = \frac{P_{in}}{(I_{in})^2} = \frac{P_o/G_p}{I_{out}^2/G_i^2} = R_L \frac{G_i^2}{G_p};$$

for grounded base and $f \ll f_{max}$:

$$G_i \approx 1,$$

$$R_{in} \approx \frac{R_L}{G_p}.$$

Usually, measured curves of power gain versus frequency are available from the transistor data

sheets. Where these are not available, the process must be reversed to find G_p from R_{in} , which to a first approximation can be assumed to be $1/ReY_{ib}$. The latter approximation is generally not too good, however.

Any of the usual methods of interstage coupling that will match the impedances can be used. Capacitive coupling methods seem to work out better, though—probably because they can be adjusted more readily to optimum conditions. One simple coupling method is connecting the stages with a small value capacitor. The collector load impedance is then simply the parallel equivalent of the series combination of the coupling capacitor and the input resistance of the next stage. Assuming R_L/R_{in} is greater than about 25,

$$R_L = Q^2 R_{in} = \left(\frac{X_c}{R_{in}}\right)^2 R_{in} = \frac{X_c^2}{R_{in}};$$

so

$$X_c = \sqrt{R_L R_{in}}.$$

The equivalent parallel capacitance (approximately X_c) will, of course, have to be counted as part of the tank circuit capacity.

A major problem with transistor amplifiers is gain stability. The gain of such a device will vary over the temperature range. For this reason, it is desirable to operate the amplifier in a slightly saturated condition. Figure 6 shows how this

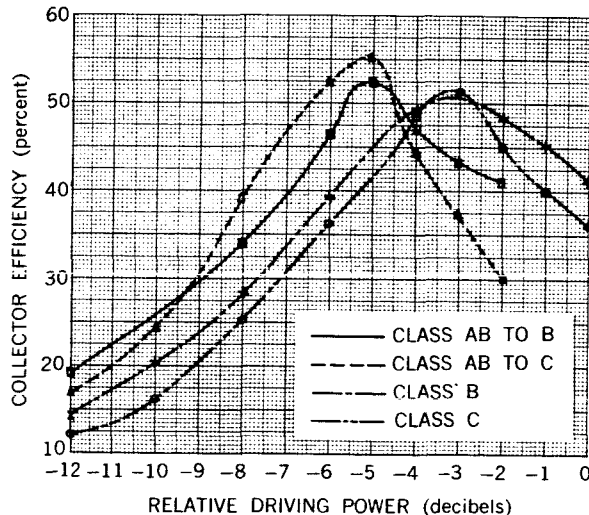


FIGURE 6.—Amplifier collector efficiency.

type of operation will sacrifice some efficiency, but it is necessary if the output power is to be independent of the ambient temperature. For the same reason, the response of the amplifier below the saturated state should be nearly linear. As seen in Figure 7 the class C amplifier is not

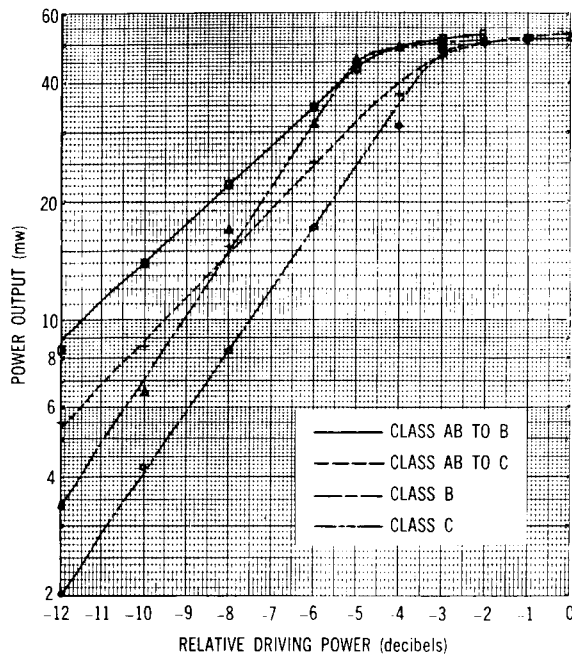
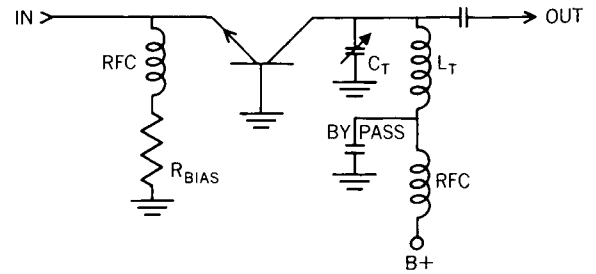


FIGURE 7.—Amplifier drive characteristics.

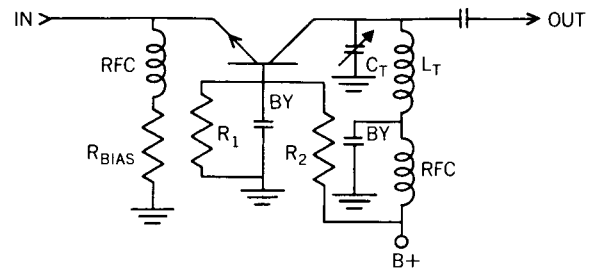
linear, small changes in drive causing larger changes in output. Thus in a transmitter employing several class C stages, a small drop in oscillator output can be greatly magnified.

The response can be made more linear (Figure 7) if a small forward bias is applied to the base as shown in Figure 8. At normal output, the self-generated reverse bias overcomes the fixed forward bias, and the operation is class C. As the drive level drops, however, the bias tends more toward forward bias, the gain increases, and a nearly linear output results. Even better regulation can be obtained if the forward bias is developed across an external diode placed in series with the base. The diode will then compensate for the change in the base-emitter voltage drop with temperature variations.

Feedback, both internal and external to the transistor, is also frequently a problem. It usually manifests itself as either outright regeneration or



CLASS C RF AMPLIFIER



CLASS C RF AMPLIFIER WITH FORWARD BIAS AT LOW SIGNAL LEVELS

FIGURE 8.—Radio frequency amplifiers.

as a "hysteresis" effect in the amplifier tuning characteristic. In theory, unilateralization is the obvious answer. In practice, though, unilateralization of large signal transistor VHF amplifiers is difficult. Because the transistor is operating near its cutoff frequency, it has a large phase shift. This phase shift requires use of somewhat more complex feedback networks, the components of which must be determined practically by trial and error. With a few precautions, however, the difficult unilateralization problem can be bypassed.

Care should be taken to eliminate all sources of feedback external to the transistor. Among common sources are inadequate shielding, excessive lead lengths (particularly the "common" transistor lead), and chassis currents. The latter two failings are frequently overlooked by novice designers. For good shielding, the "egg crate" style of packaging is recommended. Good grounding of the "common" transistor lead can be accomplished where physical conditions limit the shortness of the lead by series-tuning of the lead inductance with a bypass capacitor slightly smaller than normal. Caution is important here, so that the bypassing is not underdone. Chassis currents

are frequent sources of feedback in transmitters where large power gains are present. One example of this problem is the grounding of a transistor lead in the current path between a tank inductor bypass and the tank capacitor mounting. There is a large circulating current in this path that, in a careless effort to increase packaging density, can cause real problems.

The grounded base circuit is highly recommended over the grounded emitter for stability. While the common base circuit has only slightly less power gain at VHF frequencies and large signal conditions than the common emitter circuit, it demonstrates a much smaller internal feedback problem.

The choice of the proper transistor can be important to this problem also: generally, the higher the transistor's cutoff frequency, the lower its inherent feedback. Accordingly, where several transistors are available, the one with the highest frequency rating should be chosen. With all other things equal, this will usually turn out to be the one that is most efficient and demonstrates the best gain.

Buffer Amplifiers

Because of their internal feedback, transistors generally do not make very good buffer amplifiers. Neutralization will help, of course; but this is tricky, as was pointed out in the section on amplifiers. Figure 9 shows the effect of load

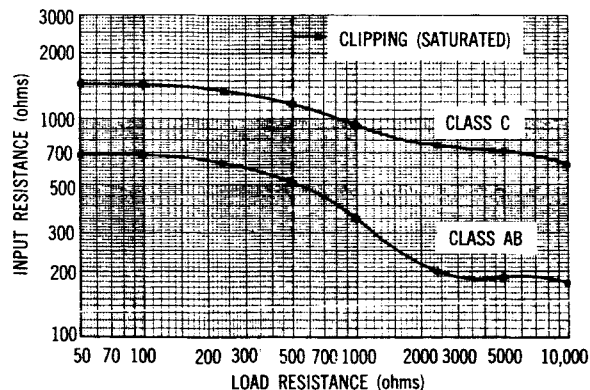


FIGURE 9.—Effect of amplifier load resistance.

resistance on the input resistance of an amplifier (Figure 10) in two classes of operation. The graph shows that the variation of input resistance

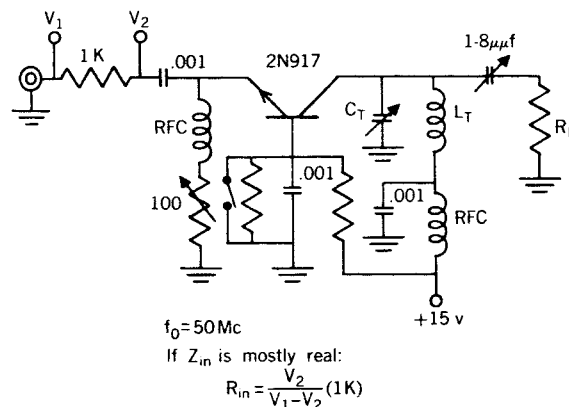


FIGURE 10.—Buffer test circuit.

becomes more severe after the amplifier saturates. For good buffer action, then, an amplifier should be loaded somewhat heavily so that it will not operate in a saturated condition. The class C mode of operation demonstrates a slight improvement in isolation over the class AB, but the superiority is slight. The extra gain of class A or AB operation may allow use of resistive padding for additional isolation with no increase in power requirements. An emitter follower or some other circuit employing negative feedback can be used where good isolation is important. These circuits however, have very low power gain, and overall efficiency will suffer somewhat. Emitter followers are also a little temperamental at VHF, sometimes developing parasitic oscillations under capacitive loads.

Frequency Multipliers

Since transistors are very nonlinear devices in large signal operation, they can be used effectively as frequency multipliers. The use of large reverse bias will enhance the efficiency and power gain in this type of operation because the harmonic distortion of the output current waveform is increased by the shorter conduction angle. Collector tank circuits with higher Q also should be used, so that effective rejection of the fundamental driving frequency is accomplished. Because of the fact that the available power gain at VHF for large signal amplifiers is already low, the gain of a frequency multiplier will be only 3 or 4 db usually; and it will decrease as the conduction angle is shortened. A compromise therefore must usually be made

between efficiency of the stage and power gain (which will affect the overall system efficiency). At lower frequencies, performance, of course, increases. In a practical sense, however, multiplication greater than tripling in a single stage is too inefficient to be useful in a transmitter. Multiplier strings should, if possible, be arranged so that the desired order of frequency multiplication can be accomplished with cascaded doublers and triplers.

SAMPLE TRANSMITTER DESIGN

Proposed Requirements

As an example of VHF transistor transmitter design, take a requirement for a telemetry transmitter in an earth satellite where rugged environmental stresses are encountered and advanced performance is required. Such a unit must be small, light in weight, efficient, reliable, operate over a wide temperature range in a vacuum, be mechanically sturdy, and generally demonstrate performance that is consistent with the state of the art. The initial requirements are, then:

- Efficiency, >25 per cent
- Weight, <1 lb.
- Temperature, 0° to 100°C
- Spurious emission, >60 db below carrier
- Output frequency, 136.500 Mc

System Design

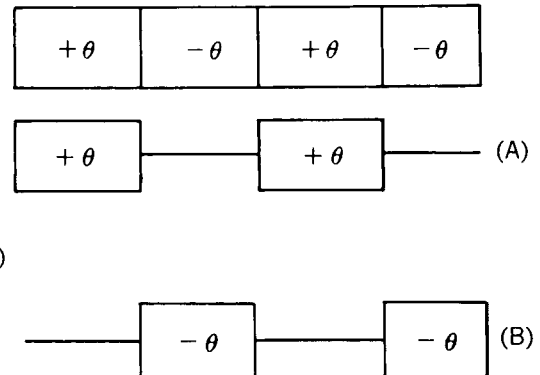
Modulation Method

The first part of the system to decide upon is the type of modulation to be used. Of the various methods in common use today, phase modulation (PM) was chosen for several practical reasons. First of all, PM is readily compatible with the synchronous detection systems presently in use by the NASA tracking stations. This linear detection process allows use of post-detection

filtering to best advantage. Second, PM, by reason of its constant amplitude carrier wave, permits the transmitter RF circuits to be adjusted for maximum efficiency without regard to linearity. Third, with the constant amplitude square wave deriving from pulse-type encoding systems, the proportion of output power in the sidebands can be easily adjusted from 0 to 100 percent—depending on requirements—by setting the index of modulation.

For this transmitter, it was decided to make the power in the carrier equal to that in each of the principal sidebands. This condition provides a large proportion of power in the information sidebands, yet it leaves sufficient power in the carrier for tracking purposes. The proper modulation level can be adjusted quite easily by observing the spectrum on a panoramic display unit.

To find the required modulation index for square wave modulation, the constant amplitude, phase-switching carrier is broken into two amplitude-modulated components:



Each component is a constant phase, amplitude-modulated carrier that can easily be analyzed by Fourier series, and whose spectra can be added to obtain the spectrum of the original phase-modulated wave:

$$\begin{aligned}
 A = & \sin(\omega_c t + \theta) \left\{ \frac{k}{2} + \frac{2k}{\pi} \left[\cos \omega_m t + \frac{1}{3} \cos 3\omega_m t + \frac{1}{5} \dots \right] \right\} \\
 & - \frac{k}{2} \sin(\omega_c t + \theta) + \frac{k}{\pi} \left\{ \cos [(\omega_c - \omega_m) t + \theta] + \cos [(\omega_c + \omega_m) t + \theta] \right. \\
 & \left. + \frac{1}{3} \cos 3 [(\omega_c - \omega_m) t + \theta] + \frac{1}{3} \cos 3 [(\omega_c + \omega_m) t + \theta] + \dots \right\} .
 \end{aligned}$$

$$B = \frac{k}{2} \sin(\omega_c t - \theta) + \frac{k}{\pi} \left\{ \cos [(\omega_c - \omega_m) t - \theta] + \cos [(\omega_c + \omega_m) t - \theta] \right.$$

$$\left. + \frac{1}{3} \cos 3 [(\omega_c + \omega_m) t - \theta] + \frac{1}{3} \cos 3 [(\omega_c - \omega_m) t - \theta] + \dots \right\} ,$$

$$A + B = k \cos \theta \sin \omega_c t + \frac{2k}{\pi} \sin \theta \left\{ \sin(\omega_c - \omega_m) t \right.$$

$$+ \sin(\omega_c + \omega_m) t + \frac{1}{3} \sin [3(\omega_c - \omega_m) t]$$

$$\left. + \frac{1}{3} \sin [3(\omega_c + \omega_m) t] + \dots \right\} .$$

Setting the amplitude coefficients of the principal sidebands and the carrier equal,

$$k \cos \theta = \frac{2k}{\pi} \sin \theta ,$$

$$\frac{\sin \theta}{\cos \theta} = \tan \theta = \frac{\pi}{2} ,$$

$$\theta = 57.5^\circ \approx 1 \text{ radian} .$$

To find the proportion of power in the carrier and sidebands, the total radiated power is needed. This can be found by setting $\theta=0$, whereupon all of the power is in the carrier:

$$P_{\text{total}} = (k \cos 0)^2 = k^2 ,$$

$$\frac{P_c}{P_{\text{total}}} = \frac{(k \cos \theta)^2}{k^2} = \cos^2 57.5 ,$$

$$P_c = 29\% P_T ,$$

$$\frac{P_1}{P_{\text{total}}} = 2 \frac{\left(\frac{2k}{\pi} \sin \theta\right)^2}{k^2} = \frac{8}{\pi^2} \sin^2 57.5^\circ ,$$

$$P_1 = 58\% P_{\text{total}} ,$$

$$P_{>1} = P_{\text{total}} - P_c - P_1 = 13\% P_T .$$

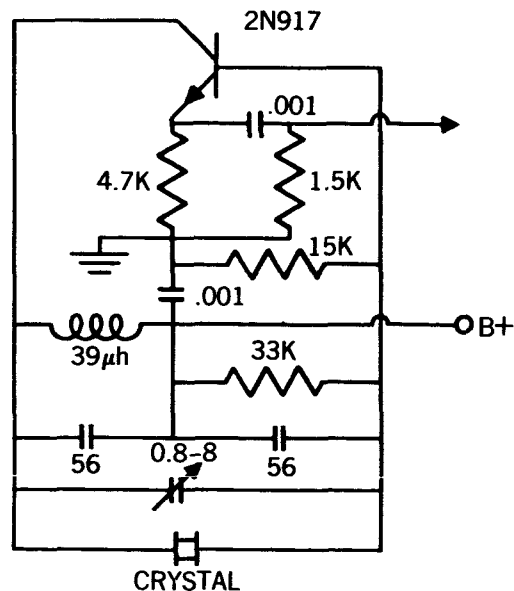
Power Output

Because there are so many factors relating to total system performance that affect the calculation of the required power output of the trans-

mitter, it is impossible to produce an exact number without knowing more about the particular application involved. The output of this sample design was therefore chosen arbitrarily as 1 watt. Figure 11 shows the available power at the receiver for a typical antenna system and a 1-watt transmitter at various ranges.

Circuit Details

Oscillator



In an earth-satellite telemetry application, the Doppler shift due to the satellite velocity will have a large effect on the apparent transmitter frequency as seen from the ground. There is, then,

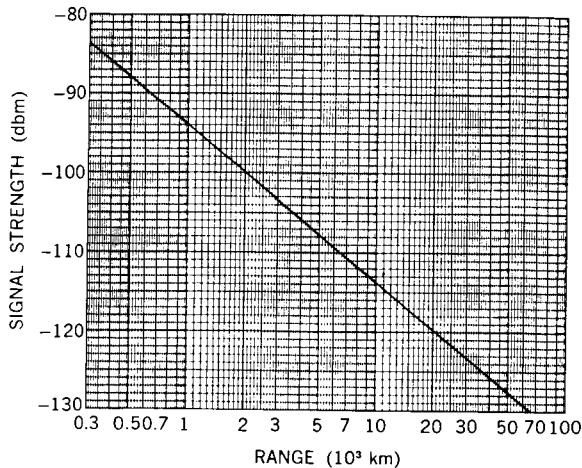


FIGURE 11.—Signal strength vs. range. $P_T=1$ watt; $G_T=-3$ db; $G=+20$ db; system loss $=-6$ db; $P_r=A_{\text{eff}}S$, with $S=P_T/4\pi r^2$ and $A=G_r\lambda^2/4\pi$.

little to be gained by requiring a long-term frequency stability and calibration of the transmitter oscillator to much closer limits than the Doppler shift itself. If the frequency tolerance is set equal to this velocity shift for an "average" orbit,

$$\text{Tolerance} = \frac{\Delta f}{f_0} = \frac{\nu \cos \phi}{c} \approx \pm 0.002\%$$

where

- ϕ = the angle between the observer and the velocity vector,
- ν = velocity of satellite,
- c = velocity of light.

A frequency tolerance of ± 0.002 percent can be obtained with commercially available quartz crystals without using temperature-controlled ovens or frequency compensation techniques.

High orders of short-term stability are required, however, because of the phase-lock detection system commonly in use for satellite tracking and telemetry reception. Steps should be taken, therefore, to prevent frequency shifts due to modulation or supply voltage transients.

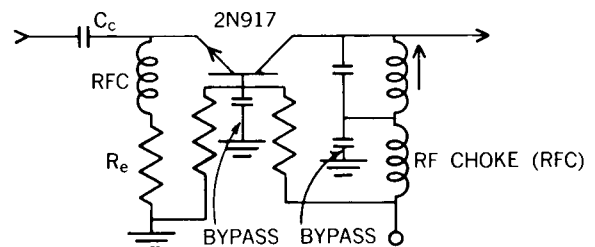
To meet these requirements, a modified Pierce oscillator [Figure 4(c)] was chosen. As is shown in Figures 3 and 5, this circuit provides good isolation of the oscillator frequency from circuit elements and loading effects. The output impedance provides the good load isolation. The absence of any tuned elements in the circuit other than

the crystal enhances the crystal's control over the frequency of oscillation, which improves the stability.

The Pierce oscillator is essentially a crystal-controlled Colpitts circuit. The crystal oscillates in an anti-resonant mode, and feedback is controlled by the ratio of the two series capacitors across the crystal and the emitter load resistance. The series equivalent of these two capacitors, added to that of the variable trimmer, determines the crystal load capacity, which is usually set to $32 \mu\text{f}$. The variable trimmer allows close calibration of the oscillator frequency by "pulling" the crystal. A 2N917 transistor is used because it has a maximum frequency of oscillation of 1000 Mc. This high cutoff frequency means that the inherent phase shift in the device at low frequencies is negligible and better stability will result.

Because the oscillator has no tuned circuits, oscillation can occur only on the fundamental mode of the quartz crystal. This limitation sets an upper limit on the oscillator frequency of 20 Mc, since this is the highest frequency fundamental crystal that can be obtained. The lowest order, non-prime submultiple of the output frequency of 136.5 Mc is 17.0625 Mc. Therefore, a frequency multiplication of 8 is required between the oscillator and the transmitter output.

Buffer Amplifier



The output from the oscillator was experimentally measured at about 1 mw. Since this output is insufficient to drive a class B or C amplifier stage, the buffer amplifier must be operated class AB. A 2N917 transistor is used because of its high gain, high frequency rating, and efficiency. The design procedure follows:

$$V_{cc} = \frac{1}{2} V_{cb} = +15 \text{ volts .}$$

Very little information on large signal operation is given by the manufacturer. The only known quantities are:

$$P_{in} = 1 \text{ mw} ,$$

$$R_{in} \approx \frac{1}{Re Y_{ib}} = 300 \text{ ohm}$$

and the relations

$$P = \frac{V_c^2}{2R_L} = \frac{(V_{cc} - V_{min})^2}{2R_L} ,$$

$$G_p \approx \frac{R_L}{R_{in}} ,$$

which, when combined, yield

$$P_o^2 = \frac{V_c^2 P_{in}}{2R_{in}} .$$

For this circuit,

$$P_o = 14 \sqrt{\frac{10^{-3}}{600}} = 18 \text{ mw} ,$$

$$R_L = \frac{V_c^2}{2P_o} = \frac{(14)^2}{2(18)} = 5.5 \text{ K} ,$$

$$X_T = \frac{R_L}{Q} = \frac{5.5 \text{ K}}{10} = 550 \text{ ohm} ,$$

$$f_0 = 17.0625 \text{ Mc} ,$$

$$C_T = 17 \mu\text{mf} \text{ (this includes collector capacity and that reflected from the coupling network),}$$

$$L_T = 5 \mu\text{h} .$$

Assuming $Q_T \approx 100$,

$$\eta_T = 1 - \frac{Q_{loaded}}{Q_T} = 0.9 ,$$

$$P_o' = 0.9 (18) = 16 \text{ mw} .$$

For a rough approximation of the dc power required, assume the collector efficiency is typically

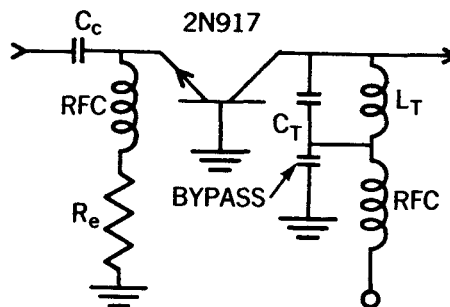
50 percent.

$$P_{dc} = \frac{P_o'}{0.5} = 32 \text{ mw} ,$$

$$I_{dc} = \frac{P_{dc}}{V_{cc}} = \frac{32}{15} = 2.1 \text{ ma} .$$

The bias network is adjusted to produce a collector current of about 0.5 ma under no-signal conditions.

First Frequency Multiplier



The required 8-times-frequency multiplication is most easily achieved with three cascaded frequency doublers. The first frequency multiplier doubles the 17.0625 Mc output from the buffer amplifier to 34.125 Mc. A 2N917 transistor is used.

$$R_{in} = \frac{1}{Re Y_{ib}} = 300 \text{ ohms} ,$$

$$X_c = \sqrt{R_L R_{in}} = 1.3 \text{ K} ;$$

therefore,

$$C_c = 7.2 \mu\text{mf} .$$

If the amplifier were a straight-through class B stage,

$$P_{o,1} = V_c \sqrt{\frac{P_{in}}{2R_{in}}} = 14 \sqrt{\frac{16}{2(0.3)}} = 72 \text{ mw} .$$

But, since the stage is operated class C over a short conduction angle, the gain will be about 2 db smaller (see Figure 7). An exact figure for the increased drive requirement can be obtained only by Fourier analysis of the input waveform, which

cannot be done under the circumstances. Experience provides the only answer.

$$P'_{o,1} = P_{o,1} - 2 \text{ db} \approx 50 \text{ mw} ,$$

$$I_{c,1} = \sqrt{2} \frac{P_{o,1}}{V_c} = 5.1 \text{ ma, rms} .$$

For a frequency doubler, the conduction angle should be about 110 degrees (Reference 1). Fourier analysis shows the ratio of fundamental current in the collector pulses (assuming a square law device) to be (Reference 1):

$$I_{c,2} = \frac{0.25}{0.30} I_{c,1} = 4.25 \text{ ma} ,$$

$$P_{o,2} = \frac{V_c}{\sqrt{2}} I_{c,2} = 42 \text{ mw} .$$

Assuming

$$\eta_T = 0.9 ,$$

then

$$P'_{o,2} = 38 \text{ mw} .$$

For better temperature stability it is desirable to run the doubler in a saturated condition. To accomplish this state, about 1 db of gain is wasted and the output is designed for 30 mw.

$$R_L = \frac{V_c^2}{2P_o} = 3.3K ,$$

$$X_T = \frac{R_L}{Q} = 330 \text{ ohms} ,$$

$$f_0 = 34.125 \text{ Mc} ,$$

$$L_T = 1.5 \mu\text{h} ,$$

$$C_T = 14 \mu\text{mf} \text{ (including collector and reflected capacity).}$$

To obtain the required conduction angle of 110 degrees,

$$\left(V_{in-peak} \right) \left(\cos \frac{\theta}{2} \right) = V_{bias} + V_{diode} ,$$

$$V_{in} = \sqrt{2} \sqrt{P_{in}} R_{in} = 3.2 \text{ volts} ;$$

so

$$V_{bias} = 3.2 \cos(55^\circ) - 0.7 = 1.1 \text{ volts} .$$

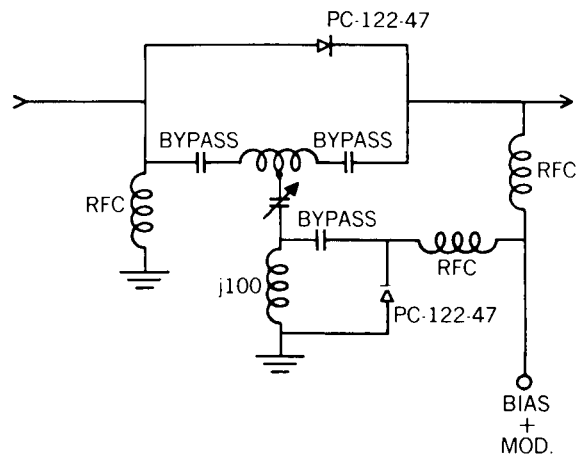
Experiment indicates that a typical collector efficiency for a doubler stage is about 40 percent.

$$P_{dc} = \frac{P_{o,2}}{\eta_c} = \frac{30}{0.4} = 75 \text{ mw} ,$$

$$I_{dc} = \frac{P_{dc}}{V_{cc}} = 5 \text{ ma} ,$$

$$R_e = \frac{V_{bias}}{I_{dc}} = \frac{1.1}{5} = 220 \text{ ohms} .$$

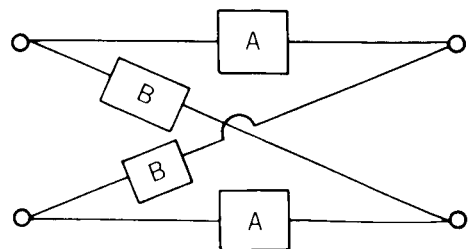
Modulator



The phase modulator uses voltage-variable capacitors in a transformation of a lattice all-pass network (Reference 2). The circuit development follows the following steps:

$$A = jX ,$$

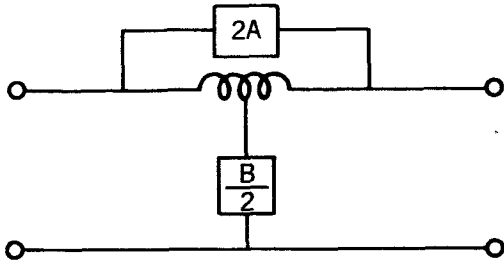
$$B = \frac{Z_0^2}{A} .$$



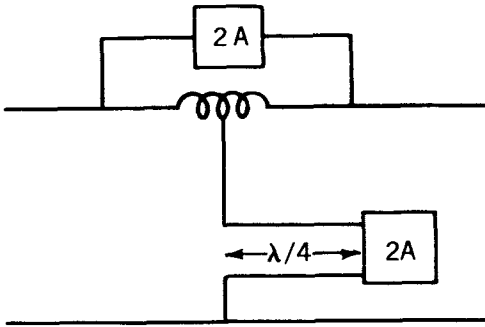
which can be transformed to

$$\alpha = 0,$$

$$\beta = 2 \tan^{-1} \left(\frac{X}{Z_0} \right).$$



To make a phase modulator of this network, the phase must be varied electronically. This can be done by making A and B electronically variable; but A and B are opposite reactances, between which a reciprocal relation must be maintained. These conditions can be met, however, if a quarter-wave transmission line of impedance Z_0 is inserted in B:



If A is a variable capacitance diode, the phase can be made electronically variable. By shunting the diode with some fixed inductance, the non-linear C versus V function can be made to compensate for the tangent phase-function nonlinearities. Linearity can be maintained for a modulation index of well over 1 radian. In a practical circuit, the quarter-wave transmission line is made from lumped elements, and the transformer provides the shunt inductance.

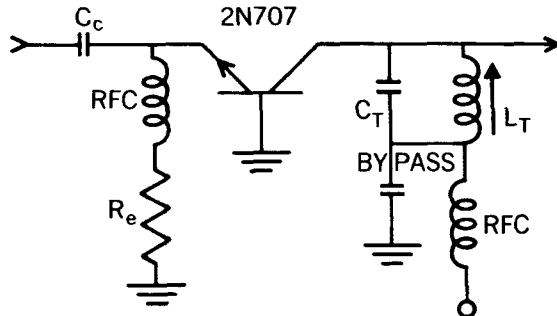
Since the reverse bias voltage on the diodes is only a few volts, the power handling capacity of the modulator is small if the diodes are not to conduct. The maximum RF signal power is:

$$P = \frac{V_d^2}{2R_0} \text{ watts.}$$

Obviously, if the device is to handle even a few tens of milliwatts, its characteristic impedance must be low. A practical Z_0 is on the order of 100 ohms.

For the reason just stated and because the frequency multipliers will also multiply the modulation index (thus requiring less from the modulator), it is desirable to place the modulator as near the oscillator as practical. On the other hand, modulating too close to the oscillator will cause incidental FM from lack of isolation. In this transmitter it is placed between the first and second frequency doublers. Thus, the oscillator is protected by the buffer and first doubler. Yet the power level is still only 30 mw, and the second and third doubler will multiply the modulation index by a factor of 4. For an output index of 1 radian, only 0.25 radian is required from the modulator. Since $Z_0 = 100$ ohms and $V_{bias} = -4$ volts, then $P_{max} = 80$ mw and $C_c = 8.5 \mu\text{f}$. Because the loss = 2 db, then $P_{out} = 19$ mw.

Second Frequency Multiplier (Doubler)



Looking ahead to the requirement of a higher voltage supply for the final amplifier, it would be wise to use higher voltage transistors for all of the remaining stages. If this is done, only one supply voltage would be required for the transmitter since the three low level stages already designed could be zener-regulated at the lower voltage with only a small loss in efficiency. A regulated voltage for the oscillator and buffer would be desirable for stability anyway. A convenient value for the supply voltage might be +24 volts, since this seems to be almost a standard voltage in many electronic systems—particularly in military systems.

With this requirement in mind, a 2N707 transistor is a good choice. The field of possible tran-

sistors to choose from is restricted, by the upper temperature limit, to silicon devices, which have not quite caught up with the state of the art in germanium transistors. Unfortunately, very little information is available from the manufacturer about the 2N707 transistor; they state only that at 100 Mc it will develop 300 mw output with 7.5 db gain. With such inadequate data, the most reliable (and fastest) procedure is to breadboard the amplifier and determine the results experimentally. For a very rough approximation, R_{in} can be calculated from the data given and the assumption made that it is independent of frequency (which it is not).

$$R_{in} = \frac{R_L}{G_p} = \frac{V_c^2}{2P_o G_p} = \frac{(19)^2}{2(300)(6)} = 100 \text{ ohms} .$$

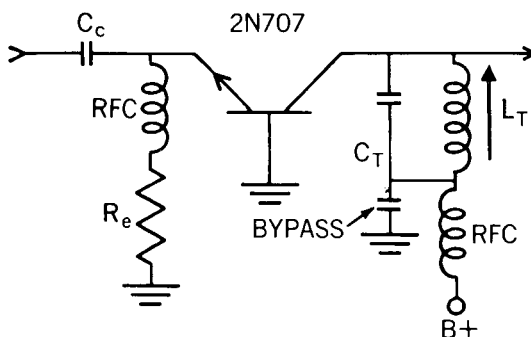
This impedance, conveniently, is just the right value for the modulator load, and no matching is required. Through a process like that used for the design of the first doubler, and allowing 2 db surplus gain for temperature and voltage variations, a rough design can be made for the second doubler:

$$\left. \begin{array}{l} P_{o,2} = 60 \text{ mw} , \\ R_L = 4.4 \text{ k} \\ C_T = 10.5 \mu\mu\text{f} \\ L_T = 0.5 \mu\text{h} \end{array} \right\} Q = 20$$

$$I_{dc} = 6 \text{ ma} ,$$

$$R_e = 47 \text{ ohms} .$$

Third Frequency Multiplier (Doubler)



Design for the third doubler proceeds exactly like that of the second doubler:

$$R_{in} \approx 100 \text{ ohms} ,$$

$$C_c = 3.6 \mu\mu\text{f} ,$$

$$P_{o,2} = 110 \text{ mw} ,$$

$$R_L = 2.4 \text{ K} ,$$

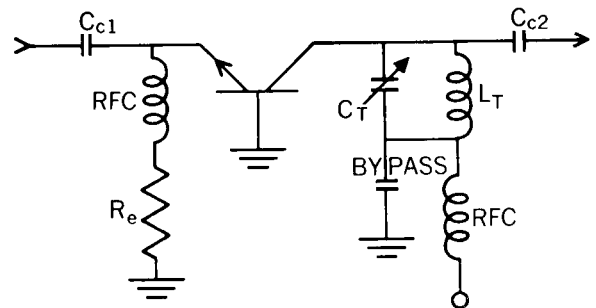
$$C_T = 10 \mu\mu\text{f} ,$$

$$L_T = 0.15 \mu\text{h} ,$$

$$I_{dc} = 11 \text{ ma} ,$$

$$R_e = 90 \text{ ohms} .$$

Final Amplifier



At this point in the design, we will jump to the final amplifier and work backward, the reason being that this stage must be designed for a specific predetermined output power. The parameters of the other stages must be adjusted to meet the requirements of this final amplifier. An RCA TA2267 transistor was selected for the circuit. This transistor has a collector breakdown voltage of 60 volts and would therefore work best with $V_{cc} = 30$ volts. Performance should be satisfactory at 24 volts, however.

Conveniently, the manufacturer provides measured curves showing the gain of the device versus frequency and output power. With a power output of 1 watt and a frequency of 136 Mc, these curves show a typical drive requirement of 250 mw

for class C operation. While the curves are measured for a common emitter configuration, the power gain should not be radically different for grounded base operation at this frequency. So, for this circuit,

$$R_L = \frac{(V_{cc} - V_{min})^2}{2P_o} = \frac{23^2}{2(1)} = 265 \text{ ohms} ,$$

$$X_T = \frac{R_L}{Q} = 26.5 \text{ ohms} ,$$

$$C_T = 44 \mu\mu f ,$$

$$L_T = 0.03 \mu h ,$$

$$X_{c,2} = \sqrt{R_L R_{out}} = \sqrt{265(50)} = 115 \text{ ohms} ,$$

$$C_{c,2} = 10 \mu\mu f .$$

Since G_p —as read from the curves—is for “class C” operation, the “class B” G_p is probably about 6.5 instead of 4. So,

$$R_{in} = \frac{R_L}{G_p} = 40 \text{ ohms} ,$$

$$X_{c,1} = \sqrt{R_{L-driver} R_{in}} = \sqrt{\frac{V_c^2}{2P_{in}} R_{in}} = 200 \text{ ohms} ,$$

$$C_{c,1} = 6 \mu\mu f .$$

Choosing

$$\theta = 140^\circ \text{ (see Reference 1),}$$

$$V_{in} = \sqrt{R_{in} P_{in}} = 3.1 \text{ volts} ,$$

$$V_{bias} = \sqrt{2} V_{in} \cos\left(\frac{\theta}{2}\right) - 0.7 = 0.8 \text{ volt} .$$

Assuming

$$\eta_c \approx 0.5 ,$$

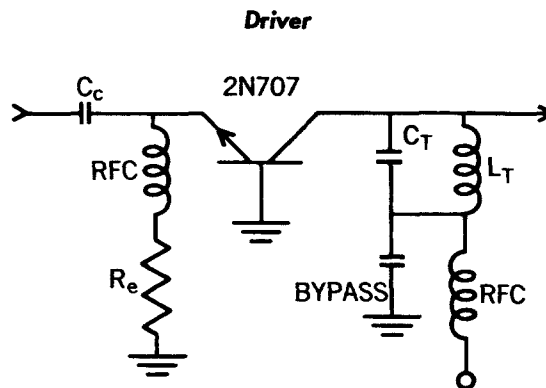
then

$$P_{dc} \approx \frac{P}{\eta_c} = 2 \text{ watts} ,$$

$$I_{dc} = \frac{P_{dc}}{V_{cc}} = \frac{2}{24} = 83 \text{ ma} ,$$

$$R_e = \frac{V_{bias}}{I_{dc}} = \frac{0.8}{0.083} = 10 \text{ ohms} .$$

Since the RF choke required in series with R_e has a resistance of 2 ohms, the actual R_e required is approximately 8 ohms.



Observing that the final amplifier requires 250 mw driving power and that the third doubler only puts out about 100 mw, it becomes obvious that an extra stage of amplification is required. Since another amplifier will certainly provide more than the 4 db of gain required, a considerable surplus of gain will be available. Because of the loss of efficiency it is not advisable to overdrive the final power amplifier, so the excess power gain must be taken up in lower power stages where the cost in efficiency will not be so high. Accordingly, the driver stage is designed for a power output of 250 mw:

$$P_o = 250 \text{ mw}, \quad f_0 = 136.5 \text{ Mc},$$

$$\eta_T = 0.9 \text{ (assumed)}, \quad R_{in} = 100 \text{ ohms (approx.)},$$

$$P_o' = 280 \text{ mw}, \quad G_p \approx \frac{R_L}{R_{in}} \approx 10 \text{ db (class B),}$$

$$R_L = 950 \text{ ohms}, \quad P_{dc} \approx \frac{P_o}{0.5} \approx 500 \text{ mw},$$

$$C_T = 12 \mu\mu f,$$

$$L_T = 0.1 \mu\text{h}, \quad I_{dc} \approx \frac{P_{dc}}{V_{dc}} = \frac{500}{24} = 21 \text{ ma.}$$

Since so much surplus gain is available, the driver can be operated with a very short conduction angle to trade some of its extra gain for better efficiency. Let

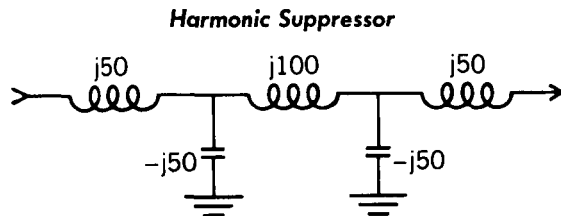
$$\theta = 100^\circ,$$

$$V_{in} = \sqrt{2P_{in} R_{in}} = \sqrt{20} = 4.5 \text{ volts peak},$$

$$V_{bias} = 4.5 \cos(50^\circ) - 0.7 = 2.2 \text{ volts},$$

$$R_e = \frac{V_{bias}}{I_{dc}} = \frac{2.2}{0.021} = 100 \text{ ohms}.$$

The rest of the extra gain should be used in the frequency multiplier stages in a similar fashion by increasing the bias on these stages and therefore decreasing their conduction angle.



Since transistors are inherently such good harmonic generators, some means must be provided to attenuate this spurious radiation to prevent interference with other services. Good harmonic suppression can be achieved simply by insertion of an artificial half-wave transmission line in the transmitter output. At the output frequency, the antenna load is transmitted faithfully through to the antenna and the output energy is delivered to the antenna with negligible loss. At the harmonic frequencies, however, the artificial line is no longer matched nor is it still one-half wave long. The harmonic energy is attenuated by the mismatch thus produced. At the same time, the impedance of the shunt capacitors in the line is decreased and the impedance of the series inductors is increased by the higher frequency of the harmonics. This action tends to bypass what

harmonic energy is in the line to ground. Actual measured values for one of these artificial lines are:

Harmonic (N)	Freq. (Mc)	Loss (db)
1	136.5	0.2
2	273	33
3	410	53
4	548	62
5	685	56
6	820	58
7	955	61

The flattening of the attenuation curve at the fourth harmonic is due to series resonance of the standoff-type shunt capacitors with their lead inductance and capacitive coupling around the series inductors. An identical unit built with shielded sections and feed-through capacitors measured 84 db rejection at 955 Mc, and was still increasing. But, since harmonics of such high orders should be negligible in the transmitter output, the extra attenuation provided is probably not worth the increased complexity of construction.

Mechanical Considerations

To provide good shielding, a compartmented chassis milled from an aluminum block is used, as shown in Figures 12(b) and (c). The compartment walls, in addition to shielding the various stages, also stiffen up the whole chassis against mechanical vibration. Printed circuit construction is used where feasible to increase ease of fabrication and uniformity of performance. For better resistance to mechanical stress and vibration, the whole chassis can be filled with a low dielectric encapsulating material, such as Eccofam FP. Repairs or modifications of a completed unit can then most easily be made by replacement of whole modules, although the potting material can be removed by careful picking with a small knife or scribe.

Thermal Problems

Since equipment in a satellite or space vehicle must operate in a vacuum, special attention must

be paid to the problem of removing the heat dissipated by the transistors. With no air to circulate, heat can be removed only by radiation or direct conduction; but, since

$$P_{\text{rad}} \sim \Delta T^4,$$

the radiation mechanism is efficient only at elevated temperature gradients. The only practical method for heat transfer in the unit is therefore direct conduction. A thermal path must be provided from each transistor case to the aluminum chassis, which should be fastened directly to the spacecraft structure. Thermal paths can be supplied on the printed circuit boards by the copper lamination, to which the transistor case can be fastened through a clamp of some fashion. If the transistor case is common to one of the leads, electrical isolation must be provided by the heat

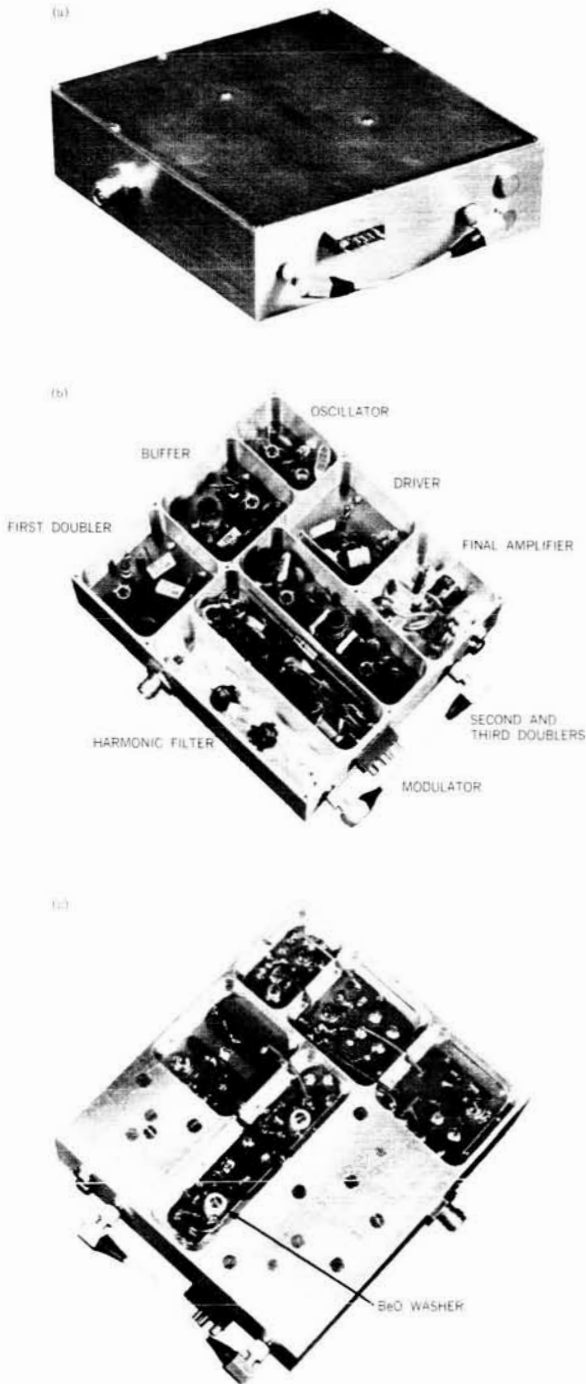


FIGURE 12.—Transmitter: (a) oblique view; (b) top view; (c) bottom view.

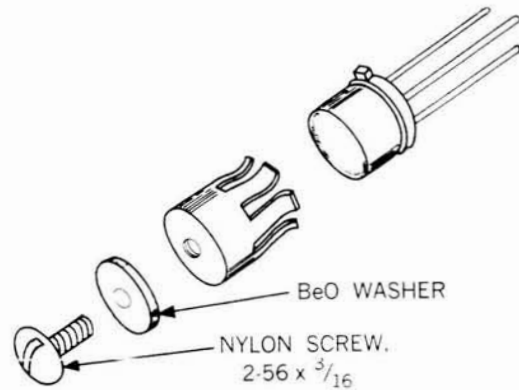
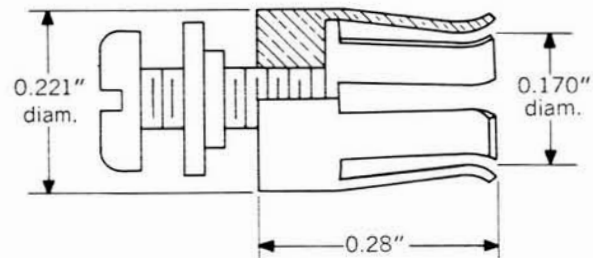


FIGURE 13.—Heat dissipator and retainer for TO-18 transistors.

sink without degrading the thermal path too much. An excellent material for use in such a situation is beryllium oxide, which has a coefficient of thermal conductivity about equal to that of aluminum and yet demonstrates a high resistivity

and low dielectric loss. Much use has been made in the past of thin slices of mica, Teflon, or Mylar for electrically insulated heat sinks. Although the thermal conductivity of these materials is poor, the fact that they are cut very thin limits the extent of the thermal path through them and therefore the temperature gradient. The very low thickness of the material causes a large capacity from the transistor case to ground, however, which limits their application in VHF circuits.

In this transmitter, the transistors are mounted in a clamp of the type shown in Figure 13. The clamp is then bolted to the copper lamination of the printed circuit board with a BeO washer as a spacer. The clamp is mounted as near to the chassis as feasible to keep the thermal path short (see Figure 14). Assuming that the thickness of the copper is 2/1000 inch and the thermal path length in the copper is 1/4 inch, the "thermal resistance" is

$$R = \frac{\Delta T}{P} = \frac{l}{kA}$$

where

l = path length in cm,

A = area in $\text{cm}^2 = \text{wt}$,

k = coefficient of conduction,

$$R = \frac{0.25(2.54)}{2(2 \times 10^{-3})(0.25)(2.54)^2} \approx 100^\circ\text{C/watt}.$$

Obviously, this method of heat sink is adequate only for stages dissipating 100 mw or less. The temperature gradient across the other components of the heat path can be neglected, compared with this relatively high value. To provide the required heat capacity, the higher powered stages are clamped directly to the chassis. This technique complicates the construction procedure somewhat but is necessary to keep the transistors adequately cooled. Where large amounts of heat are dissipated, even the areas of contact between the heat-sink components can become critical heat problems. In a vacuum, with no air to fill the microscopic voids in the contact surfaces, the actual area of conduction may be only 1 percent of the geometrical area. To reduce this problem, all areas of contact in a heat sink for a high power transistor should be cemented.

As an example of the thermal design of a transistor amplifier, we will examine the third doubler stage:

$$R_{\text{transistor}} = 150^\circ\text{C/watt} ,$$

$$R_{\text{heat sink}} = 100 ,$$

$$R_{\text{total}} = 250 ,$$

$$P_{\text{dissipated}} = P_{\text{dc}} - P_o = 260 - 110 = 150 \text{ mw} ,$$

$$\Delta t = RP = 250(0.15) = 37.5^\circ\text{C} ,$$

$$T_j (\text{max. junction temp.}) = 175^\circ\text{C} ;$$

therefore,

$$T_{\text{chassis max.}} = T_j - \Delta T = 137.5^\circ\text{C} .$$

A reliable design always will include safety factors so that degradation of components, accidental operation just outside of the specified temperature range, and unforeseen operating conditions might not cause failure of the transmitter.

The final amplifier should always be designed so that it can dissipate the full dc input power as heat, in case the unit is operated without the antenna connected or under conditions of high standing wave ratio.

Table 1
Transmitter Performance.*

T(°C)	P ₀ (dbw)	I (ma)	Δf ₀ (cps)	FM (cps)	AM (%)
0	-0.7	140	+450	0	5
10	-0.6	142	+403	1	3
25	-0.5	148	-78	1	0
40	-0.4	150	-200	2	0
60	-0.2	155	-350	1	4
80	-0.2	152	-174	2	4
100	-0.3	148	+840	2	2

*f₀ = 136.500 Mc
η = 28%

θ_p = 1 radian
V_p = +24 volts ±10%

Weight = 247 gm, unpotted
Estimated weight potted <300 gm,

2nd harmonic = 55 db below carrier
3rd harmonic = 80 db below carrier

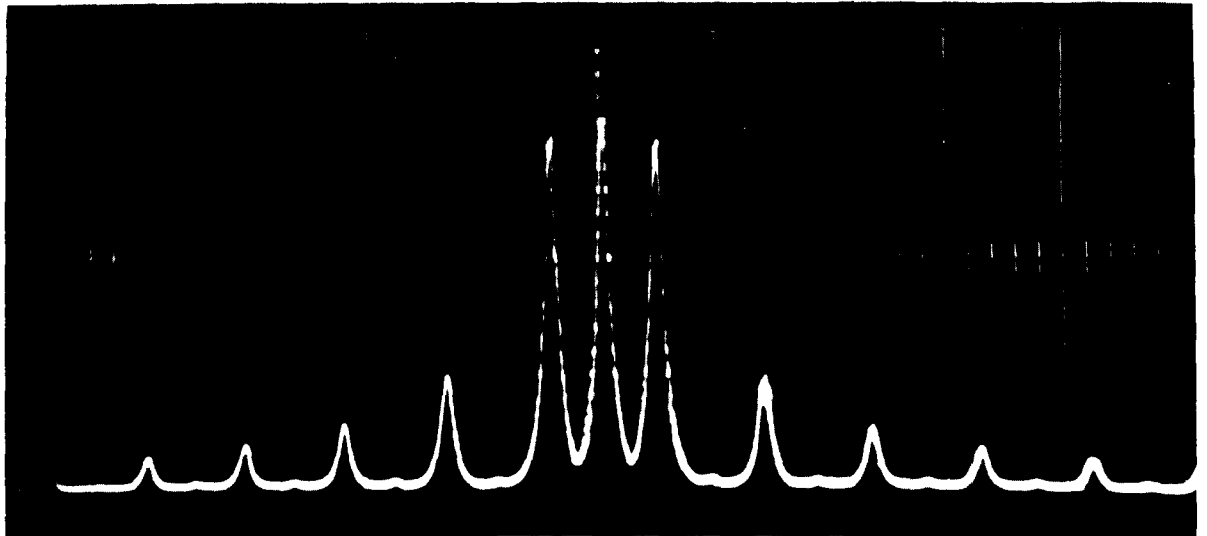


FIGURE 15.—Spectrum for square wave modulation.

REFERENCES AND BIBLIOGRAPHY

1. TERMAN, F. E., "Radio Engineering," 3rd ed., New York: McGraw-Hill, 1947.
2. FISHER, A. J., TALBERT, W. R. and CHITTENDEN, W. R., "Telemetry Transmitter for Radiation Satellite," *Electronics*, 33(19):68-69, May 6, 1960.
3. BENETEAU, P. J., "Design of High Power VHF Transistor Amplifiers," Fairchild Application Note APP-10, October 1959.
4. Lo, W., et al., "Transistor Electronics," Englewood Cliffs, N. J.: Prentice-Hall, 1955.
5. PIERCE, J., "The Design of VHF Power Amplifiers," Texas Instruments Application Note, April 1961.
6. RHEINFELDER, W. A., "Common Emitter vs. Common Base Operation," Motorola Application Note AN123.
7. "Radio Frequency Applications," Pacific Semiconductors Incorporated Power Transistor Application Note No. 2.

APPENDIX A

A_{eff}	effective capture area of antenna	R_L	collector load resistance
C	capacity	R_{out}	output signal resistance
C_c	interstage coupling capacity	R_o	real part of Z_o
C_T	tank circuit capacity	S	flux density
c	velocity of light	T	temperature
f	frequency	T_{chassis}	ambient chassis temperature
f_o	nominal frequency	T_J	maximum allowable transistor junction temperature
G_i	current gain	t	thickness
G_p	power gain	V_{bias}	required bias voltage
G_r	receiving antenna gain	V_c	peak collector signal voltage
G_T	transmitting antenna gain	V_{cb}	maximum collector-to-base voltage
$I_{c,1}$	fundamental frequency collector current	V_{cc}	collector supply voltage
$I_{c,2}$	2nd harmonic collector current	V_{ce}	maximum collector-to-emitter voltage
I_{dc}	direct current collector current	V_d	diode reverse bias voltage
I_{in}	input signal current	V_{diode}	forward bias necessary to cause current to flow
I_{out}	output signal current	V_{min}	minimum collector voltage
jX	reactance	V_{peak}	peak voltage
k	constant of proportionality	W	width
L_T	tank circuit inductance	X_C	reactance of capacitor in tank circuit
l	path length (cm)	X_c	reactance of coupling capacitor
P	power	X_L	reactance of inductor in tank circuit
P_c	power in carrier only	X_T	reactance of tank circuit elements
P_{dc}	direct current power	Y_{ib}	common base input admittance
P_{in}	driving power	Z_o	characteristic impedance
P_o	output power	a	attenuation
$P_{o,1}$	fundamental frequency output power	β	phase shift
$P_{o,2}$	2nd harmonic output power	η_c	collector efficiency
P_r	receiver power	η_T	tank circuit efficiency
P_{rad}	radiated power	θ	modulation index of phase modulation (pages 9–11, 28); conduction angle (pages 16, 21–22)
P_T	transmitted power	λ	wavelength
P_{total}	power in carrier and sidebands combined	ν	velocity of satellite
P_1	power in principal sideband	π	2.1416, approximately
$P_{>1}$	power in sidebands or order greater than 1	ϕ	angle between observer and satellite
Q	quality factor	ω_c	angular carrier frequency
Q_{loaded}	load Q of a tank circuit	ω_m	angular modulation frequency
Q_T	unloaded Q of a tank circuit		
\mathcal{R}	thermal resistance		
R_o	emitter bias resistor		
R_{in}	input signal resistance		

ON THE DESIGN OF PFM TELEMETRY ENCODERS*

H. D. WHITE, JR.

Pulse Frequency Modulation (PFM) is a time-division-multiplex telemetry system, especially suitable for small satellites because of its efficient use of transmitter power as a function of bit rate. The Explorer XII (Energetic Particles) satellite utilized a practical PFM system in which the analog-to-digital conversion was done essentially on the ground.

The encoding techniques used in PFM satellite telemetry systems are described herein, with emphasis on the simplicity of synthesizing and mechanizing the encoding. Several encoders are discussed (e.g., the Explorer XII and the UK-1 encoders, and a general-purpose digital data processor). In addition, checkout and bench-calibration techniques that facilitate integration between experiments and the encoder are included.

INTRODUCTION

Pulse Frequency Modulation (PFM) is a time-division-multiplex telemetry system, especially suitable for small satellites because of its efficient

use of transmitter power as a function of bit rate. A typical PFM encoder receives both analog and digital information from the satellite experiments (Figure 1). The encoder electronically commu-

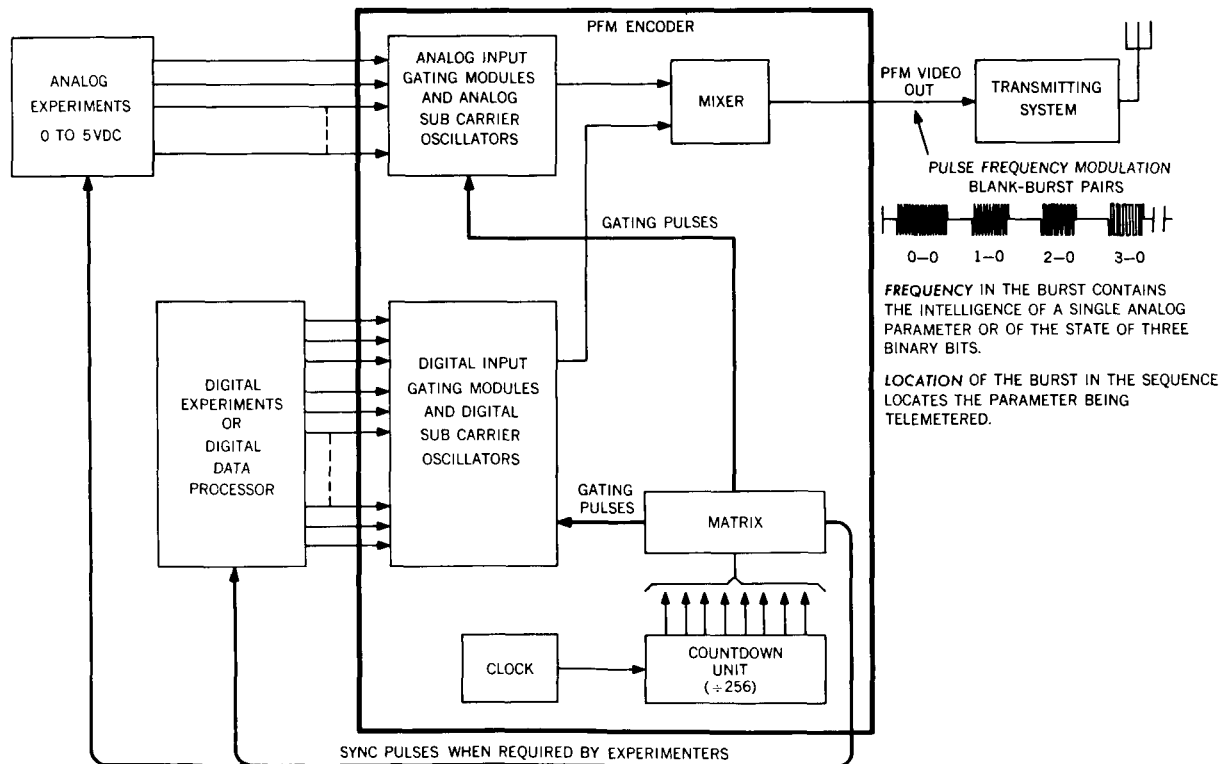


FIGURE 1.—PFM encoder.

*Published as *NASA Technical Note D-1672*, June 1964; also published in the *Proceedings of the 1962 National Telemetry Conference*, Sheraton-Park Hotel, Washington, D.C., May 23, 1962, 1 (3-4).

tates the information, either one analog parameter or 3 digital bits at a time, and encodes it into a series of time-multiplexed PFM bursts and blanks. The pulsed frequency bursts are derived from a set of pulsed subcarrier oscillators with a variable frequency range of 5 to 15 kc. The burst frequency contains the intelligence of a single analog parameter or of the state of 3 binary bits. The location of the burst in the telemetry sequence identifies the parameter being measured. The encoder also generates any required synchronizing (timing) pulses for the experiments.

The timing for the PFM system is obtained by counting down a clock frequency in a string of binary counters (called a countdown unit). The states of the binary counters in the countdown unit are then sampled by appropriate logical circuit elements in a matrix to produce the required commutation-gating pulses and synchronizing pulses.

From the theoretical signal-to-noise standpoint, the PFM system has certain marked advantages over some of the other commonly used satellite telemetry systems (Reference 1). In addition, this system has practical advantages in the simplicity of satellite encoding, which will be illustrated in the discussions that follow.

BASIC PFM SYSTEM AND FORMAT

General Description

A characteristic of the PFM system is that both the basic format and bit rate may be changed easily to accommodate different satellite missions.

The basic PFM telemetry system is a particular form of time-division multiplexing in which the intelligence being telemetered is contained in the frequency of a sequential series of 10-millisecond pulses (*bursts*) separated by 10-millisecond intervals (*blanks*). The blanks precede the bursts. The pulse frequency is derived from a set of pulsed variable-frequency subcarrier oscillators, each having a frequency range of 5 to 15 kc.

The satellite telemetry system works in conjunction with a ground data-reduction system in which a bank of 120 contiguous bandpass filters, each 100 cps wide, are used to reduce the system noise bandwidth to that approximating one of the 100-cps filters. The entire bank of filters covers the frequency range of 4 to 16 kc.

Essentially, the PFM burst activates one of the 120 contiguous filters. If only one pulse frequency is present at a time, only one of the 120 filters will be excited. The filter having the greatest signal is determined by means of an auction circuit. All the other filters, containing noise only, are biased off; and the one filter containing signal plus noise is gated to the output. With reasonable signal-to-noise conditions, the filter output can be passed on to a discriminator for making more precise measurements; the magnetometer data of Explorer X (1961 κ) were sometimes read to a precision of 10 bits by this technique.

For the general use of PFM, a discriminator is not used; the signal is quantized into one of the hundred discrete 100-cps values in the 5 to 15 kc range by recording only the number of the particular filter that contains the greatest signal. Thus the basic analog accuracy of the standard PFM system is approximately 1 percent from the input of the satellite encoder to the recorded filter number on the ground. For the PFM encoders presently on board satellites, a 1-percent accuracy of analog information channels is the best that can be achieved. But, as noted, a 1/10-percent precision is possible with the use of discriminators at the filter band output.

Some interesting features of this system become apparent. If a 1-percent accuracy is assumed equivalent to almost 7 bits, then analog information is sent 7 bits at a time in a single PFM burst and, effectively, the analog-to-digital conversion is done on the ground. In regard to the satellite telemetry system, two important considerations result:

- (1) The complicated analog-to-digital conversion is not required within the satellite.
- (2) For a given bit rate, the transmitter power may be substantially reduced over that required for the same accuracy in an un-coded PCM system.

Analog Subcarrier Oscillator

The analog oscillator (Reference 2) is the name applied to the device that converts an input voltage of 0 to +5 volts* dc into a frequency output of 15 to 5 kc. Zero volts corresponds to approxi-

*Or 0 to -5 volts dc in the UK-1.

mately 15 kc, while 5 volts corresponds to approximately 5 kc. The analog oscillator is designed to operate in the temperature range of -10°C to $+60^{\circ}\text{C}$.

Several inputs may be gated into a single oscillator. To maintain a 1-percent system accuracy, it is necessary to calibrate each analog experiment with its associated oscillator. The input impedance to the oscillator, when several inputs are commutated into it, is approximately 250 kilohms at approximately -2.3 volts when the analog parameter is being telemetered, with a slightly active network at a high impedance when the oscillator is gated off. If the output impedance of the experiment is reasonably low (less than 1 kilohm), the input impedance of the encoder has a negligible effect on the output frequency of the oscillator. The input impedance of the analog oscillator, when only one input is applied (e.g., no commutation of inputs), is greater than 2 megohms at approximately $+7$ volts when the analog parameter is being telemetered, with a slightly active network at a high impedance when the oscillator is off.

The usual procedure followed in analog oscillator allocation is to give each experimenter his own oscillator. If the experimenter has many analog inputs (as in the case of the UK-1), the oscillator(s) is placed in the experimenter's package (approximately one for each eight inputs); this reduces the interface and calibration problems. When an experimenter has only a few inputs (as in Explorer XII† or the Interplanetary Monitoring Probe), the oscillator is placed in the encoder; this reduces the number of leads required. A main criterion for the oscillator is that its most likely failure be *off*, instead of *on*. Thus, when each experimenter has his own oscillator, a catastrophic event in an experiment affects only that one experiment. If an experimenter desires 1-percent analog data, the interchange of oscillators requires a recalibration.

Digital Subcarrier Oscillator

A pulsed digital oscillator that accepts 3 bits of information and encodes this as one of eight discrete frequencies in the band from 5 to 15 kc has been developed at Goddard Space Flight Center. The eight possible frequency levels correspond to

†(1961 v 1) satellite.

the eight possible combinations of 3 bits. Thus digital information already in a digital form may be easily encoded and transmitted 3 bits at a time.

Digital-input gating modules have been developed to accept accumulator outputs 3 bits at a time, read them out as a discrete frequency, and then take the next 3 bits and read them out (18-bit and 15-bit accumulators have been read out in this manner in the UK-1 encoder). To use this feature, the two binary states must be 0 volts at less than 56 kilohms output impedance, and -4 to -5 volts at less than 56 kilohms output impedance from the experiment's accumulator. In the Interplanetary Monitoring Probe (IMP), the accumulation and read-out is done in the encoder in a digital data processor. The digital data are read out 3 bits at a time in a PFM burst.

Basic Format and Frame Synchronization

As previously stated, PFM is a time-division-multiplex system. The basic format consists of 256 channels arranged 16 channels in a frame, with 16 frames in a sequence. Frame synchronization is achieved by two methods, as follows:

1. The first channel of all frames (designated as 0-all) consists of an asymmetric blankburst interval, where a 5-millisecond blank is followed by a 15-millisecond burst. All the other channels consist of symmetric blank-burst intervals where a 10-millisecond burst follows a 10-millisecond blank. The system is synchronous in that all blank-burst intervals total 20-milliseconds.

2. The frequency in channel zero for all frames (0-all) is also devoted to frame synchronization. The exact format is as follows:

Channel-Frame	Information Telemetered
0-0	000 digit
0-1	4.5 kc
0-2	001 digit
0-3	4.5 kc
0-4	010 digit
0-5	4.5 kc
0-6	011 digit
0-7	4.5 kc
0-8	100 digit
0-9	4.5 kc
0-10	101 digit
0-11	4.5 kc
0-12	110 digit
0-13	4.5 kc
0-14	111 digit
0-15	4.5 kc

Since each channel takes 20 milliseconds, a complete telemetry sequence of the 256 channels takes 5.12 seconds.

Information Nomenclature

An X-Y nomenclature is used to designate the location of a sample in the sequence, where X is the channel number and Y is the frame number. For example, "0-all" means channel 0 for every frame and "3-4" means channel 3 during frame 4. The 16 channels and the 16 frames are each labeled 0 through 15. This nomenclature also is used in the location of synchronizing pulses from the encoder.

Information Rate

In the basic PFM system the sampling rate is 50 cps, which amounts to 20 milliseconds for a blank-burst pair. Since an analog channel contains approximately 7 bits (1-percent accuracy through the system) and a digital channel contains 3 bits (1 of 8 possible frequencies), the information rate depends on the ratio of analog channels to digital channels. The following examples are given:

Channel Ratio	Information Rate
All analog	$= 50 \times 7 = 350$ bits/sec
All digital	$= 50 \times 3 = 150$ bits/sec
$\frac{1}{2}$ analog	$\left. \begin{array}{l} \frac{1}{2} \text{ analog} \\ \frac{1}{2} \text{ digital} \end{array} \right\} = 50 \times 5 = 250$ bits/sec
$\frac{1}{2}$ digital	

Synchronizing Pulses Supplied to the Experimenter

No standard synchronizing (timing) pulses are available in a given encoder because the possible combinations of pulses are extremely large. The pulses from the digital data processor to the experimenter constitute the only exception.

For a given satellite encoder, the experimenter's requirements are analyzed during the initial design planning stage and the experimenter requests those "sync" pulses necessary to program his experiment.

For the most part, any synchronizing pulse can

be easily generated in the encoder if it is to start at the beginning of any channel (i.e., the beginning of the blank) and end at the end of any channel. These pulses are, of course, synchronized with the encoder. The experimenter must specify his requirements at a suitable stage of the design because changes are very difficult after the encoder is built. The use of welded modules, as discussed in the next section, limits changes.

Construction Utilizing Standard Modules

The construction now used in the basic PFM encoder has been the result of a gradual development. The first PFM modules, used in Explorer XII, were printed-circuit modules of binary counters and subcarrier oscillators. Many of the components, however, were directly mounted on printed-circuit boards. The UK-1 encoder, which followed, was so complex that the packing density had to be substantially increased; and therefore an almost all welded-module construction was used. The welded modules were mounted by soldering to the printed-circuit boards.

Certain advantages of welded modules other than increased packing density became apparent. One of these advantages was that each module was separately checked out with dynamic tests so that, when the modules were interconnected, every individual component had been dynamically tested before final test. This reduced the system test time considerably and increased reliability in that cards, once assembled, never have to be disassembled because of the initial installation of a faulty component. Another advantage that proved significant was that the increased packing density resulted in fewer circuit boards which, in turn, resulted in fewer interconnecting cables.

The disadvantage of welded modules is that changes and repairs are not easily made. Fortunately, the encoder contains very few critical circuits and will operate over wide changes in parameters. Standard modules are now available, and more have been developed for the Interplanetary Monitoring Probe (IMP). The welded technique construction will be carried one step further on the IMP encoder than was done on the UK-1 encoder: All interconnections between the welded modules will be welded instead of soldering the welded modules on printed-circuit boards.

PFM ENCODER FOR EXPLORER XII

Four digital oscillators and five analog oscillators were used in the Explorer XII encoder (Figure 2). One oscillator was used for each experiment, and two oscillators were used for the 16 performance parameters. Since eight of the 16 channels were digital and the other eight were analog, the bit rate was approximately 250 bits per second. Neither welded construction nor modular techniques were extensively used.

Commutation of analog input lines to a single analog oscillator was achieved. The three magnetometer inputs were gated into the magnetometer analog oscillator, and the three ion-and-electron (I&E) inputs were gated into the I&E analog oscillator. Performance parameters PP_0 through PP_{15} were split up, with PP_0 through PP_7 gated into a single analog oscillator and PP_8 through PP_{15} gated into another analog oscillator. The clock was not crystal controlled, and the countdown unit and matrix were both made exclusively of germanium transistors. Figure 2 shows

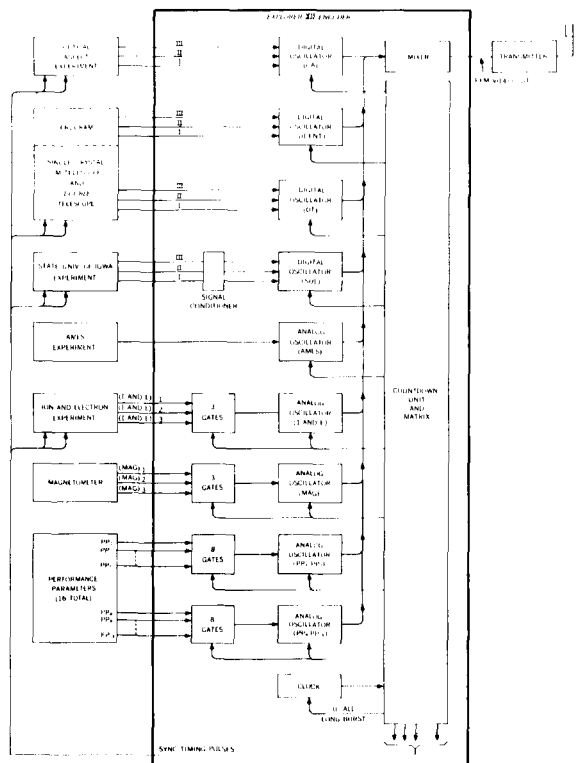


FIGURE 2.—Explorer XII PFM encoder.

that each digital experiment supplied only three input leads to the encoder. Thus the digital experimenters commutated the outputs of their accumulators, 3 bits at a time, into the encoder.

The encoder was packaged in three 5×7 inch printed-circuit cards. One additional 5×7 inch card was used to package the performance parameter signal-conditioning circuitry and an encoder power converter.

Designed to operate in a temperature environment of from 0° to $+50^\circ\text{C}$ (tested from -10° to $+60^\circ\text{C}$), the power drain of the encoder and converter was approximately 270 milliwatts. The weight of the three encoder cards after potting, plus the encoder power converter, was approximately 3 pounds.

As Figure 2 indicates, the Explorer XII PFM encoder is relatively simple and illustrates the starting point of the basic PFM system. Figure 3 shows that channel zero was not devoted to frame synchronization.

The encoder operated well and was built on schedule. Much of the design philosophy was dictated by expediency in that little time was available for development. Thus much of the circuitry and packaging techniques were similar to those used in Vanguard III (1959 η 1).

PFM ENCODERS FOR THE UK-1 SATELLITE

Figure 4 is a block diagram of the UK-1 encoder, which was required to telemeter the information of 100 input lines from the various British experiments. In addition, two encoders were required: one called the high speed (HS) encoder, and the other called the low speed (LS) encoder. The output from the HS encoder directly modulates the transmitter (real-time data). The output from the LS encoder is tape-recorded for one complete orbit at $1/48$ the information rate of the HS system. On command from a ground station, the taped output is played back 48 times faster than the recorded speed. The taped playback then modulates the transmitter. Thus the output from both systems (HS and LS) will have the same bandwidth when received at a ground station.

Design Criteria

In order to meet the size and weight requirements of the UK-1 payload, the design approach

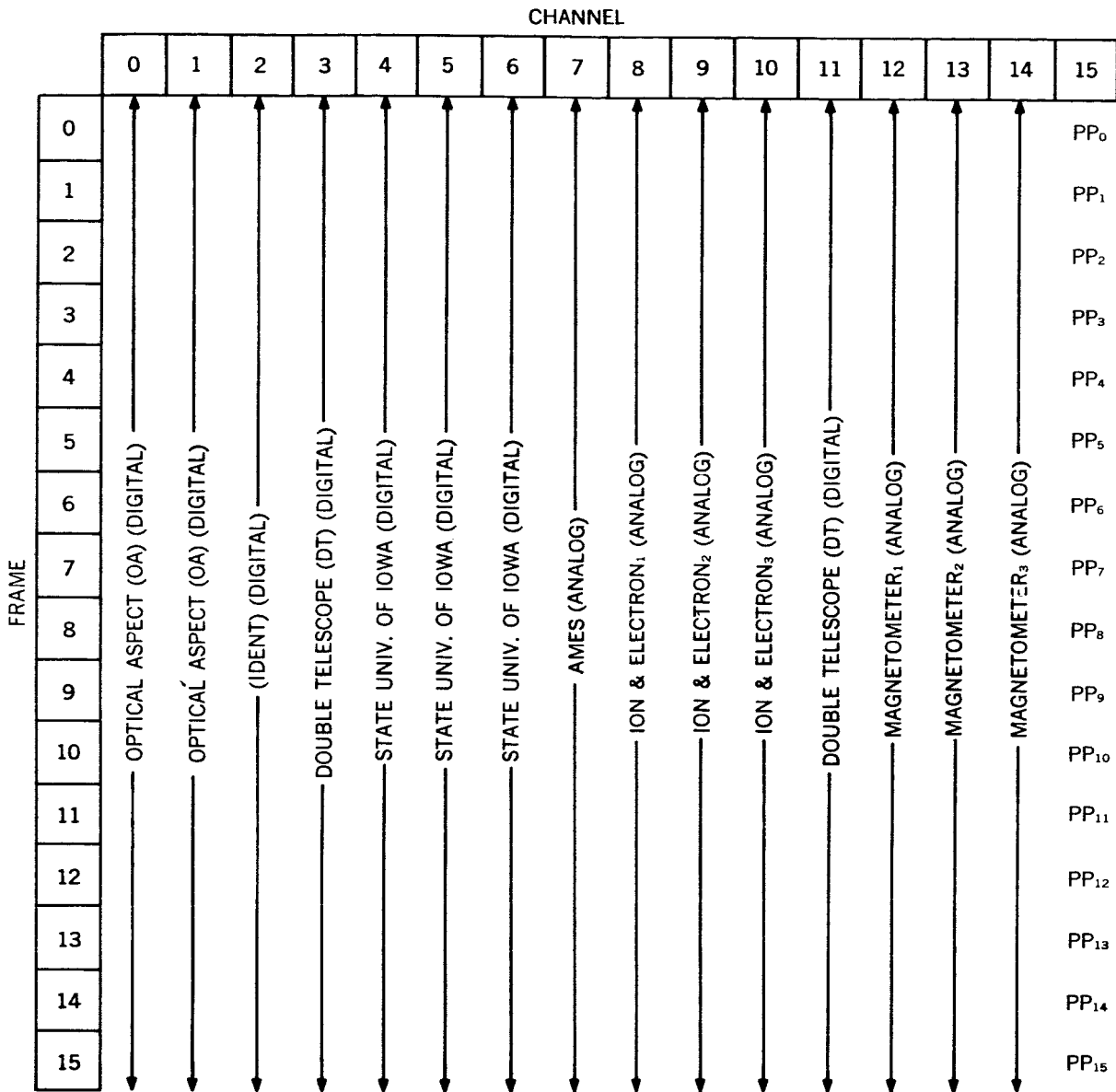


FIGURE 3.—Explorer XII telemetry channel allocation.

for this encoder had to differ somewhat from the encoder used in the Explorer XII. The payload experiments are divided among three British universities: The University College of London (UCL), the Imperial College (the cosmic-ray experiment), and The University of Birmingham (the electron density experiment). The number of information lines required from each experimenter was such that it was desirable to locate the

subcarrier oscillators and input-gating circuits in the experimenters' electronic stacks: two oscillator cards for UCL, one oscillator card for the electron-density experiment, and one oscillator card for the cosmic-ray experiment.

Figure 5 shows the placement of the subcarrier oscillator cards. This configuration has two advantages: (1) The information lines from the experiment remain short; and (2) the experiment

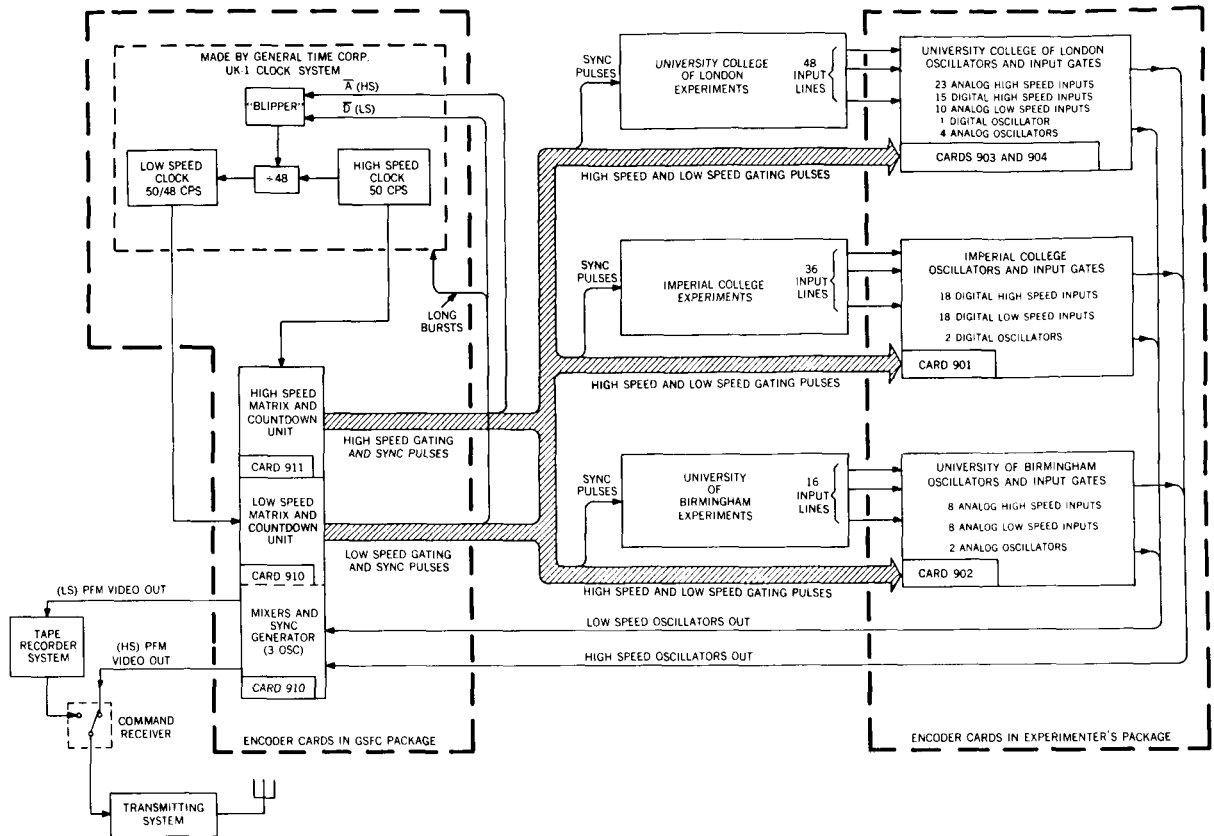


FIGURE 4.—UK-1 PFM encoder.

electronic stack is a nearly self-contained unit, which facilitates system integration.

During the initial design phases of the experiments and of the encoder, the Atlantic Ocean separated the two systems; thus an integration problem could easily have resulted. Matrix simulators (simple to build, ship, and use, but of unsophisticated design) were sent to the United Kingdom with appropriate pre-prototype subcarrier-oscillator cards and test specifications. The interface between the experiments and the encoder worked very well with this technique. Much of the success must be accredited to the outstanding cooperation and keen interest shown by the UK experimenters, particularly Dr. A. P. Willmore, of the University College of London.

Subcarrier-Oscillator Cards

Figure 6 illustrates most of the circuitry used in the subcarrier-oscillator cards. The card contains two analog-oscillator modules, one digital-oscillator module, one 8-negative analog-input-

gate module, one 7-negative analog-input-gate module, and one 15-bit digital-input-gate module. Thus the card handles the information from 15 analog lines and 15 digital lines from some of the UCL experiments. All gating modules are of welded construction and are mounted on a 5-inch-diameter printed-circuit card. The card size was chosen to fit with the rest of the cards in that particular UCL electronic stack. In the analog gating modules the n-p-n transistors perform the switching functions with about a 30-millivolt drop when *on* and essentially an open circuit when *off*. The p-n-p transistors invert the standard logic feeding the cards, which accounts for the fact that the analog lines from the UK-1 experiments go from 0 to -5 volts dc.

No measurable crosstalk has been observed from adjacent inputs when the system is in operation. There is, however, a 0.1 percent jitter in the encoding and decoding system that may mask slight crosstalk effects. The linearity and accuracy over the temperature range of 0° to +40°C

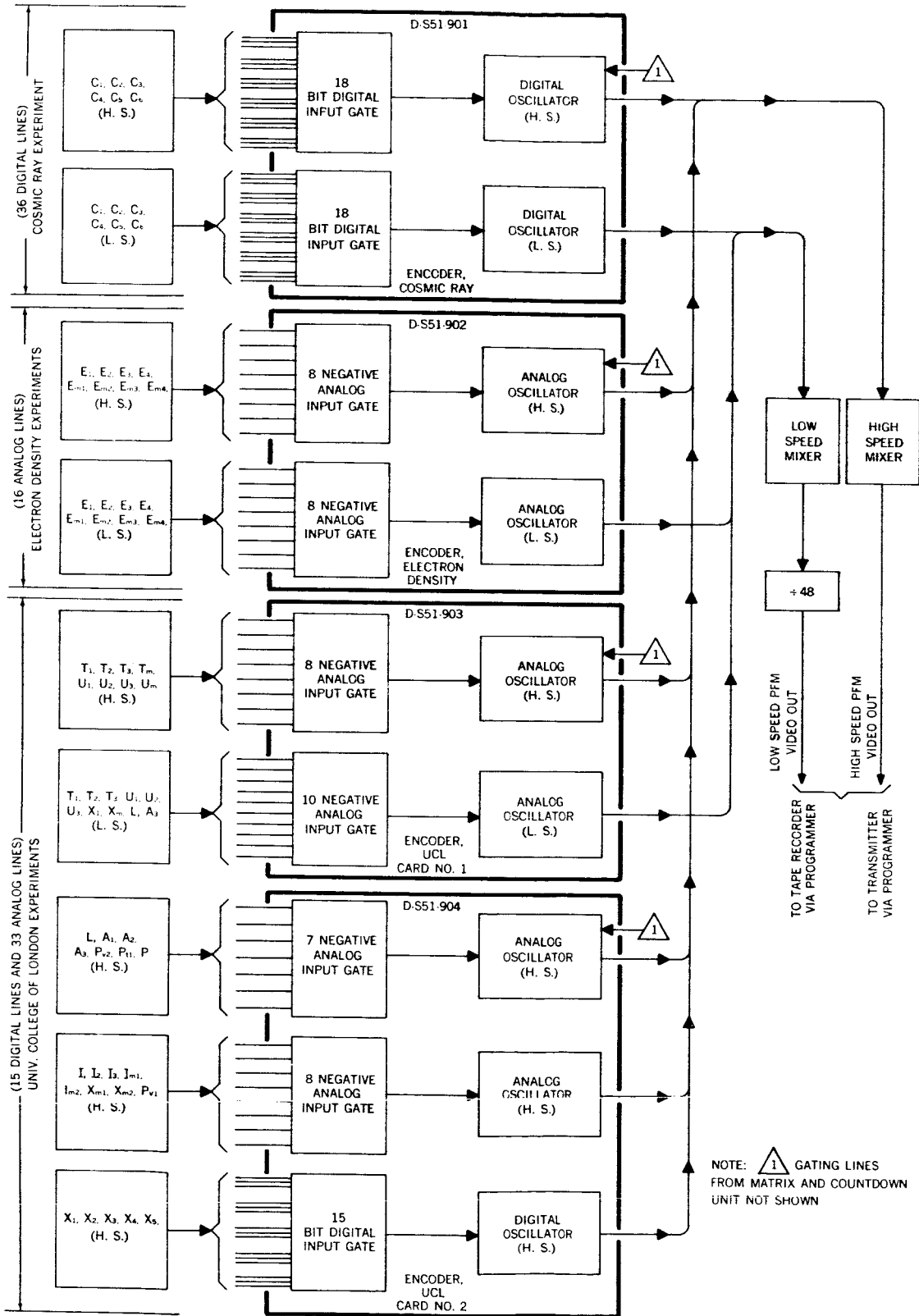


FIGURE 5.—UK-1 encoder cards located in experimenter's stack.

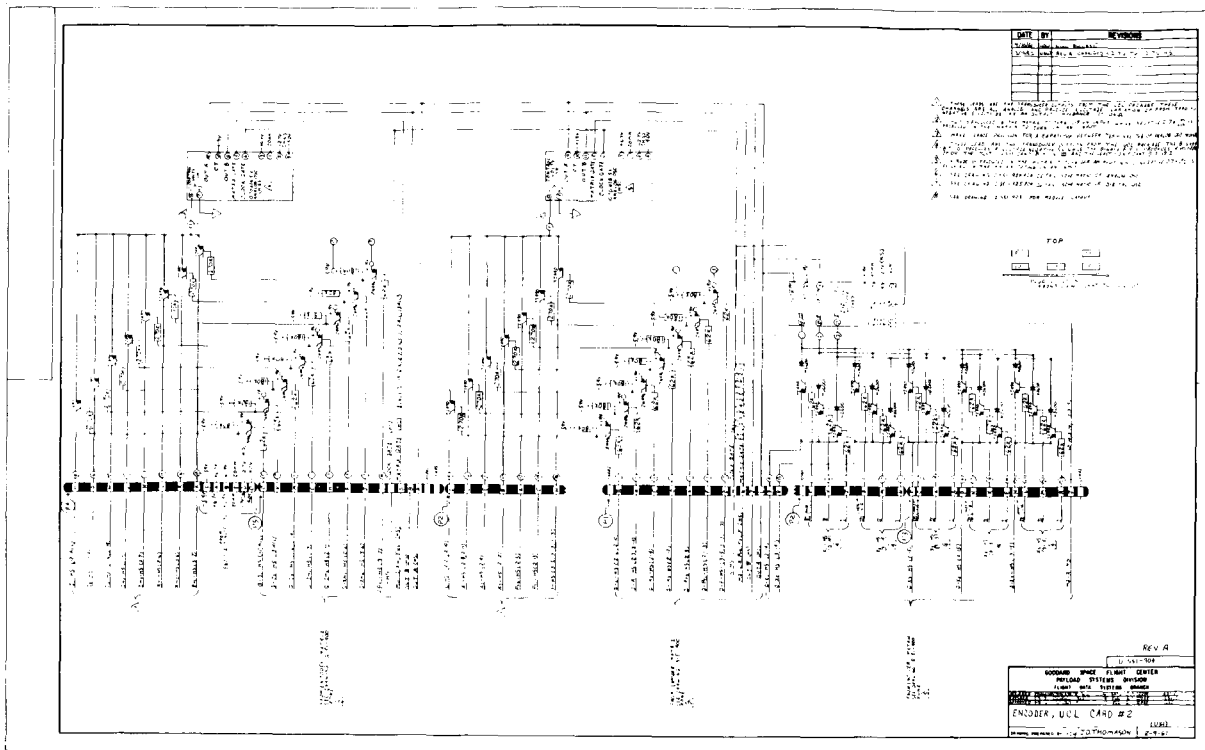


FIGURE 6.—UK-1 encoder, UCL card 2.

is of the order of 1 percent. The long-term repeatability (1 month) is of the order of $\frac{1}{2}$ percent, and the short-term repeatability is approximately that of the 0.1 percent jitter.

The method of gating does not provide a constant input impedance to the experiments, and it is not a passive network. It does, however, have the decided advantage of simplicity and, by making these cards an integral part of the experimenter's electronics package, some of the interface calibration problems are eliminated.

It should be noted that no common analog-to-digital converter is involved, since the quantization is done on the ground by the 120 comb filters. This entire job could, of course, have been done with four oscillators (one HS analog, one LS analog, one HS digital, and one LS digital) instead of the nine selected, by gating appropriate functions into one oscillator. This was not done for several reasons: First, it was desirable to give each experimenter his own oscillator (since only one oscillator is *on* at a time, there is not significant increase in power by having more than one oscillator); second, if it is assumed that the most likely

oscillator failure occurs when *off*, not *on*, no single oscillator failure would be catastrophic. In addition, if an n-p-n input-gating transistor shorts out, crosstalk will result among that channel and the others associated with its oscillator. The oscillator configuration was chosen, then, as a compromise between size and (hopefully) reliability.

The general approach toward reliability was not to design toward the minimum number of components but to design toward the minimum number of catastrophic components. The choice of oscillators is an example of this approach in that, if an oscillator fails when *off*, only those experiments associated with that oscillator are lost. The same is true in the case of a shorted input-gating transistor. It also should be noted that an open input-gating transistor will result in the loss of only the analog channel associated with it. In the five complete UK-1 encoder systems that have been built to date, the input-gating transistors have not opened; but on one occasion an input-gating transistor shorted, probably as a

result of an excessive voltage spike applied to the input channel.

Again, Figure 6 shows that 15 digital lines are gated 3 bits at a time and commutated into a single digital oscillator. This means that the accumulator for this particular experiment (UCL x-ray) is scanned 3 bits at a time by the encoder. The experimenter, therefore, was not required to shift out the contents of his accumulator 3 bits at a time as was done in the Explorer XII encoder. The next logical step will be taken in the Interplanetary Monitoring Probe (IMP) encoder, where the experiment accumulation will be done in the encoder. A digital data processor designed for this purpose will be described later.

Matrix Design

Figure 7 shows the schematic diagram of the UK-1 high speed (HS) matrix card, which functions to generate all the HS encoder gating functions and provide the necessary HS timing pulses required by the experiments. The card is entirely made up of welded modules, and all transistors are silicon. No attempt has been

made to design a standard matrix with a standard format for different satellites. The main reason for tailor-making the matrix for each payload is that the number of input lines to be telemetered and the required number of synchronizing pulses vary considerably with different satellites, so that it has been thus far impractical to attempt to design an efficient all-purpose matrix card.

The X-Y nomenclature (discussed in the section "Basic PFM System") makes for rapid matrix design, however, since the Boolean equations may be written, by inspection, from the desired telemetry format. More important, perhaps, is the fact that an efficient telemetry format may be derived by use of the X-Y nomenclature and its easy conversion to a Boolean expression during the initial planning phases of the satellite. For example, say it is required to generate a function for channel 8 on frame 9 only: The X-Y nomenclature would be 8-9 and, by inspection, the Boolean function is $E \bar{F} \bar{G} \bar{H} A \bar{B} \bar{C} D$, where $E \bar{F} \bar{G} \bar{H}$ denotes the channel location and $A \bar{B} \bar{C} D$ denotes the frame location. All the equations may be written in this manner and "reduced" by

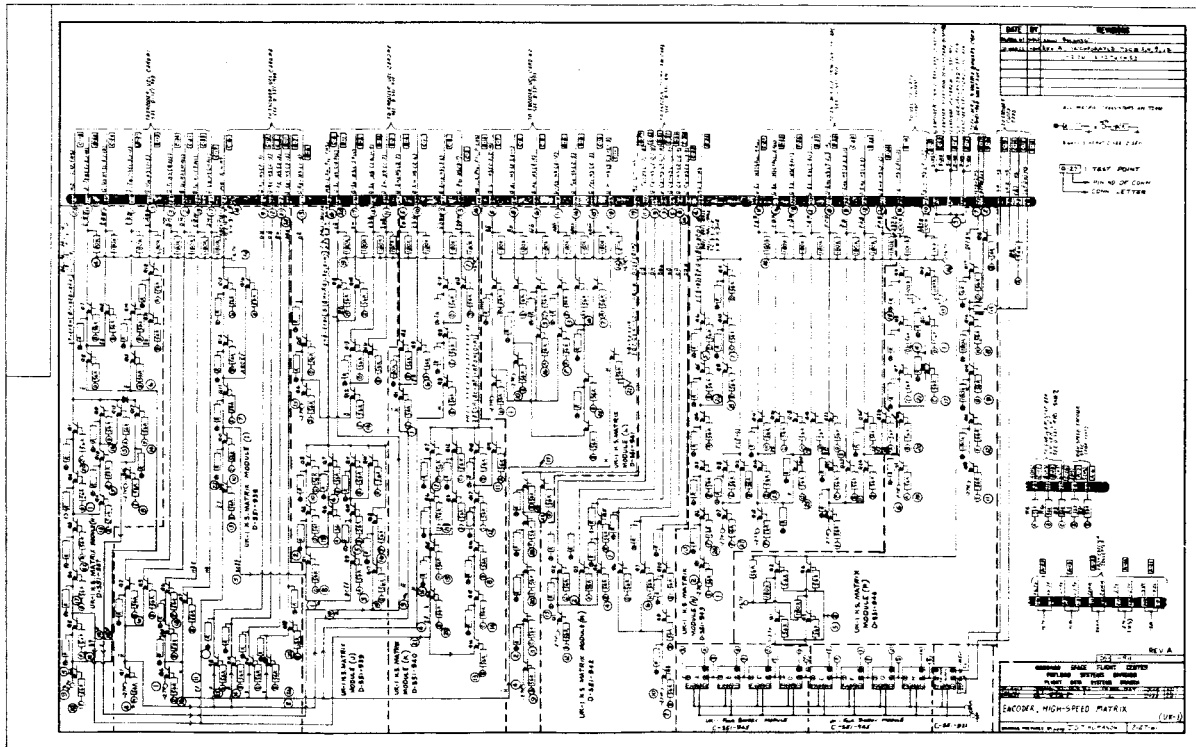


FIGURE 7.—UK-1 encoder, HS matrix card.

Boolean algebra. The equations are reduced, not in their simplest form, but in terms of the simplest form consistent with the circuit logical elements used.

Some techniques for reducing and checking the equations, which may or may not be new, have been developed for this type of logical circuitry. This paper will not dwell on the mechanization and synthesization of the equations except to say that Boolean algebra is a powerful tool in this application. As an example, the high speed and low speed matrix equations were written and mechanized in less than 30 hours. Thus the tailormade matrix may be rapidly and accurately designed, and has great flexibility in that any format may be efficiently put together. Whatever timing pulses the experimenters need may be generated for them. This is often of considerable value in overall payload reliability in that encoder timing pulses supplied to the experiment may reduce the complexity.

Clock and Low Speed Encoder Design

Since two encoders are utilized in the UK-1 satellite, reliability can be increased if the two encoders are made independent of each other (provided the experiments do not depend on both encoders operating simultaneously). In this system, only the UCL x-ray experiment required synchronization between the encoders. The design philosophy was to make the two systems independent and very loosely "sync" them through the clock circuitry. In other words, the high speed and low speed countdown units were not hard-coupled to each other.

The 48:1 relation between the LS and HS encoders was reached by assuming that the tape recorder for the LS encoder could be acquired by an interrogating ground station for at least 2 minutes during playback. The orbit takes approximately 100 minutes; thus, the tape should be played back approximately 50 times faster than it was recorded. The 48:1 relation was chosen because it is easier to work with than 45:1 or 50:1. In addition, since the LS encoder is of 2-frame duration (32 channels) and the HS encoder is 16 frames (256 channels), a 48:1 reduction will cause exactly six high speed sequences to occur for one low speed sequence; thus they may be synchronized. Figures 8 and 9 depict the telem-

etry channel allocations for the LS and HS encoders, respectively.

The circuit elements in the LS and HS encoder are identical and therefore use the same circuit design for flip-flops, logical elements, analog oscillators, and digital oscillators.

The clock signal to the LS encoder is made 48 times slower than the HS clock signal. The LS oscillators are mixed to form a composite video in the 5 to 15 kc range and are divided by 48 to supply the proper frequency deviation to the tape recorder. Thus the LS encoder is merely slowed down by a factor of 48.

The clock was designed and manufactured to Goddard Space Flight Center (GSFC) specifications by the General Time Corporation, Stamford, Connecticut. A simplified block diagram of the clock is shown in Figure 4. The clock is controlled by a 20-kc crystal and is divided down to synchronize a free-running multivibrator at a lower frequency. The signal is further divided to give the desired 50-cps HS clock rate, which is then divided by 48 to give the 50/48 LS clock rate. An additional input from a so-called "blipper" circuit is inserted in this last divide-by-48 circuit. An output from the "blipper" circuit occurs once every low speed sequence (approximately 30 seconds) when the encoders are not in synchronization, to advance the "sync" by 20 milliseconds. It may therefore take up to 2 hours to achieve synchronization. Once encoder synchronization is obtained, it remains in sync. This is what is meant by "loosely" synchronizing the encoders; it is the desired result. All flight-unit clocks for the UK-1 have functioned well.

Fabrication

The six UK-1 encoder cards and associated bench test equipment were designed, and pre-prototype (see Figure 10) cards were built at GSFC. Detailed electrical test specifications were written, and the pre-prototype cards were checked out and calibrated at GSFC. A prototype and three sets of flight unit cards (total of 24 cards) were built, electrically checked out, integrated, and calibrated to the GSFC electrical test specifications by Electro-Mechanical Research, Inc. This company also fabricated all the welded matrix modules, the subcarrier oscillators used in the prototype, and the three sets of

		CHANNEL															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
FRAME	0	S	X ₁	X _{m1}	E ₄	E ₁	E ₂	E ₃	X ₁	C ₁	C ₂	C ₃	C ₄	X ₁	A ₃	C ₅	C ₆
	1	S	X ₁	X _{m1}	E _{m4}	E _{m1}	E _{m2}	E _{m3}	X ₁	T ₁	T ₂	T ₃	L	X ₁	U ₁	U ₂	U ₃

Chart 3

FIGURE 8.—UK-1 low speed encoder, telemetry channel allocation.

		CHANNEL																
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
FRAME	0	S	C ₁	C ₂	C ₃	E ₃	E ₄	E ₁	E ₂	T ₁	T ₂	U ₁	U ₂	I ₁	I ₂	E ₁	E ₂	
	1		C ₄	C ₅	C ₆													
	2		T ₃	L	E _{m1}													
	3		T _m	A ₁	E _{m2}													
	4		U ₃	A ₂	E _{m3}													
	5		U _m	P _{v2}	P													
	6		I ₃	X _{m1}	X _{m2}													
	7		I _{m1}	I _{m2}	P _{v1}													
	8		C ₁	C ₂	C ₃													
	9		C ₄	C ₅	C ₆													
	10		T ₃	L	E _{m1}													
	11		T _m	A ₁	E _{m2}													
	12		U ₃	A ₃	E _{m4}													
	13		P	P _{t1}	P													
	14		I ₃	X ₁	X ₂													
	15		X ₃	X ₄	X ₅													

CHART 2

FIGURE 9.—UK-1 high speed encoder, telemetry channel allocations.

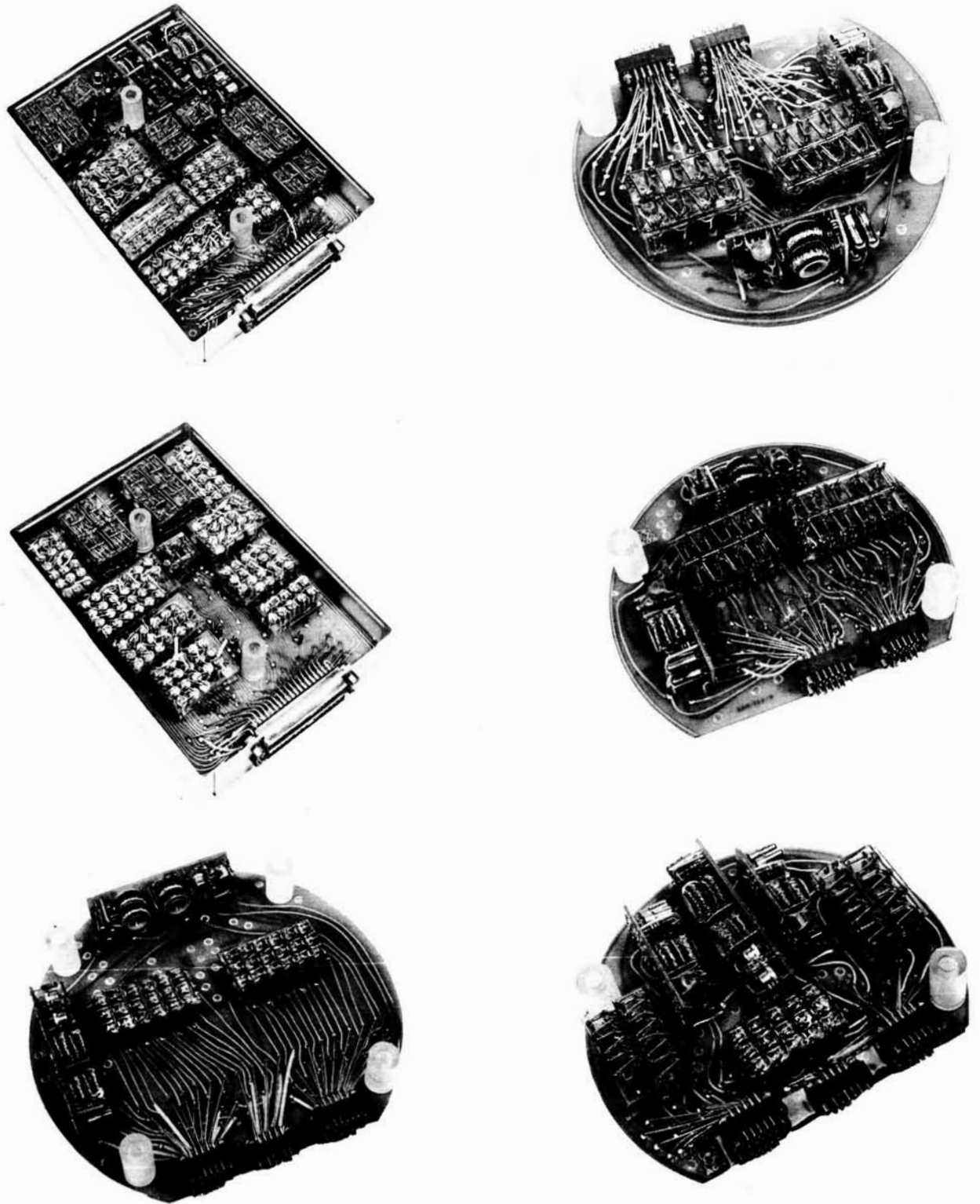


FIGURE 10.—UK-1 encoder cards (pre-prototype). (Not to scale.)

flight unit cards. The power converters for all the encoders were designed and fabricated by GSFC.

The finished product for the entire UK-1 encoding system, therefore, was a joint effort of Goddard Space Flight Center, Electro-Mechanical Research, Inc., and General Time Corporation.

THE "IMP" PFM ENCODING SYSTEM

As an example of how the basic PFM encoding system may be modified to accommodate different satellite requirements in bit rate vs. range, the Interplanetary Monitoring Probe (IMP) encoding system is described briefly.

The IMP system, now in the initial design state, will consist of two parts: a digital data processor (DDP), and an encoder.

Differences Between IMP and Basic PFM Systems

The IMP encoder is a modification of the basic PFM (Pulse Frequency Modulation) encoder system that was successfully flown on Explorer XII and on the UK-1 real-time HS encoder system. Table 1, and the list that follows, explains most of the modifications in the IMP system.

The six main modifications in the IMP system are:

1. The sampling rate has been reduced by a factor of 16 (i.e., the basic PFM sampling rate is

50 cps, but the IMP sampling rate is 50/16 cps).

2. The subcarrier-oscillator frequency deviation has been reduced by 16 (the basic PFM frequency deviation is 5 to 15 kc, but that of the IMP is 5/16 to 15/16 kc).

3. The time for a complete sequence has been increased by 16 (the basic PFM sequence time is 5.12 seconds, but in the IMP it is $5.12 \times 16 = 81.92$ seconds).

4. Telemetry sequences are essentially subcommutated (i.e., three normal sequences are followed by a fourth sequence, the Rb sequence). (See Figure 11.) Another way of stating this is that the basic PFM system consists of 256 channels, while the IMP encoder consists of 1024 channels as far as the ground station is concerned. This sequence commutation is not done in the encoder but is done in the programmer, with control pulses supplied by the encoder.

5. The blanks have been eliminated in some of the analog frames. In the basic PFM system, each channel consists of a blank followed by a burst; in the IMP, channels 1 through 15 of frames 3, 7, 11, 13, 14, and 15 contain no blanks but are 4.80-second bursts. In all other frames of the normal sequences (0, 1, 2, 4, 5, 6, 8, 9, 10, and 12), channels 1 through 15 consist of 0.16-second blanks followed by 0.16-second bursts. The fourth sequence (Rb sequence) will probably consist of an 81.92-second burst but with an 80

TABLE 1.—Comparison of IMP and Basic PFM Encoder Systems

Function	Basic PFM System	IMP Encoder "Normal" Sequence
Sampling rate	50 cps	50/16 cps
Subcarrier frequency deviation	5 to 15 kc	5/16 to 15/16 kc
Sequence time	5.12 seconds	*81.92 sec
Information channel construction	10-millisecond blank followed by 10-millisecond burst	0.16-sec blank followed by 0.16-sec burst, or continuous 4.8-sec burst
Data handling	By experimenter	By Digital Data Processor

*In the IMP, sequences are subcommutated (see Figure 11).

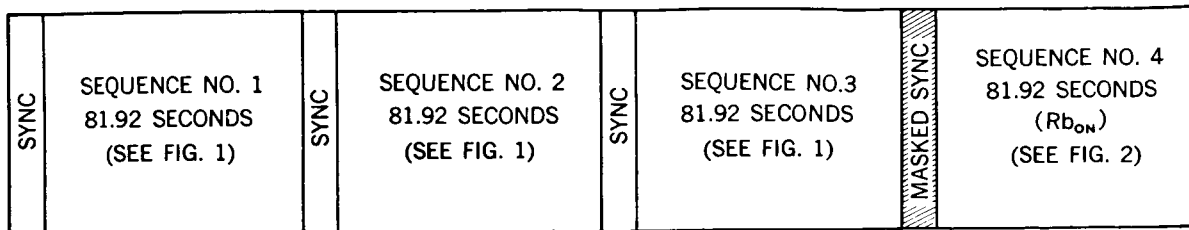


FIGURE 11.—IMP telemetry sequence format.

percent attenuation of the Rb signal at the beginning of each frame to allow a "masked sync" in the transmitter output.

6. Much of the digital data handling (accumulation and digital subcommutation) has been placed in the encoding system in a device called the DDP (Digital Data Processor). In the basic PFM system, each experimenter did his own accumulation and sometimes did his own subcommutation while, in the IMP, much of it is done by the DDP.

IMP Format

Figures 11 through 13 are timing diagrams of the IMP format. The following definitions will be helpful:

Telemetry Sequence: The time for 256 samples in a normal PFM system, the time to commutate through all the experiments.

Normal Sequence: The telemetry sequence shown in Figure 12 (all experiments except for Rb).

Rb Sequence: The telemetry sequence shown in Figure 13. Three normal sequences are followed by one Rb sequence, as shown in Figure 11.

Sequence Time: 81.92 seconds for IMP normal sequence.

IMP Encoder Design Philosophy

As previously indicated, the basic PFM format is a 256-channel encoder (divided into 16 channels by 16 frames) with a 50-cps sampling rate and a video output in the range of 5 to 16 kc with a fixed synchronization frequency of 4.5 kc. Each sample consists of a blank followed by a burst.

Many of the modules necessary for this type of format have been developed for the UK-1. The basic design philosophy for the IMP encoder has been to use as many of these modules as possible and to develop new ones only when absolutely necessary. To slow the encoder down by 16 is no

particular problem, since the clock frequency is merely changed to 50/16 cps and the oscillators are then "mixed" (only one oscillator is *on* at a time) into a composite video in the 5 to 15 kc range. This composite video is then divided by 16 to give the desired 5/16 to 15/16 kc range. A slowdown by 48 was accomplished in the UK-1 LS encoder by this technique.

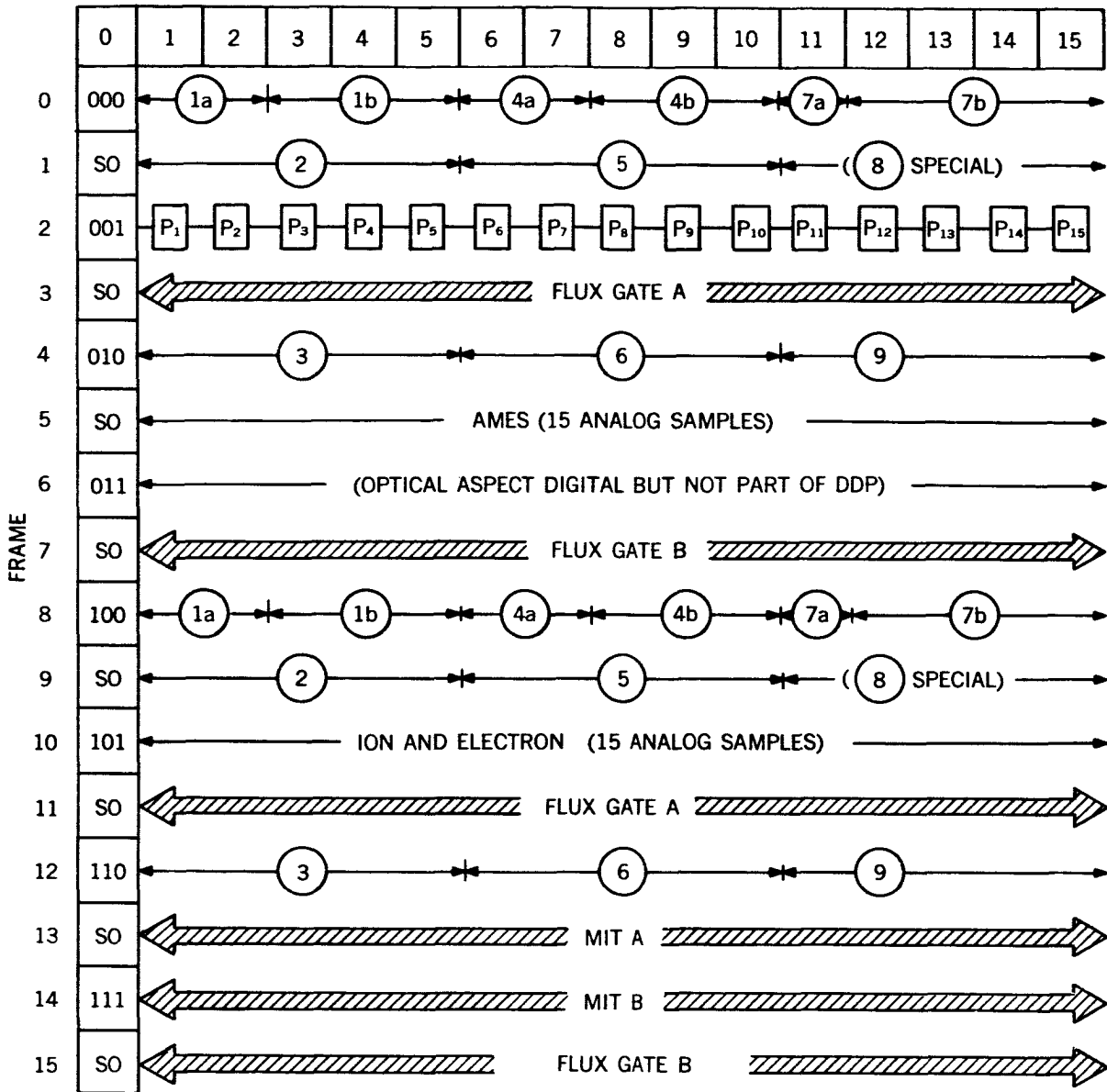
The UK-1 low speed information was then recorded on a tape recorder *in the satellite*. Upon command, the tape was played back 48 times faster than it was recorded so that the ground station received information in the 5 to 15 kc band while the encoder output, as mentioned, was 5/48 to 15/48 kc. As of this writing, the UK-1 satellite has successfully completed environmental tests, and the encoding scheme works well.*

The divide-by-16 circuit at the output of the mixer is a mixed "blessing" in that it decreases reliability but does provide a chance to make the output video coherent when using simple, proven, noncoherent oscillators.

Figure 11 shows that the telemetry sequences are actually subcommutated; in addition, during the Rb sequence it is necessary to carry the sync-frequency channel (0-all) through while the Rb signal is attenuated by 80 percent. Basically, there are two ways to handle this problem. One is to generate the entire sequence format in the encoder; this would result in a 1024-channel encoder instead of the "standard" 256-channel encoder. The resulting complexity of the encoder would be rather large because extra switching functions would have to be added to each encoder matrix function. In addition, the DDP would have to be redesigned and the existing 256-channel decoding test equipment would have to be modified.

*Since this report was written, the spacecraft was successfully launched from Cape Canaveral on April 26, 1962 on a three-stage Delta rocket.

CHANNEL



= ACCUMULATOR NO. n READOUT OF DDP (MOD A)
 = ANALOG READOUT; CONTINUOUS 4.80 SECOND BURST

P_n=PERFORMANCE PARAMETER NO. n (0.16 SECOND READOUT)
 SO=SYNC OSCILLATOR=4.5/16 KC
 COMPLETE TELEMETRY SEQUENCE=5.12 x 16=81.92 SECONDS
 DIGITAL BURSTS (3 BITS) TAKE 0.16 SECOND FOLLOWING A 0.16 SECOND BLANK
 SYNC CHANNEL BURSTS TAKE 0.24 SECOND FOLLOWING A 0.08 SECOND BLANK
 "SAMPLING RATE"=50/16 CPS

CHART 4

FIGURE 12.—IMP telemetry channel allocations for sequences 1, 2, and 3.

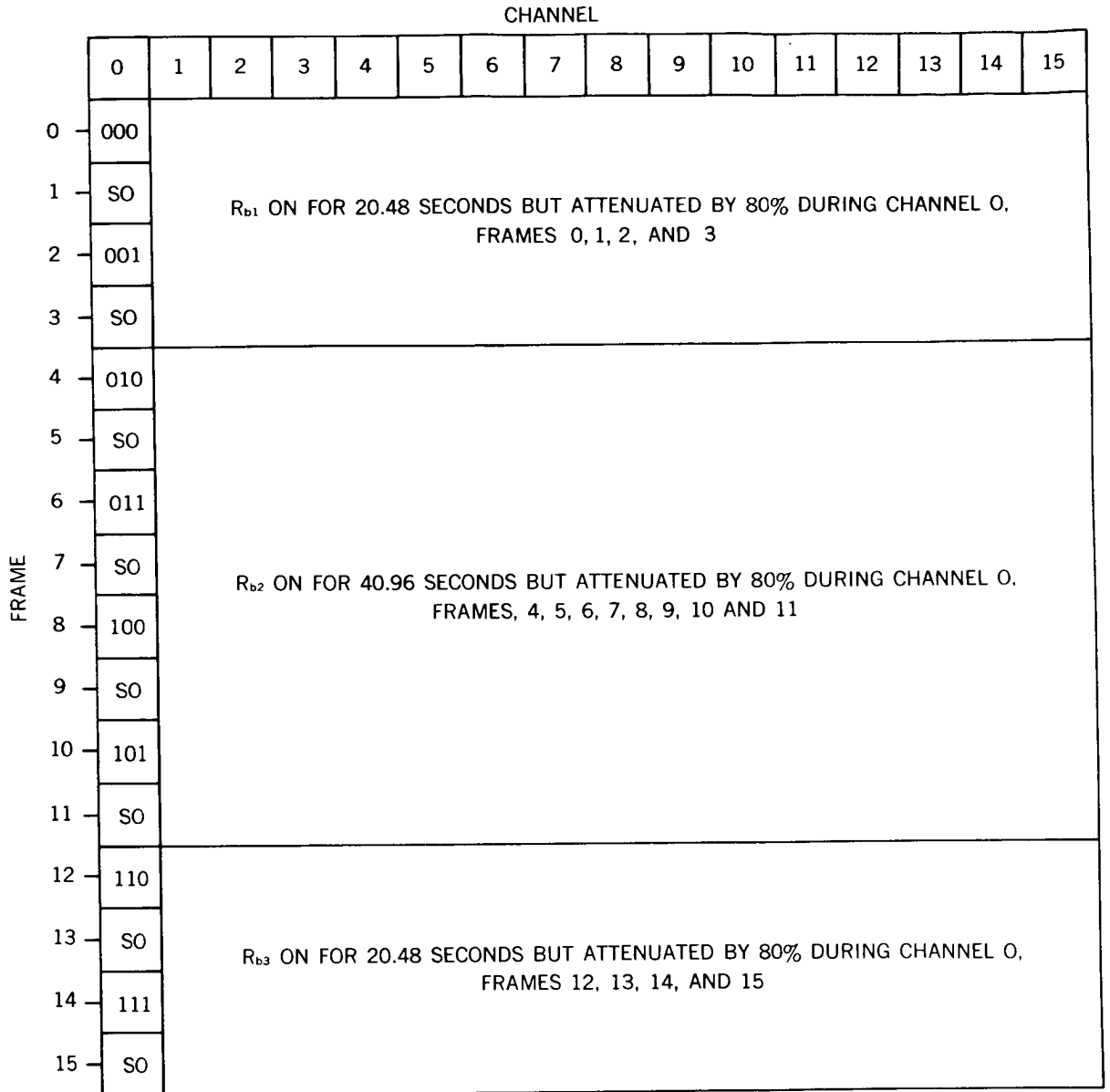


CHART 5

FIGURE 13.—IMP telemetry channel allocation for sequence 4 only.

The other approach to the problem is indicated in Figure 14, where a video-out signal is generated as if the encoder were a normal 256-channel device. In other words, the video-out signal is present even during the R_b sequence. An additional countdown by 4 is added to the 256-state countdown unit, and the additional binary counters (A1 and A2) are utilized to generate synchronizing pulses that indicate the occurrence of

sequence 4 and to initiate control circuitry for the R_b experiment. The switching is then done in the IMP programmer to result in the sequence shown in Figure 11.

Digital Data Processor

The Digital Data Processor (DDP) has been designed to accept inputs either from the digital type of experiment (e.g., counting experiments) or

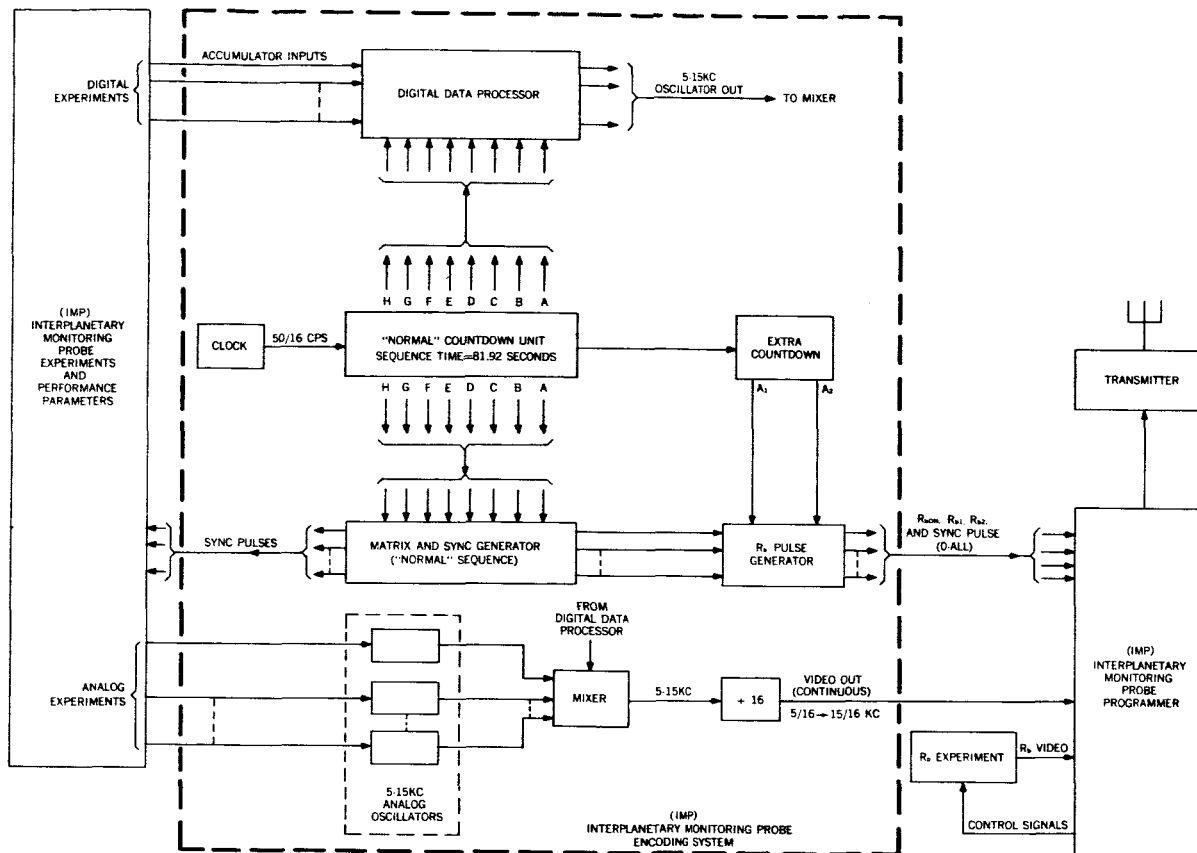


FIGURE 14.—IMP encoding system.

from the type of experiment that measures the magnitude of random discrete phenomena, such as randomly occurring pulse heights, and requires the storage of this data until it is telemetered. In the latter type of experiment, the experimenter will be required to supply a pulse train where the number of pulses supplied is proportioned to the value of the parameter being measured. In the counting experiment, the experimenter must supply a line that contains a pulse whenever his probe is activated by the event he wishes to count. Some signal conditioning is done inside the DDP. Each accumulator will accept pulses up to a 500-ke rate. Some of the accumulators are reset after each read-out while others are not reset. Every accumulator is frozen during read-out.

Another function of the DDP is to supply a digital clock with a storage capacity approximately 32,000 times the telemetry sequence time. This output will be telemetered at least once per telemetry sequence.

Essentially the DDP will consist of 12 accumulators, read out twice per telemetry sequence (see Figure 12). The accumulator capacities in the DDP are: three of 6 bits, three of 9 bits, and six of 15 bits. One of the 15-bit accumulators will be used as the digital clock. The total capacity of the DDP, then, is 135 bits.

The logical design is such that the entire DDP may consist of three identical cards (see Figure 15). The three cards are called sets 1, 2, and 3, respectively. Each of the three cards contains: four accumulators (45 bits); a PFM digital oscillator; and a generalized matrix that supplies the scanning, freeze, and reset functions necessary for accumulator operation. The accumulators are packaged 3 bits at a time with their associated output gate circuits. Thus, each of the cards is a self-contained unit; and connecting different countdown functions to the plug (external to the card) will generate different scanning, freeze, and reset functions via the generalized matrix. Three

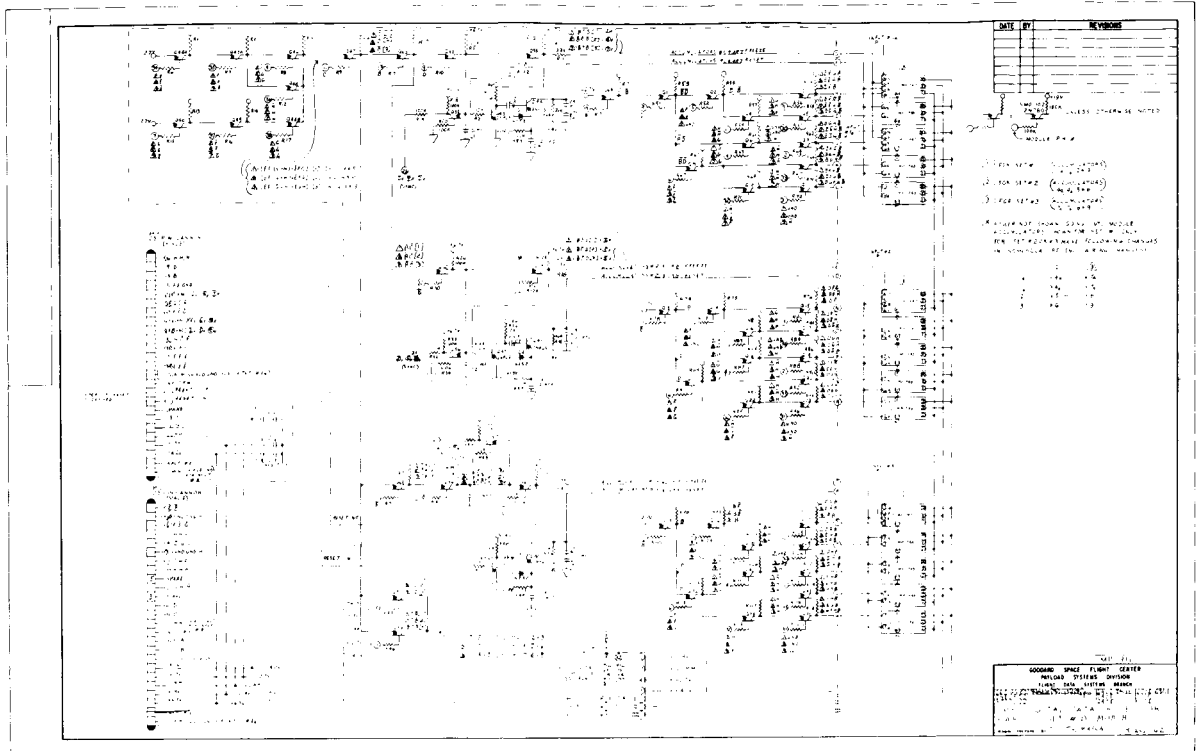


FIGURE 15.—IMP digital data processor, mod B.

sets of Boolean equations were written, simplified, and arranged so that the generalized matrix would solve the three sets of equations with the same hardware.

CONCLUDING REMARKS

PFM telemetry lends itself to relatively simple encoding techniques. The development of these techniques has been directed toward making PFM encoders more easily used by satellite experimenters. It has been demonstrated that the

basic PFM format can be readily changed to adapt PFM telemetry to varying space applications.

REFERENCES

1. ROCHELLE, R. W., "Pulse Frequency Modulation," NASA Technical Note D-1421, 1962. Also published in: *IRE Trans. on Space Electronics and Telemetry* SET-8(2):107, June 1962.
2. PAULL, S., "A Magnetic Core Voltage-to-Frequency Converter," NASA Technical Note D-1677, 1962.

IV. SENSORS

THE TIROS LOW RESOLUTION RADIOMETER*

FRANK BARTKO, VIRGIL KUNDE, CLARENCE CATOE, AND MUSA HALEV

The Tiros II, III, and IV (1960 π 1, 1961 ρ 1 and 1962 β 1) meteorological satellite experiments included a low resolution infrared radiometer of broad spectral and spatial response designed to measure the planetary heat budget of the earth. The radiometer consisted of two thermistor detectors, each mounted at the base of a reflective cone. One detector was coated black and was sensitive to reflected solar and long wave terrestrial radiation. The second detector was coated white and was sensitive primarily to the long wave terrestrial radiation. Simultaneous evaluation of the energy balance equation for each detector yields the earth's apparent blackbody temperature and albedo.

The first portion of this paper presents the physical and optical characteristics of the radiometer, the calibration, and the procedure for determining the coefficients of the energy balance equations from the calibration data. The second portion of the paper discusses the procedure for reducing the data and presents an error analysis of the radiation data from a Tiros III orbit. Efforts to improve the performance of the low resolution radiometer and a summary of the status of the low resolution radiometer program are also presented.

LIST OF SYMBOLS

A	Albedo of the earth, defined as the reflectance of the area on the earth viewed by the radiometer, averaged spatially over the solid angle defined by the field of view, and averaged spectrally over the solar spectrum	k_2'	Thermal coefficient of proportionality, equal to $K/\sigma\epsilon_t$
B	Elevation angle of the sun	k_3	Reflected solar radiation coefficient ($^{\circ}\text{K}^4\text{-cm}^2/\text{watt}$), equal to $K_3/\sigma\epsilon_t$
B'	Angle between the incoming light beam and the plane of diffuse paper	P	Pressure
C	Thermopile constant for diffuse radiation	$R_c(\lambda)$	Spectral reflectance of a reflecting cone
C_1	Thermopile constant for beam radiation	R_h	Housing thermistor network resistance
f	Frequency of the oscillator	R_N	Network resistance
H_s	Solar radiation incident upon the white paper of the calibration apparatus after being modified by the earth's atmosphere and a reflecting mirror	R_{N_b}	Network resistance for the black detector
K	Thermal conductive coefficient, watts/ $^{\circ}\text{K-cm}^2$	R_{N_w}	Network resistance for the white detector
K_1	Thermal coefficient of proportionality	R_o	Resistance of constant value
K_2	Thermal coefficient of proportionality	R_{ob}	Resistance of constant value for the black detector network
K_3	Reflected solar radiation coefficient	R_{oh}	Resistance of constant value for the housing thermistor network
k_1	Thermal coefficient of proportionality, equal to K_1/ϵ_t	R_{ow}	Resistance of constant value for the white detector network
k_2	Thermal coefficient of proportionality, equal to K_2/ϵ_t	$R_p(\lambda)$	Spectral reflectance of a paint coating on the thermistor detector
		R_t	Thermistor detector resistance
		R_{tb}	Black detector resistance
		R_{tw}	White detector resistance
		R_1	Reference resistance, near the upper limit of the bandwidth
		R_3	Reference resistance, near the lower limit of the bandwidth
		r_p	Spectral reflectance of diffuse paper

*Published as NASA Technical Note D-614, September 1964

S	Solar constant
S_i	Sensistor resistance
S_j	Sensistor resistance
S_1	Sensistor resistance for the detector leg of the black detector network
S_2	Sensistor resistance for the fixed resistance leg of the black detector network
S_3	Sensistor resistance for the fixed resistance leg of the white detector network
S_4	Sensistor resistance for the detector leg of the white detector network
T_b	Black thermistor detector equilibrium temperature
T_c	Housing temperature of the medium resolution radiometer
T_E	Oscillator temperature
T_e	Apparent earth blackbody temperature, corresponding to the thermal emission received from the earth's surface and intervening atmosphere, averaged spatially over the solid angle defined by the field of view and averaged spectrally over the thermal spectrum
T_q	Quartz window temperature
T_s	Radiometer housing temperature
T_t	Thermistor detector equilibrium temperature
T_w	White thermistor detector equilibrium temperature
t	Transmittance of the quartz window
V_{bb}	Thermocouple voltage representing the blackbody cavity temperature
V_1	Thermopile response due to H_s
V_2	Thermopile response due to W_{sp}
V_3	Thermopile response due to W_{bp}
W	Total radiant emittance from the earth
W_{bp}	Radiant emittance of the diffuse paper due to tungsten bulb irradiance
W_{sp}	Radiant emittance of the diffuse paper due to solar irradiance
$\alpha_t(\lambda)$	Spectral absorptance of the painted thermistor detector
ϵ_t	Spectral emissivity of the detector
σ	Stefan-Boltzmann constant
$\phi(\lambda)$	Spectral response of the radiometer
$\phi'(\lambda)$	Auxiliary spectral response function, equal to $r_p\phi(\lambda)$

INTRODUCTION

The Tiros II, III, and IV (1960 $\pi 1$, 1961 $\rho 1$, and 1962 $\beta 1$) meteorological satellites each contain two television cameras, a medium resolution scanning radiometer, a low resolution non-scanning radiometer, and a University of Wisconsin heat balance experiment. The resolution description (medium or low) refers to both the spectral and spatial optical characteristics of the radiometers. The television cameras and medium resolution radiometer have been described elsewhere (References 1-4). This paper is concerned only with the low resolution non-scanning radiometer. The mean orbital characteristics for the Tiros satellites with an operational low resolution radiometer on-board are listed in Table 1.

Table 1

Tiros Mean Orbital Characteristics.

Satellite	Launch Date	Perigee Altitude (km)	Apogee Altitude (km)	Anomalistic Period (min)	Orbital Inclination (degrees)
Tiros II	11/23/60	620.6	731.3	98.27	48.53
Tiros III	7/12/61	740.5	815.6	100.41	47.90
Tiros IV	2/8/62	710.3	844.7	100.40	48.30

The low resolution radiometer was designed to measure thermal and reflected solar radiation emanating from the earth. From these two quantities, the apparent blackbody temperature and the albedo of the earth can be determined. A study of these data over an extended period of time shows, in principle, how the incoming solar energy is utilized by the earth. Such quantitative measurements could hopefully provide a firm foundation for a better study of the earth's overall energy budget. The basic principles of the experiment have been described by Hanel (Reference 5). The radiometer characteristics, the calibration procedures, and the reduction of calibration and orbital radiation data are presented here. The limitations of the radiometer and efforts to remove these limitations are also discussed.

The physical basis of the experiment is provided by the consideration of the steady state

energy gains and losses experienced by a thermistor detector exposed to the earth from a satellite about 765 km above the surface. The low resolution radiometer consists of two such thermistor

detectors housed in the Tiros satellite structure. In the vacuum environment, the steady state thermal balance of each detector is given by Hanel (Reference 5) as

$$\epsilon_t \sigma T_t^4 = K_1 \sigma T_e^4 + K_2 \sigma T_s^4 + K(T_s - T_t) + K_3 SA \sin B \quad (1)$$

where the left side is the radiative energy loss of the thermistor detector at an equilibrium temperature T_t and emissivity ϵ_t . The first and second terms on the right represent the radiative coupling of the detector with the earth at an apparent blackbody temperature T_e and the radiometer housing at a temperature T_s , respectively. The third term on the right represents the conductive coupling with the satellite structure, and the fourth term is the earth's reflected solar radiation, where S is the solar constant, A is

venient form

$$T_t^4 = k_1 T_e^4 + k_2 T_s^4 + k_2' (T_s - T_t) + k_3 SA \sin B \quad (2)$$

In Equation 2 the albedo is the reflectance of the area on the earth viewed by the radiometer, averaged spatially over the solid angle defined by the field of view and averaged spectrally over the solar spectrum. The apparent earth temperature is the blackbody temperature corresponding to the thermal emission received from the earth's surface and intervening atmosphere, averaged spatially in the same manner as for the albedo and averaged spectrally over the thermal spectrum. When the boundary conditions $S=0$ and $T_t = T_e = T_s$ are applied, then $k_1 + k_2 = 1.0$.

If the T_t 's and T_s 's are measured and the k 's are determined by calibration, the solution of two such equations for two detectors of sufficiently different optical characteristics will yield the desired values of T_e and A . The instrument considered here was designed to provide for each detector values of approximately 0.9 for k_1 and about 0.1 for k_2 .

The low resolution radiometer is mounted on the baseplate of the Tiros satellite so that it views the earth along the direction of the satellite spin axis. The location of the radiometer can be seen in Figure 1. The spin-stabilized satellite is space oriented so that the radiometer has its field of view filled by the earth for a maximum time of 1/5

the albedo, as defined below, and B is the solar elevation angle at the center of the viewed area. It is necessary to include the coupling of the detectors with the satellite since it is impossible to thermally isolate the detectors completely from the satellite structure. The thermal coefficients of proportionality, K_1 , K_2 , and K , and the proportionality coefficient for reflected solar radiation, K_3 , include the geometrical and optical characteristics of the detectors.

Dividing Equation 1 by $\epsilon_t \sigma$ yields the con-

of the orbital period. The geometry of the viewing area at different parts of the orbit is shown

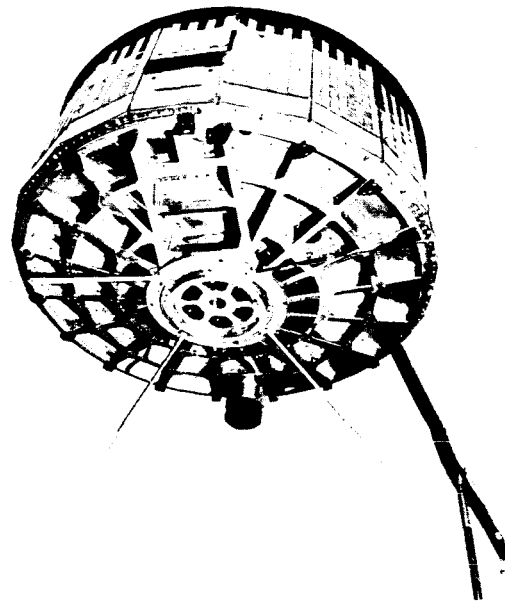


FIGURE 1.—Tiros meteorological satellite. A wide angle television lens protrudes from the baseplate and the round aperture near the top of the photograph is the location of the low resolution radiometer.

in Figure 2. The orientation and shielding provided by the satellite are such that the detectors should never receive direct solar radiation.

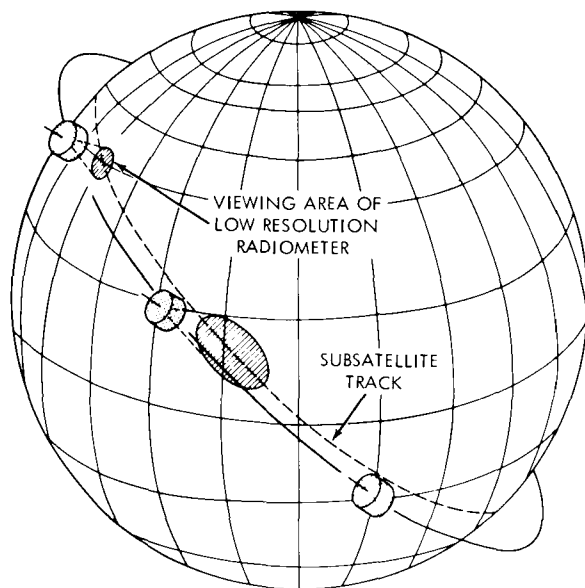


FIGURE 2.—Examples of the viewing area of the low resolution radiometer.

DESCRIPTION OF THE RADIOMETER

The radiometer structure is illustrated in Figure 3 which shows a front view of the instru-

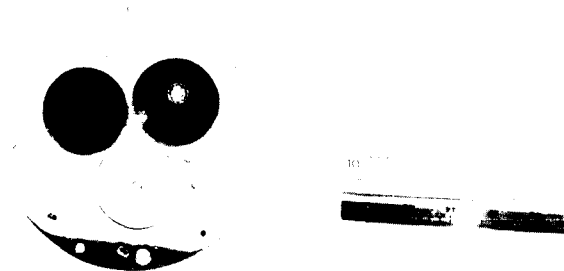


FIGURE 3.—The low resolution radiometer. The white thermistor detector can be seen at the base of the aluminumized cone on the right. The black thermistor detector is located in the cavity on the left.

ment (Reference 4). Two highly aluminumized truncated Mylar cones are mounted on a 3 in. diameter, gold-plated, aluminum plate. At the base of each cone is mounted a thermistor detector fastened to the mylar base by a grid of fibers of

low thermal conductivity. The fiber strength has proved to be more than sufficient as evidenced by numerous successful vibration tests on the radiometer. The two thermistors comprise the heart of the instrument. The resistance of these detectors is a function of the energy flux which they absorb. The response is the value of the resistance of the detector, which is further processed by additional electronics and transmitted to the earth. Further details will be described below. Each of the two platinum electrical leads attached to each detector winds about one of the mounting grid fibers and passes through the base of the cone wall to a network of three additional resistances. Since the response is the value of the detector resistance, three additional resistances are provided to compensate for effects due to housing temperature changes on the detector. The network is illustrated by Figure 4. R_N is the resul-

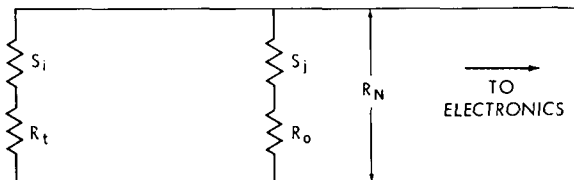


FIGURE 4.—Schematic of the resistance network for the thermistor detector of the low resolution radiometer. R_N is the resultant resistance of the four-element network consisting of resistances of the thermistor detector R_t , two sensitistors S_i and S_j , and a fixed resistance R_o .

tant resistance of the four element network consisting of resistances of the thermistor detector (R_t), two sensitistors (S_i , S_j), and a fixed resistance (R_o).

The resistance values and their temperature coefficients are chosen to give the best compensation for housing temperature effects on the detector. The wire-wound resistance R_o is given a fixed value which is independent of temperature. The response of each detector is actually found by calibrating R_N as a function of housing temperature and target temperature and then calculating R_t from R_N . The following relationship holds between the various elements of the resistance network:

$$R_t = \frac{R_N (R_o + S_j) - S_i (R_o + S_j - R_N)}{R_o + S_j - R_N} \quad (3)$$

The radiometer housing temperature is determined from the resistance R_h of a network consisting of a thermistor and a fixed resistor connected in parallel. A back cover plate is provided to protect the delicate construction of the cones and the electrical components. The surfaces of the cover plate are highly polished to minimize radiation coupling with the satellite. The total weight of the instrument is 46.3 gm. A schematic of the radiometer is shown in Figure 5.

The most important optical characteristics of

the radiometer are its spectral sensitivity, field of view, and response time. It was attempted to tailor each of these quantities to the desired objectives of the experiment and to make them compatible with the Tiros satellite.

The spectral characteristics of the radiometer were determined by coatings of paint applied to each detector. These characteristics were based on the following considerations. The total radiant emittance, W , emanating from the earth consists of two spectral components, expressed

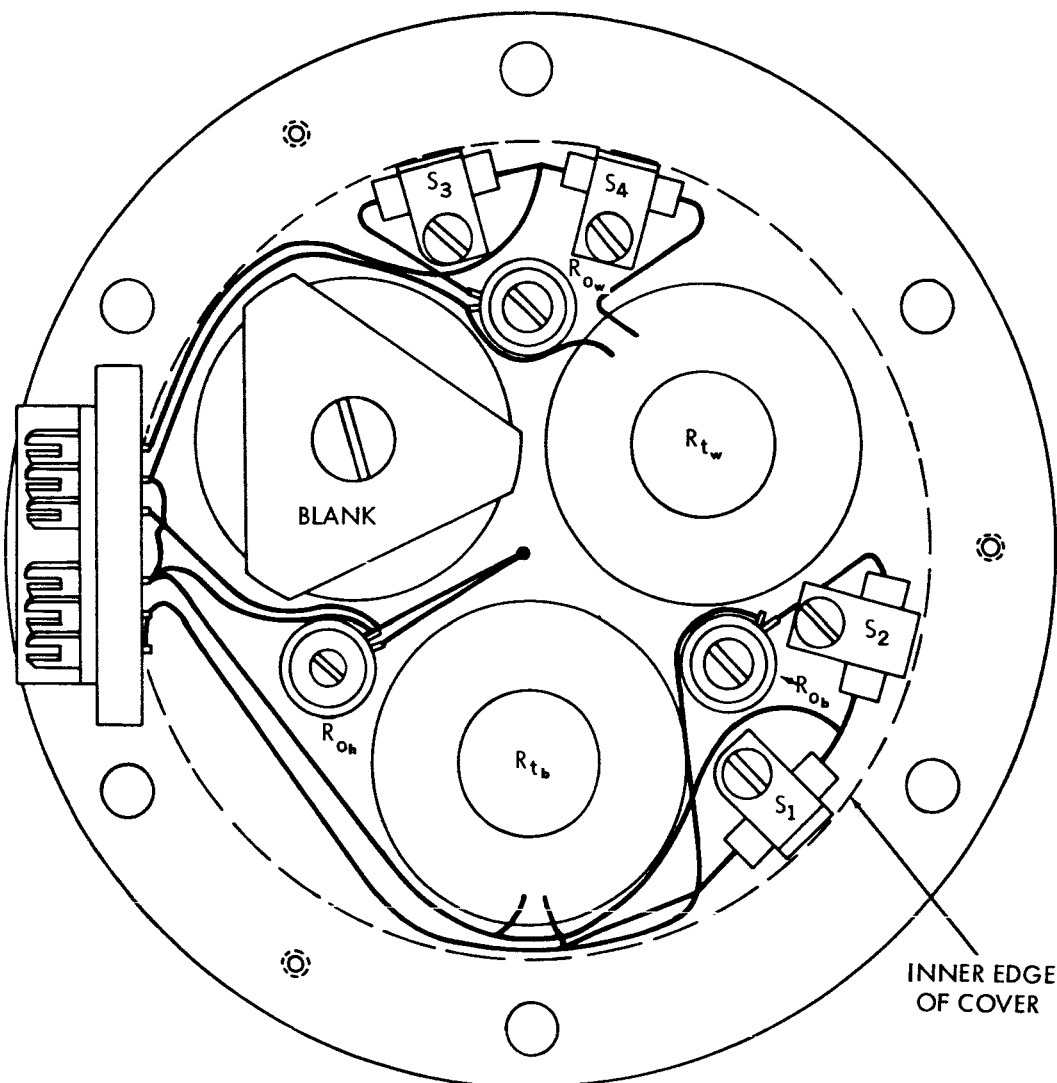


FIGURE 5.—Schematic of the low resolution radiometer. R_{tb} , S_1 , S_2 and R_{ob} comprise the network resistances for the black cone and R_{tw} , S_3 , S_4 , and R_{ow} comprise the network resistances for the white cone. The housing temperature T_s is determined from the resistance R_h of a separate thermistor and an associated fixed resistance R_{oh} .

quantitatively by

$$W = W_{\text{earth emission } (\lambda > 4\mu)} + W_{\text{reflected solar radiation } (\lambda < 4\mu)}$$

The values of these components can be obtained by measuring each separately or measuring the total and one of the components. In the present case the total energy and the terrestrial emission were selected for measurement. One thermistor was coated with a black paint designed to absorb, as uniformly and completely as possible, radiation over all wavelengths of the spectrum. The second thermistor was painted white in order to absorb primarily the terrestrial emission and be spectrally blind to the reflected solar radiation from the earth. The differences in temperatures of these thermistors then provide a means for calculating the values of T_e and A .

The field-of-view characteristics of each detector are determined by the angular aperture of the cone and the ratio of cone length to detector radius. The half-angle of the cones is about 27 degrees and was chosen to: (1) optimize the radiation exchange of the detectors with the earth's surface, (2) provide integrated values for quantities that are averaged over large areas, and (3) view an area on the earth comparable in size with that viewed by the television cameras. At the altitude of the Tiros satellites (≈ 765 km), when the satellite spin axis is coincident with the local vertical of the earth, circular areas of approximately 870 and 1050 km in diameter are viewed by the radiometer and wide-angle television camera, respectively.

The time response of each detector is directly proportional to its thickness. This includes the thickness of the actual detector and that of the paint. The value chosen for the time constant had to be compatible with the radiometer field of view and the orbital velocity of Tiros. The design figure for the radiometer was about 30 sec and measurements indicate a value of about 35 sec.

CALIBRATION MEASUREMENTS

The values of the thermal and reflected solar radiation coefficients of the heat transfer equation must be known as accurately as possible. Since these coefficients involve the geometric and optical parameters of the radiometer in a complex fashion,

an empirical method is the best means for obtaining their values.

While the radiometer is viewing the earth, the energy flux incident upon the radiometer consists of the thermal emission from the earth and its atmosphere (4 to 40μ) and the reflected solar radiation (0.2 to 4μ). It is quite difficult to simulate these components in a realistic fashion with a single standard calibration source. As a result, the calibration was separated into two parts, each of which could be handled by separate standard calibration sources. In each part, the calibration consisted of placing the radiometer in a small vacuum chamber and exposing each detector to a target of known and variable radiant emittance. A vacuum environment is necessary since the heat balance of the detectors had to be controlled by radiative processes during calibration just as it is in the vacuum environment in orbit. If air were present about the detector, the conduction of heat to and from the detector by the air would overshadow the heat transfer by radiation. Vacuum chamber pressures of 10^{-5} mm Hg were used for the calibration of the instrument. The housing temperature of the radiometer became a parameter and was varied through the expected temperature range for each value of target radiant emittance for the thermal radiation calibration. The reflected solar radiation calibration was performed at one single housing temperature only, since corrections for the entire range of housing temperatures had already been determined in the thermal calibration.

Thermal Calibration

The first portion of the calibration was designed to simulate the expected thermal emission from the earth. The radiometer was exposed to a blackbody cavity of known temperature. The whole apparatus was placed in a vacuum chamber as illustrated in Figures 6 and 7. The radiometer was mounted in good thermal contact to a metal plate with cutouts in the center which permit each detector to view the cavity directly. The temperature of the metal plate and radiometer could



FIGURE 6.—The laboratory calibration apparatus. A low resolution radiometer is positioned in front of the black-body cavity. The upper part of the photograph shows a second radiometer in position for the reflected solar radiation calibration.



FIGURE 7.—The calibration stand and vacuum chamber. The radiometer mounted for the reflected solar radiation calibration can be seen through the quartz window of the vacuum chamber.

be varied by circulating liquids of various temperatures through a tubing soldered onto the plate. The housing temperature range was varied from -20°C to $+60^{\circ}\text{C}$, by using alcohol cooled with various amounts of dry ice for the cold temperatures and heated by an electrical heater for the high temperatures. The blackbody temperature, -100°C to $+60^{\circ}\text{C}$, was likewise controlled by circulating liquids; it was measured by a sensitive voltmeter with the use of a calibrated thermocouple placed inside the cavity. Liquid nitrogen for the lowest temperatures, alcohol with varying amounts of dry ice, and electrical heaters were used to vary the cavity temperature.

The procedure, then, consisted of holding the temperature of the blackbody cavity constant and varying the temperature of the housing. The measurements were then repeated for each detec-

tor with a different value of the blackbody temperature, until the complete range was covered. A set of calibration curves was then developed showing the response for each detector as a function of the housing temperature with the cavity temperature as a parameter. Curves are shown in Figure 8 for a typical white detector.

Figure 9 illustrates the procedure for the thermal radiation calibration measurements. The calibration was designed to simulate the actual electrical environment existing in Tiros. Each detector can be connected through a switching box to a subcarrier oscillator. The network resistance of the detector then controls the output frequency of the oscillator within the frequency band 425–475 cps. The frequency output of the oscillator passes to a frequency counter and is recorded by a digital printer. The response of the

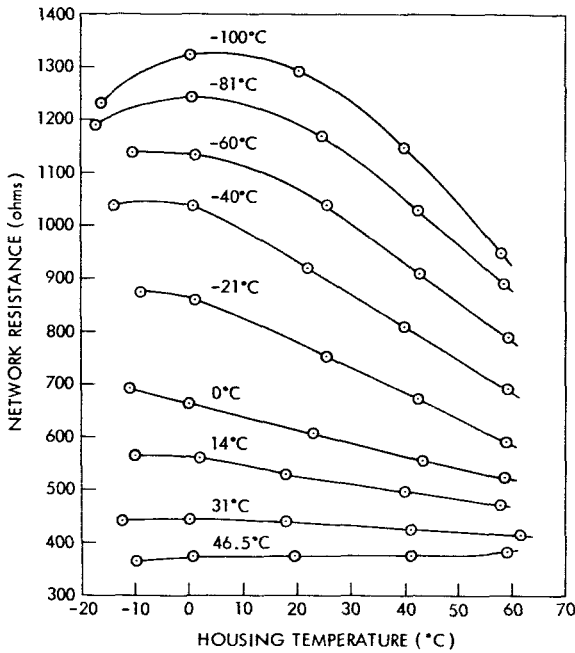


FIGURE 8.—Thermal test data for a white detector. The parameter for the curves is the apparent earth blackbody temperature.

frequency-controlled subcarrier oscillator was calibrated at the same time. Therefore, immediately after the measurement, a decade box was connected to the input of the oscillator and the resistance R_N which corresponds to the measured frequency was found.

Figure 10 illustrates the calibration curve for the housing thermistor network resistance. This calibration can be made in terms of R_h , rather than of the thermistor resistance itself, because there are no temperature effects on the fixed resistor.

The unreduced thermal calibration data consisted of R_h , V_{bb} , R_{Nb} , and R_{Nw} , where the subscripts b and w represent the black and white detectors. These must be converted to their respective detector equilibrium temperatures. The first two conversions are accomplished readily, but the last two require more extensive reduction. The value of R_h is taken directly from a thermistor housing temperature curve as is shown in Figure 10. The value V_{bb} is converted by a calibration chart for the thermocouple. The

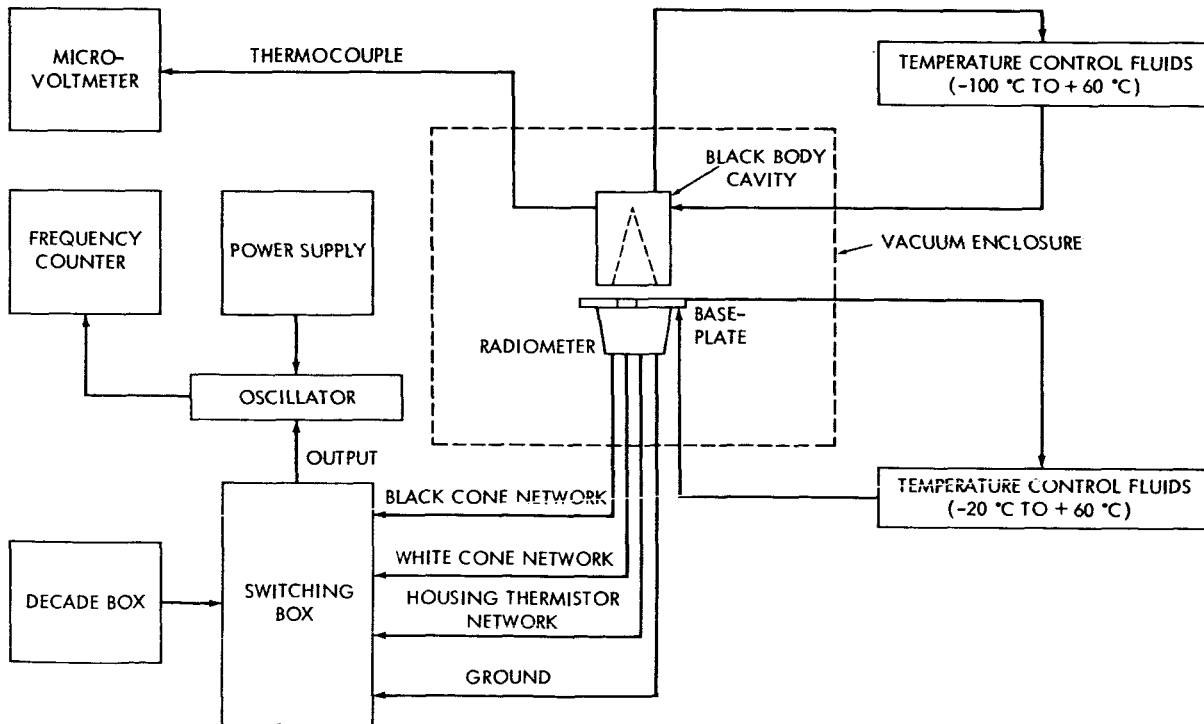


FIGURE 9.—Schematic diagram of the thermal radiation calibration setup.

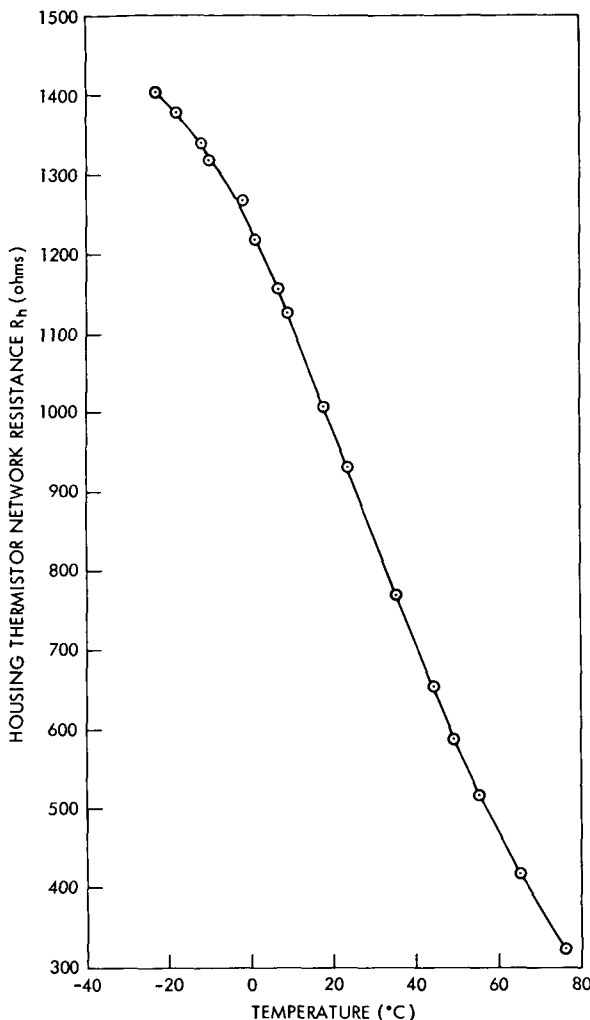


FIGURE 10.—Radiometer housing thermistor network response as a function of temperature.

reduction of the detector network resistances is discussed later.

Reflected Solar Radiation Calibration

The second portion of the calibration involved an attempt to simulate the reflected solar radiation from the earth. The calibration consisted of exposing each detector to a diffuse light source of known and variable radiant emittance. As before, the radiometer was placed in the vacuum chamber, but in a position to view a large sheet of white diffuse paper through a quartz window. The illuminated paper was placed in a position such that it filled the fields of view of both detectors. The source of radiation used to illuminate

the white diffuse paper consisted of an array of eight blue-dyed tungsten bulbs. The tungsten bulbs were used for two reasons: (1) A greater variation in radiant emittance could be attained by the use of an array of bulbs as contrasted with a single value obtained by using the sun directly at any given time, and (2) for practical calibration purposes, an artificial source was necessary to avoid delay caused by unpredictable weather conditions. A great disadvantage occurs however since the spectral energy distribution is different from that of the sun. These bulbs were used to simulate solar radiation under the assumption of a flat spectral response for the radiometer over the spectral range of the solar radiation and the bulbs, so that the radiometer response depended only on the total energy incident on it. (The validity of this assumption is discussed later when the spectral response measurements are presented.) Hence, the network resistances of the black and white detectors were calibrated against the total energy output of the bulbs, expressed in terms of the solar constant by using an Eppley thermopile as the standard of measurement. First, however, the thermopile had to be calibrated in terms of the existing experimental conditions. This was necessary because of differences in the types of sources viewed in the original calibration of the thermopile and in the radiometer calibration. The original thermopile calibration refers to beam radiation whereas the radiometer calibration uses diffuse radiation. Because of these differences, the original calibration data for the thermopile, supplied by the manufacturer, could not be used directly to convert its output voltage to total energy received and it was necessary to calibrate the thermopile for radiation from the same diffuse target as seen by the radiometer. The experimental arrangement is illustrated schematically in Figure 11. A 30 in. diameter aluminized, optically flat mirror directed a beam of sunlight onto a sheet of diffuse white paper. By assuming that the paper obeys Lambert's Law, measurement of the angle (B') between the incoming beam of solar radiation and the plane of the paper, combined with a knowledge of the solar irradiance as modified by the mirror reflectance H_s and the spectral reflectance of the diffuse paper r_p , permitted the radiant emittance of the paper W_{sp} to be calculated.

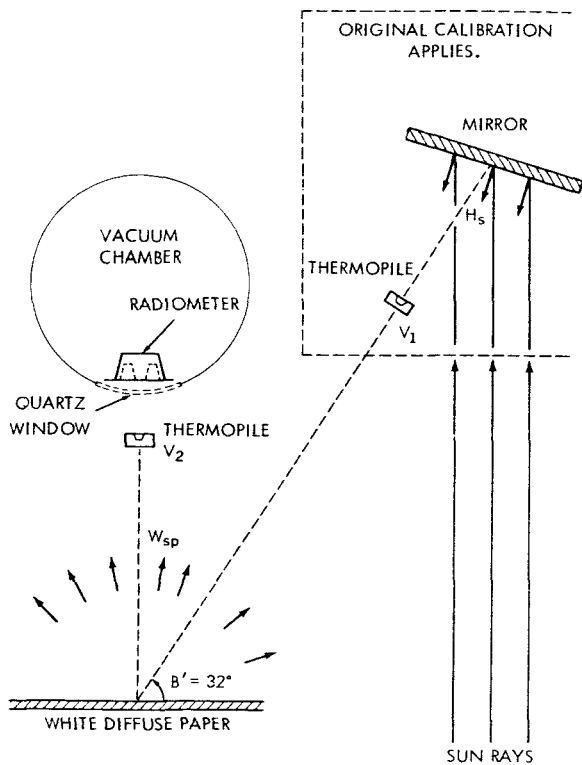


FIGURE 11.—Two-dimensional schematic of the calibration of the thermopile for diffuse radiation in terms of reflected solar radiation.

The thermopile was first placed to view the incoming solar beam. This gave a measure of H_s at the time of the measurement. The thermopile response V_1 was converted to H_s by

$$V_1 = C_1 H_s, \quad (4)$$

where C_1 is known from the original calibration data. The thermopile was next placed to view the illuminated paper in the same fashion as the radiometer, and its response V_2 was recorded:

$$V_2 = C W_{sp}, \quad (5)$$

where C is the constant to be determined. Since the paper was assumed to obey Lambert's Law,

$$W_{sp} = H_s \sin B' \bar{r}_p, \quad (6)$$

where \bar{r}_p is the average reflectance (over the solar spectrum) of the white diffuse paper. Substituting Equation 6 into Equation 5 gives

$$V_2 = C(H_s \sin B' \bar{r}_p). \quad (7)$$

Hence, by knowing V_2 , B' , and \bar{r}_p , and measuring H_s with Equation 4, the value of C can be found.

With C known, Equation 5 defines the calibration of the thermopile for diffuse radiation. Since the thermopile response is assumed to be spectrally flat, Equation 5 is assumed to be linear, as shown in Figure 12.

The array of eight tungsten bulbs was then placed to illuminate the white diffuse paper as shown in Figure 13. The thermopile was again placed to view the diffuse source in the same fashion as the radiometer, and its response was

$$V_3 = C W_{bp}, \quad (8)$$

where W_{bp} is the radiant emittance of the diffuse paper due to the irradiance from the tungsten

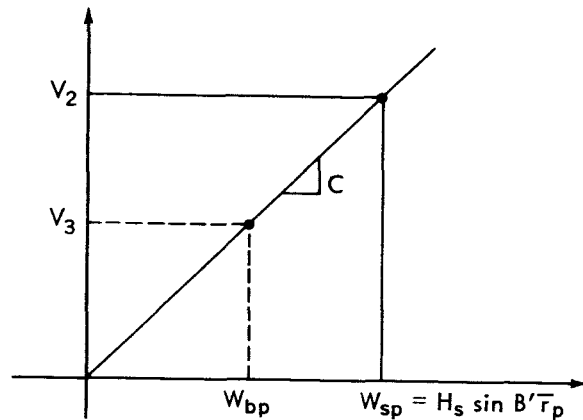


FIGURE 12.—Graphical relationship between the thermopile response and diffuse radiation.

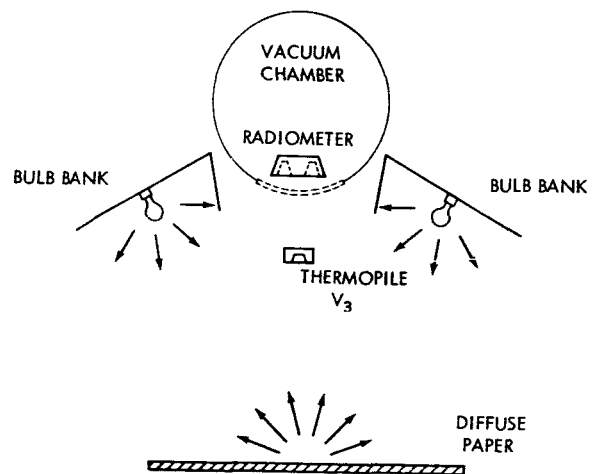


FIGURE 13.—Schematic of the reflected solar radiation calibration of the low resolution radiometer.

bulbs. In Equation 2, the heat transfer equation, the term $S A \sin B$ may be represented by W_{bp} . A correction must be made for the transmittance (t) of the quartz window. With this correction

$$S A \sin B = V_3 \frac{t}{C} \quad (9)$$

By varying the number of tungsten bulbs used, the radiant emittance W_{bp} can be varied. The calibration of the radiometer proceeds in the following manner:

1. The radiometer housing temperature, black and white detector responses, and the quartz window temperature (T_q) are recorded prior to exposure to the tungsten bulb array.

2. The bulbs are turned on until an equilibrium response is attained for each detector and the above measurements are repeated. In addition, the thermopile output V_3 is recorded.

3. The measurements of steps 1 and 2 are repeated with an increasing number of bulbs in the array.

The unreduced calibration results consist of R_{Nb} , R_{Nw} , R_h , T_q , and V_3 . These parameters must be converted to temperatures except for V_3 which is converted to reflected radiant emittance by Equation 9. The corrections for the contribution of the thermal emission from the quartz window and housing to the irradiance of the detector are discussed later.

Time Constant Measurements

The time response of the detectors was measured by using the array of tungsten bulbs described above. The radiometer was placed to view the white diffuse paper through the quartz window. The bulbs were turned on until the detectors attained an equilibrium response, and the bulbs were then turned off. The temperatures of the quartz window and the radiometer housing, as well as that of each detector network, were continuously recorded as a function of time until the changes became negligible. Figure 14 shows the results of such a measurement with the detector

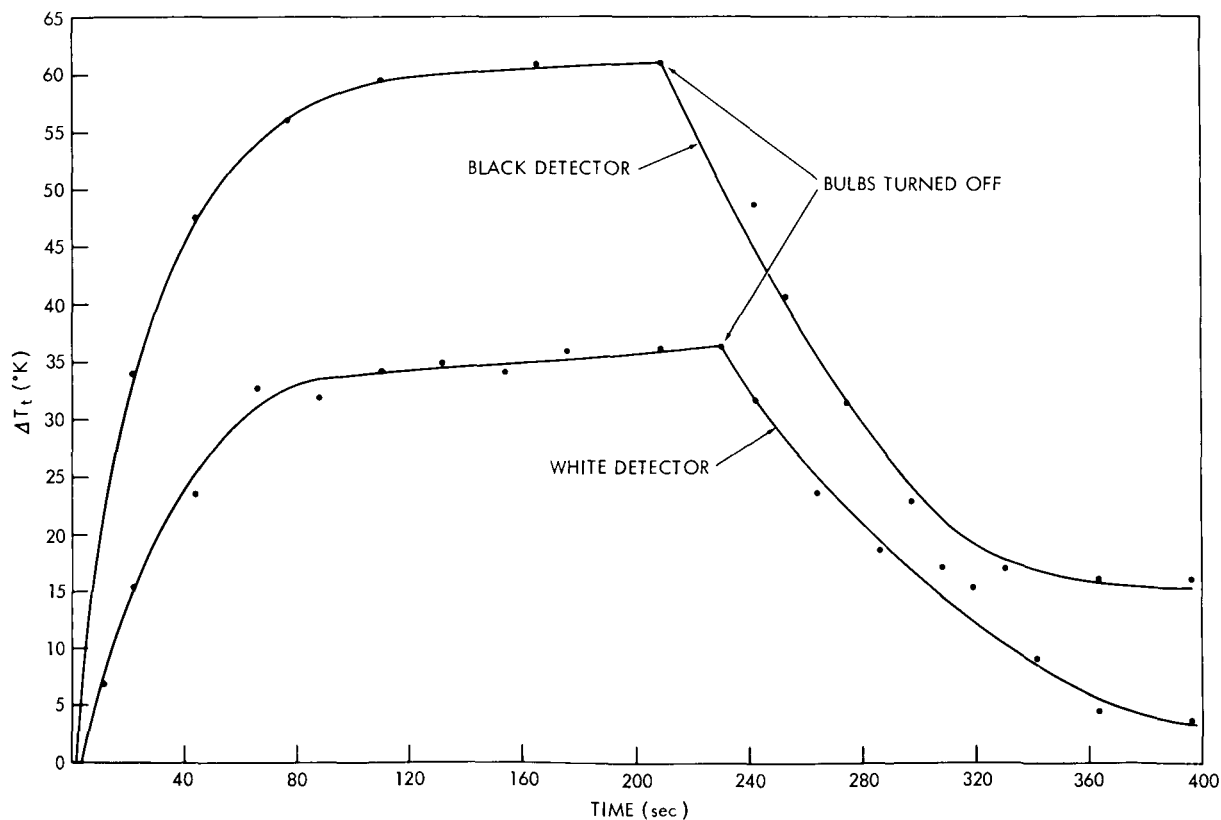


FIGURE 14.—Detector time constant determination for the low resolution radiometer. The time constant is determined from the rise portion of the curve. The housing temperature was about -20°C .

temperature differences plotted as a function of time for a housing temperature of about -20°C . The black and white detectors have very nearly the same time response and show the characteristic exponential rise and decay with time. These results give a value for the time constant of about 35 sec for the rise portion of the graph under the conditions of detector equilibrium temperature shown. The decay portion was not used in determining the time constant because the effects from the heating of the quartz window by the tungsten bulbs were judged to be too large.

Calibration of the Radiometer Components

Several other measurements were performed by the manufacturer which are necessary for the interpretation of the instrumental response. These consist of temperature-resistance calibra-

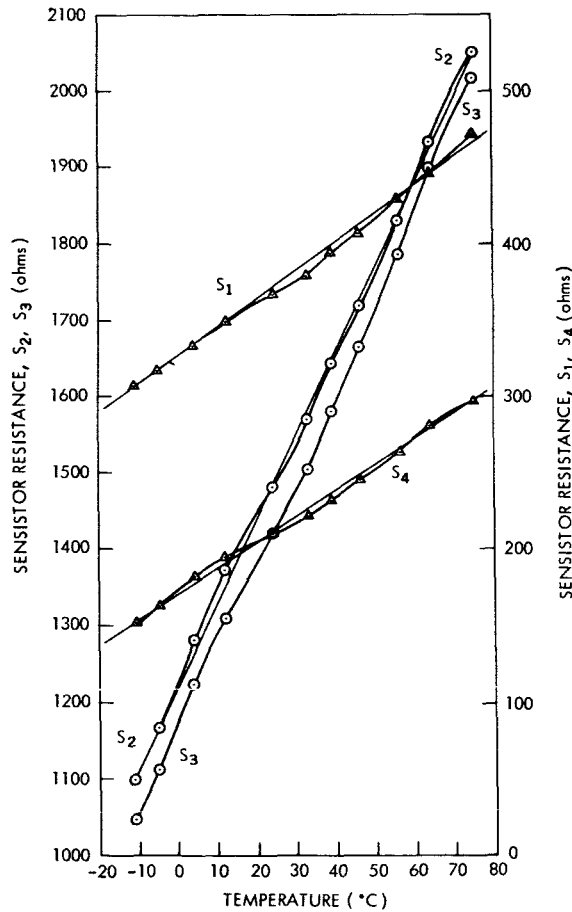


FIGURE 15.—Sensor response as a function of temperature.

tions of the detectors, sensistors, and other thermistors used in the radiometer. Figure 15 shows the sensistor response curves as a function of temperature. The sensistor resistance values are $S_i=S_1$, $S_j=S_2$ for the black detector and $S_i=S_4$, $S_j=S_3$ for the white detector. Figure 16 shows an example of the thermistor detector response as a function of temperature for the black and white detectors. Curves of the type shown in Figures 15 and 16 were used for the reduction of network resistance data to the required thermistor detector temperatures.

Subcarrier Oscillator Calibration

Generally, the resistance-controlled subcarrier oscillator used in the radiometer calibration was not flown in the satellite, and so a separate calibration for the flight oscillator was necessary. This calibration consisted of measuring the output frequencies and input resistances, with the temperature of the oscillator T_E as a parameter. A decade box was used to simulate the input resistances. The calibration curve for an oscillator is shown in Figure 17.

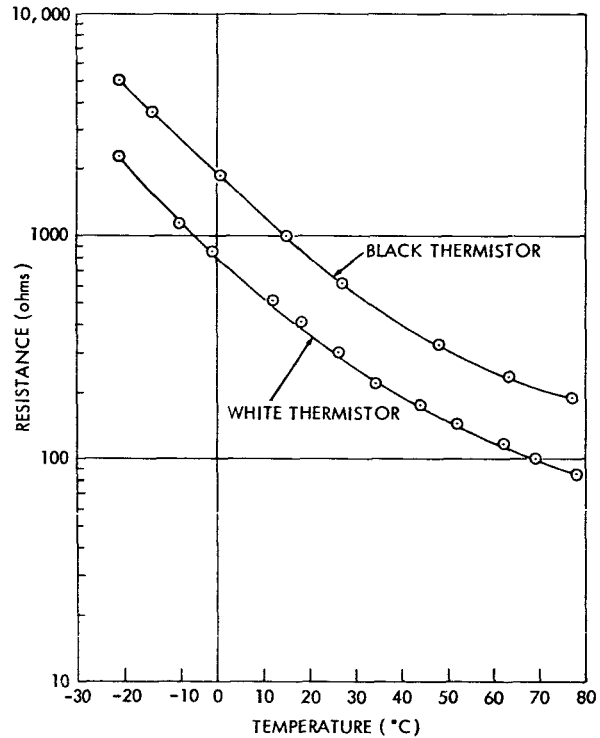


FIGURE 16.—Black and white thermistor detector response as a function of temperature.

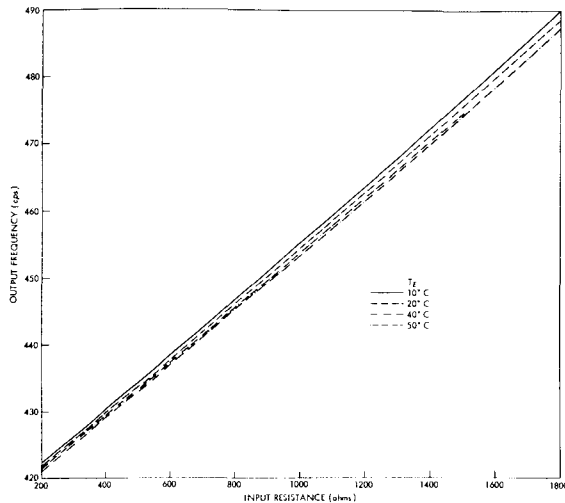


FIGURE 17.—Resistance-controlled subcarrier oscillator calibration curve. The temperature of the oscillator is shown as a parameter.

Additional Measurements

In addition to the individual radiometer calibration measurements, other measurements were performed to check, in general, the field of view, the spectral response of radiometer components, the spectral distribution of bulbs, and the degradation of spectral response due to exposures of ultraviolet radiation.

The field of view measurements were performed by the manufacturer. These were made by linearly traversing a point source of light across the field of view and measuring the radiometer response as a function of angle of the source from the optical axis. Results from such a measurement are shown in Figure 18. As was mentioned

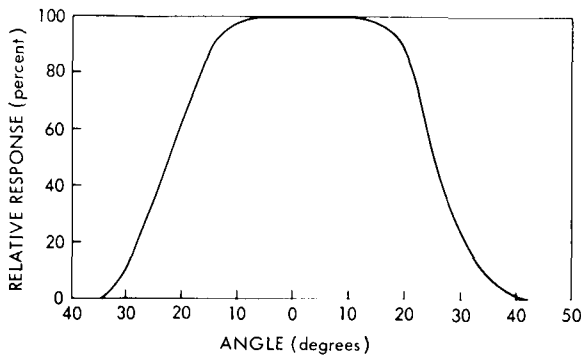


FIGURE 18.—Field of view of the low resolution radiometer.

previously, the field of view of each detector is a function of the angular aperture of the cone and also the ratio of cone length to detector radius. The measurement shown in Figure 18 gives a field of view of about 45 degrees or 0.6 steradian at the half-power points.

Spectral response measurements for the white paint coating are shown in Figure 19. The measurements were performed by the manufacturer using a Beckman DK-2 spectrophotometer for the spectral range from 0.2 to 3.5μ . The average total reflectance for this region, with the solar irradiance curve of Johnson (Reference 6) as a weighting function, is 77 percent. Reid and McAlister (Reference 7) have made measurements (Figure 20) on a similar white paint in the region of 2 to 15μ which show that the reflectance of the white paint in the region of 3.5 to 5.5μ is high. However, only 1.5 percent of the sun's total energy and 2.5 percent of the energy of a

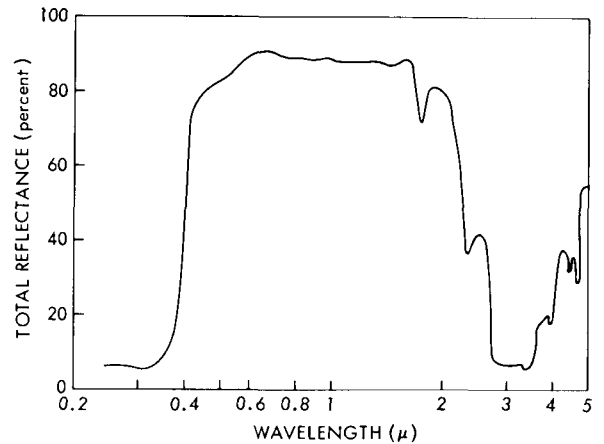


FIGURE 19.—Spectral reflectance of the white coating used on the thermistor detector for the region 0.2 to 3.5μ . The sample was 30μ thick on a backing of pure SiO_2 . The portion of the curve beyond 3.5μ is beyond the limits of the spectrophotometer used in making the measurements.

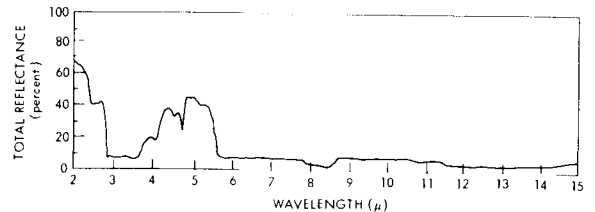


FIGURE 20.—Spectral reflectance of a white paint similar to that used for Figure 19 for the region 2 to 15μ .

300°K blackbody fall in this region, and these high reflectance values, therefore, are not critical to the measurements. Approximately 55 percent of the energy of a 300°K blackbody falls in the spectral region from 5.5 to 15 μ where the reflectance varies between 5 and 10 percent. At wavelengths greater than 15 μ , where 43 percent of the 300°K blackbody energy lies, no reflectance values are available.

The white paint coating was exposed to an ultraviolet source which would show whether any degradation of reflectance occurred. The paint samples were exposed to about 5 times the solar flux in the spectral region from 2000 to 4000A for various exposure times. Figure 21 shows the

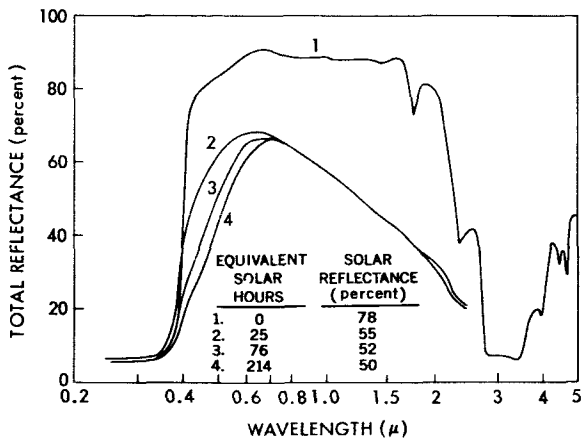


FIGURE 21.—The ultraviolet degradation of the white after 25, 76, 214 equivalent solar hours of ultraviolet exposure. The reflectance values were averaged by using the solar irradiance curve from Johnson (Reference 6) as a weighting function.

original solar reflectance of the white paint and the degradation effects of direct exposure to 25, 76, and 214 equivalent solar hours. The Tiros II and IV low resolution radiometer lifetimes were both approximately 1300 orbits and during this time each radiometer viewed the earth for about 430 hr. For an assumed mean earth albedo of 10 percent in this near-ultraviolet region, the equivalent solar exposure was about 43 hr. The exposure for Tiros III was considerably longer. Hence, long term changes in spectral response can be expected to occur during a radiometer's lifetime.

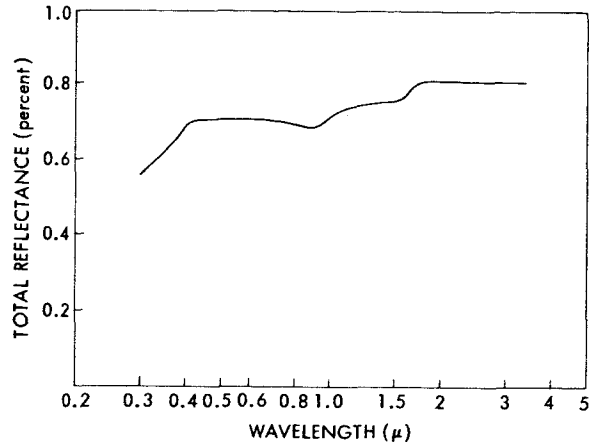


FIGURE 22.—Spectral reflectance of the cone of the low resolution radiometer for the region 0.3 to 3.0 μ .

Figure 22 shows the spectral reflectance of the aluminized mylar cone. In general, the reflectance decreases with decreasing wavelength. The diffuse spectral reflectance of the white paper which was used in the reflected solar radiation calibration, is shown in Figure 23. From 0.3 to 0.5 μ the reflectance increases, from 0.5 to 1.3 μ it is uniform, and from 1.3 to 2.8 μ it decreases. The energy distribution of the blue photoflood bulbs is shown in Figure 24. The peak of the distribution occurs at 1.25 μ which corresponds to a blackbody temperature of 2300°K. The deviation of the energy distribution in the region of 0.7 μ , from that of a blackbody function for a tungsten filament of 2300°K, is due to the peculiar transmission properties of the blue envelope of the bulbs in this region. The deviation from a blackbody distribution at the long wavelength region is due to the lower transmittance of the glass envelope.

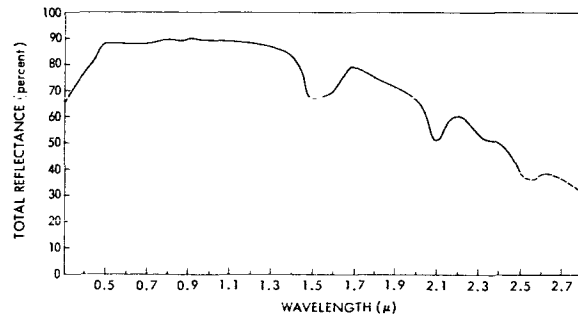


FIGURE 23.—Spectral diffuse reflectance of the white paper used in the reflected solar radiation calibration.

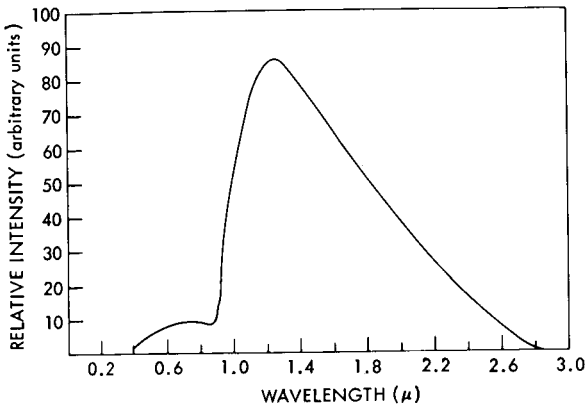


FIGURE 24.—Relative intensity of the blue photoflood bulbs used in the reflected solar radiation calibration. The energy peak is at 1.25μ which corresponds to a blackbody temperature of 2300°K .

The spectral response $\phi(\lambda)$ of the radiometer was assumed to be fairly constant with wavelength in the spectral region of 0.2 to 2.5μ . This assumption is equivalent to the radiometer response being proportional to the total energy of the source regardless of the source spectral distribution, as long as the bulk of the energy output falls in the above region. The spectral response of the radiometer $\phi(\lambda)$ is proportional to the product $R_c(\lambda) \alpha_t(\lambda)$, where $R_c(\lambda)$ is the spectral reflectance of the reflecting cone and $\alpha_t(\lambda)$ is the spectral absorptance of the painted thermistor detector. It is assumed that $\alpha_t(\lambda) = 1 - R_p(\lambda)$, where $R_p(\lambda)$ is the spectral reflectance of the paint coating on the thermistor detector. If the paint coating is too thin this will be a bad assumption. For comparison purposes an auxiliary spectral response function $\phi'(\lambda)$ will be considered which is defined by $\phi'(\lambda) = \phi(\lambda)r_p(\lambda)$. The diffuse reflectance of the white paper is included because of its effect on the spectral energy distribution of the source. The above definition is somewhat arbitrary and the $r_p(\lambda)$ could have been included as part of the sources, but for the comparison to be presented, the above definition will be used. The reflectance of the black paint is assumed to be zero for all wavelengths. The auxiliary spectral response functions pertaining to the black and white detectors are now

$$\phi'_b(\lambda) = r_p(\lambda)R_c(\lambda)$$

$$\phi'_w(\lambda) = r_p(\lambda)R_c(\lambda)[1 - R_p^*(\lambda)]$$

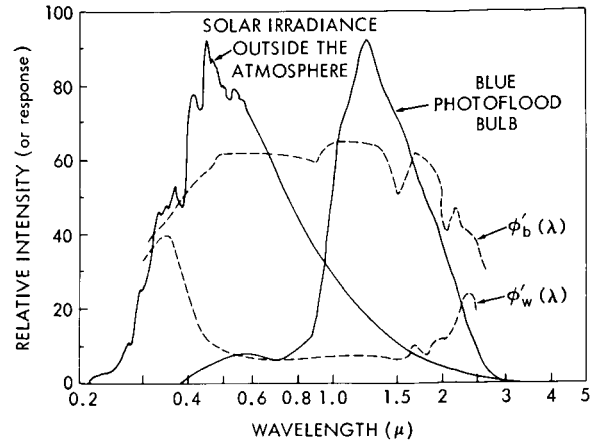


FIGURE 25.—The spectral response curves pertaining to the black and white detector are shown along with the solar irradiance curve from Johnson (Reference 6) and the relative intensity curve for the blue photoflood bulbs.

The functions $r_p(\lambda)$, $R_c(\lambda)$, and $R_p^*(\lambda)$ have been shown in Figures 19, 20, 22, and 23. The functions $\phi'_b(\lambda)$ and $\phi'_w(\lambda)$ are shown in Figure 25. The energy distribution for the photoflood bulbs and the solar irradiance curve for zero air mass by Johnson (Reference 6) are also shown in Figure 25. The peak intensities of the two curves are normalized to the same value. As can be seen from the $\phi'(\lambda)$ curves of Figure 25, the assumption of a flat spectral response is not very good near the upper and lower wavelength limits of $\phi'_b(\lambda)$ and $\phi'_w(\lambda)$. However, to a first approximation a suitable mean value for $\phi'_b(\lambda)$ and $\phi'_w(\lambda)$ may be selected. The deviation of $\phi'_w(\lambda)$ from its mean is most pronounced for wavelengths shorter than 0.4μ . Since less than 10 percent of the solar energy lies in this region, the first approximation is still reasonable. Figure 25 indicates that since $\phi'_w(\lambda)$ and $\phi'_b(\lambda)$ are not constant, the radiometers should be calibrated spectrally.

Sources of Error

The sources of error in the calibration are numerous and difficult to evaluate numerically in terms of both random and systematic effects. The net result of these errors manifests itself in the accuracy to which the coefficients of the energy balance equations can be determined.

The precision (random error) of temperature measurements of the calibration is considered to be $\pm 1.0^\circ\text{K}$ for T_t and $\pm 0.5^\circ\text{K}$ for T_s and T_e .

These result from the various conversions from resistance to temperature. The dispersion of T_t also includes a random error caused by the method of measuring resistances via frequency readings from the subcarrier oscillator. The oscillator used during calibration has a small local variation about its mean calibration curve. The calibration curve also drifts slowly in time but this systematic effect is largely eliminated since the frequency to network resistance conversions (via the decade box) are made immediately after the measurement.

The accuracy (systematic error) of the calibration is considered to be $\pm 2.0^\circ\text{K}$ for T_t and $\pm 1.0^\circ\text{K}$ for T_s and T_e . Some of the inherent inaccuracies of the calibration arise in the apparatus used to measure and record the various temperatures. The measuring devices have been described in the calibration section. In the thermal calibration, the cavity and housing temperatures are maintained by circulating liquids. Any gradient between the temperature of the point of measurement and the controlling temperature may cause a systematic error. It is also quite possible that the temperature difference would consist of random fluctuations and thus contribute to the random error.

The reflected solar radiation portion of the calibration is subject to considerably greater error than the thermal part. This is partially due to the greater number of intermediate measurements required; in addition, the types of measurements are inherently subject to greater error. The various sources of error (without regard to the random or systematic effect) are discussed below.

The white diffusely reflecting paper was assumed to be a Lambert's Law surface and later measurements indicate a slight departure from this law for large angles of incidence. For this reason, and also to avoid any specular reflectance components, large angles of incidence were avoided.

Because the field of view of the thermopile is smaller than that of the radiometer, it is necessary to irradiate the paper uniformly so that the radiometer and thermopile observe the same radiant emittance. The large field of view of the radiometer requires that the surface area of the paper be large; it becomes difficult to produce a uniform irradiance over the entire surface area.

In addition, it is difficult to maintain a uniform irradiance because the bulb voltages are not controlled accurately and because the bulbs deteriorate at different rates, thereby causing asymmetries in the bulb outputs.

The radiometer views the paper through the quartz window of the vacuum chamber. The window cannot be cooled to eliminate or minimize its thermal emission since frosting begins to appear. Hence, during the tests, the window is usually left at room temperature and any slight changes in its temperature due to absorption of infrared energy by the window from the bulbs is monitored by a thermocouple. The quartz temperature is measured before, half-way through, and after the test. A correction is made for the quartz thermal emission contribution to the detectors and the only error due to this effect would be the error in the quartz temperature measurement.

Reduction of the Calibration Data

The network resistance R_N must be converted to detector temperature T_t before the coefficients for the heat transfer equation can be evaluated. This conversion is illustrated by the portion of the flow diagram of Figure 26 enclosed by the dashed lines. The portion of the flow diagram outside the dashed lines will be discussed later. S_i and S_j , which are functions of the housing temperature, are manually determined as part of the input to an electronic computer program which reduces the network resistance to detector resistance via Equation 3. The other input element of the computer program is R_N . The computer output is the detector resistance which is manually converted to the detector temperature.

DETERMINATION OF THE COEFFICIENTS OF THE HEAT TRANSFER EQUATION

The coefficients of the heat transfer equation of each detector are determined from the calibration data. The heat transfer equation used in the determination of the coefficients is Equation 2 with the conduction term omitted. The coefficients are determined by the method of least squares and the resulting set of normal equations is solved simultaneously by using Cramer's rule for the desired coefficients. This type of solution must be checked carefully to insure that the

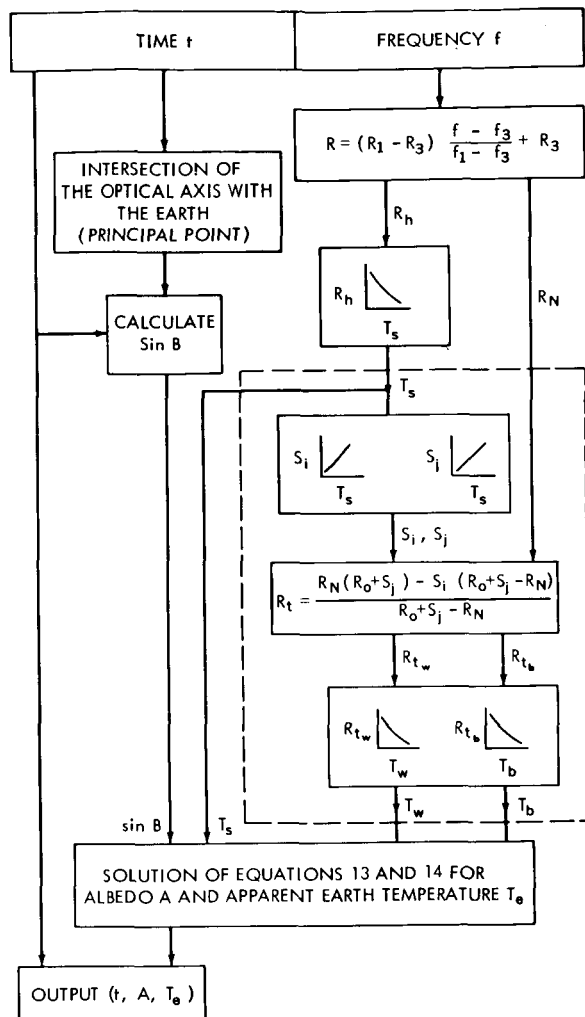


FIGURE 26.—Flow diagram for the reduction of the low resolution radiometer data.

co-efficients are not affected by normal equations closely approaching dependency and creating a solution which is nearly indeterminate. The radiative and conduction coefficients are evaluated from the thermal calibration data, and the reflected solar radiation coefficient from the reflected solar radiation calibration data.

Thermal Coefficients

The thermal calibration data, in terms of temperature, are shown in Figure 27 for the black detector of the Tiros IV radiometer. For the appropriate values of the constants k_1 , k_2 , and k_2' , a least squares method was first applied to the

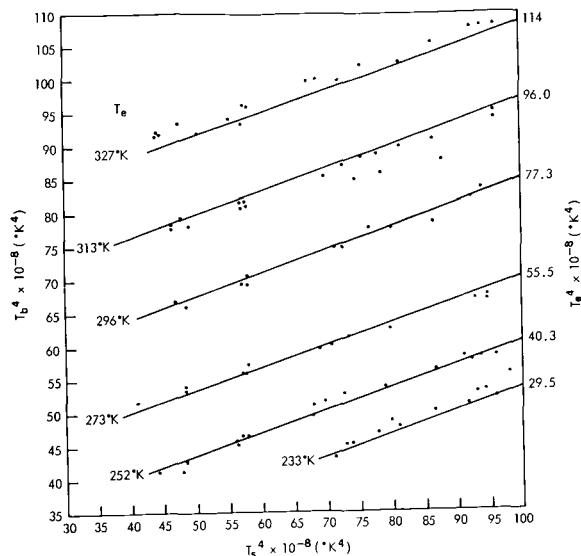


FIGURE 27.—Thermal calibration data and least squares solution (solid lines) used in determining thermal coefficients for the black detector of the low resolution radiometer on Tiros IV. The apparent earth blackbody temperature is the parameter for the various curves.

thermal portion of the heat transfer equation (2):

$$T_t^4 = k_1 T_e^4 + k_2 T_s^4 + k_2' (T_s - T_t) \quad (10)$$

In applying the least squares method the variable whose residual error is to be minimized must be chosen. The conventional choice is the variable which is to be predicted (T_e) under the assumption that the errors in the other measured quantities (T_t , T_s) are small. This was not done in the present case since the T_e measurement is fairly accurate with respect to T_t . Therefore, the residual error of T_t was minimized in these calculations.

In Equation 10 no acceptable solution of the 3 normal equations was obtained because of the unavoidable loss, through subtraction, of the 3 or 4 leading significant figures in the expressions for the k 's. The loss of significant figures reflects the nearly indeterminate state of the solution. As the input temperatures are only significant to three figures, the resulting k 's cannot be considered reliable. The cause for the unreliability of the coefficients, in terms of the three nearly linearly dependent normal equations, is that instead of one unique solution for the coefficients, an infinite number of solutions for the normal

equations exist in this nearly indeterminate situation. Therefore, a different method of solution for the coefficients must be employed or the heat transfer equation must be altered so that the least squares solution becomes more determinate.

If the conductive term of the heat transfer equation is neglected, Equation 10 becomes

$$T_t^4 = k_1 T_e^4 + k_2 T_s^4 \quad (11)$$

Mathematically, the conductive term is ignored but physically the conductive effects are absorbed in the values of k_1 and k_2 . The solution of the two normal equations for this case involves the loss of only one or two significant figures and yields coefficients that approximate the physical situation. In Figure 27 the least squares solution for all the data points is indicated by the solid line. The coefficients for this solution and their estimated probable errors are

$$k_{1b} = 0.65 \pm 0.02 .$$

$$k_{2b} = 0.34 \pm 0.02 .$$

A similar solution for the white detector of Tiros IV gives

$$k_{1w} = 0.51 \pm 0.02 .$$

$$k_{2w} = 0.49 \pm 0.02 .$$

The thermal coefficients for Tiros IV, and also those of Tiros III, are listed in Table 2. The higher value of k_{1b} with respect to k_{1w} indicates that in the infrared the black detector has a greater average spectral absorptance than the white detector. The k_1 's are lower and the k_2 's higher than called for by design with the consequence that the thermal coupling between the detector and the satellite is fairly strong com-

pared with the coupling between the detector and the earth. The detector temperature is therefore strongly influenced by the satellite temperature. The least squares probable errors are $\pm 0.9 \times 10^8$ $^{\circ}\text{K}^4$ for T_b^4 and $\pm 0.7 \times 10^8$ $^{\circ}\text{K}^4$ for T_w^4 . The probable errors for T_b and T_w are ± 0.9 $^{\circ}\text{K}$ and ± 0.7 $^{\circ}\text{K}$, respectively.

An attempt was made to estimate the conduction constant k_2' by using a unique solution for a given T_e curve. The expression for k_2' consists of terms such as $(T_t^4 - T_i^4)$ and $(T_{s_i} - T_t)$. If the probable error in T_t is $\pm 1.0^{\circ}\text{K}$ the percentage error in the term $T_{s_i} - T_t$ varies from 4 to 20 percent as $T_{s_i} - T_t$ varies from 2.5 $^{\circ}\text{K}$ to 5 $^{\circ}\text{K}$. Under these conditions k_2' depends on the difference of quantities containing considerable error and no accurate value can be determined.

In summary, the evaluation of the thermal coefficients indicates that a representative value for the conduction coefficient cannot be determined. The optimum heat transfer equation for the detectors is considered to be Equation 11 and the best values for the thermal coefficients for Tiros III and IV are listed in Table 2. As the thermal coefficients are in part functions of $R_c(\lambda)$ and $\alpha(\lambda)$, improvement of the values of the coefficients to meet design specifications involves increasing $R_c(\lambda)$ and $\alpha(\lambda)$ in the infrared.

Reflected Solar Radiation Coefficient

The reflected solar radiation coefficient k_3 is determined from the energy balance equation for the detector:

$$T_t^4 - (k_1 T_q^4 + k_2 T_s^4) = k_3 S A \sin B \quad (12)$$

The term $(k_1 T_q^4 + k_2 T_s^4)$ is the correction for thermal effects of the quartz window and radiometer housing during the calibration. Because the

Table 2

Tiros III and IV Low Resolution Radiometer Coefficients.

Satellite	k_{1b}	k_{2b}	$k_{3b} \left(\frac{^{\circ}\text{K}^4 \text{ cm}^2}{\text{watt}} \right)$	k_{1w}	k_{2w}	$k_{3w} \left(\frac{^{\circ}\text{K}^4 \text{ cm}^2}{\text{watt}} \right)$
Tiros III	0.60	0.39	1210×10^8	0.51	0.44	807×10^8
Tiros IV	0.65	0.34	2214×10^8	0.51	0.49	1153×10^8

quartz window is nearly opaque to thermal radiation, the thermal contributions to the irradiance of the detector are felt to be due to the quartz window and the radiometer housing only. T_q , T_t , T_s , and the term $S A \sin B$ are obtained from the reflected solar radiation calibration, and the coefficients k_1 and k_2 have been determined previously, leaving k_3 as the only unknown in the equation. $S A \sin B$ is determined by Equation 9. The Tiros IV reflected solar radiation calibration data are shown in Figure 28. A least squares

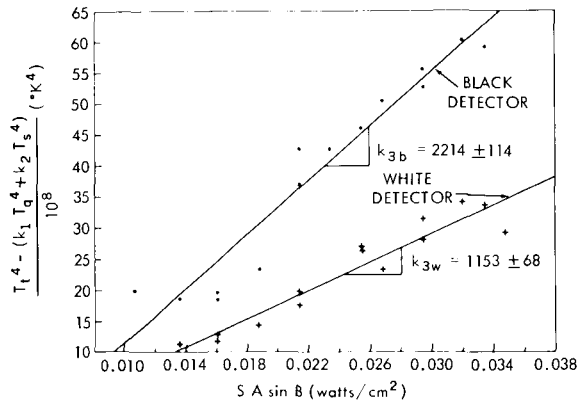


FIGURE 28.—Reflected solar radiation calibration data and least squares solution (solid lines) used to determine reflected solar radiation coefficients for the black and white detectors of the low resolution radiometer of Tiros IV.

fit used to find the values for k_3 gives $k_{3b} = (2214 \pm 114) \times 10^8 \text{ }^\circ\text{K}^4\text{-cm}^2/\text{watt}$ and $k_{3w} = (1153 \pm 68) \times 10^8 \text{ }^\circ\text{K}^4\text{-cm}^2/\text{watt}$. The magnitude of k_{3w} with respect to k_{3b} indicates that the reflectance of the white paint is poor in the visible region. The percentage error in k_3 is 5.2 and 5.9 percent for the black and white detectors, respectively. The least squares probable error in $[(T_t^4) - (k_1 T_q^4 + k_2 T_s^4)]$ is $\pm 2.9 \times 10^8 \text{ }^\circ\text{K}^4$ for the black detector and $\pm 1.7 \times 10^8 \text{ }^\circ\text{K}^4$ for the white detector. In the visible region, some indication of the spectral absorptance of the black detector relative to that of the white detector is given by the ratio of their relative response (k_{3b}/k_{3w}) which is only 1.9.

The reflected solar radiation coefficients for the Tiros III and IV wide field radiometers are listed in Table 2. A comparison of the coefficients indicates a large difference between the numerical k_3 values for Tiros III and IV. The values are $1210 \times 10^8 \text{ }^\circ\text{K-cm}^2/\text{watt}$ vs. $2214 \times 10^8 \text{ }^\circ\text{K-cm}^2/$

watt for k_{3b} and $807 \times 10^8 \text{ }^\circ\text{K-cm}^2/\text{watt}$ vs. $1153 \times 10^8 \text{ }^\circ\text{K-cm}^2/\text{watt}$ for k_{3w} . The difference is due either to different calibration conditions or a different spectral response between radiometers. The calibration conditions which could vary are the reflectance of the paper, the transmittance of the quartz window, or the color distribution of the photoflood bulbs. If any of the above variations did occur, it is not likely that the magnitude of the variation would be large enough to cause the observed differences. A difference in the spectral response for the two radiometers would be a more reasonable explanation since k_3 is, in part, a function of $R_c(\lambda)$ and $\alpha(\lambda)$ in the visible region. Since $\alpha(\lambda)$ should be fairly constant in the visible region, especially for the black detector, it would seem that the cone reflectance is responsible for this effect. This was borne out by an inspection of several other radiometers, which showed large variations in the thicknesses of the aluminum coatings, some being partially transparent.

DATA PROCESSING AND REDUCTION

The telemetry of data from the satellite has been previously described (Reference 8). The designated bandwidth for the telemetry of the low resolution radiometer information is 425 to 475 cps (designated channel 6). The resistance-controlled, phase-shift subcarrier oscillator for channel 6, with output frequency nearly linear with respect to input resistance, converts the input resistances to frequencies. Channel 6 is a time-sharing channel, the time sharing controlled by a mechanical commutator which switches the appropriate resistor into the oscillator for a 6 sec sampling period. There are 5 commutated subchannels and the fifth includes a group of 7 to be subcommutated. Included in the first 4 commutations are the wide field sensor data and the radiometer housing temperature data. The fifth commutation includes resistances which represent environmental temperatures, instrumentation cannister pressure, and fixed resistor telemetry calibrations. The latter consist of fixed resistors near the lower (R_3) and upper (R_1) limits of the bandwidth which give reference points for calibration of the channel 6 frequencies. The oscillator converts the input resistance into output frequency signals which are amplified and recorded on a magnetic tape located in the satellite. A

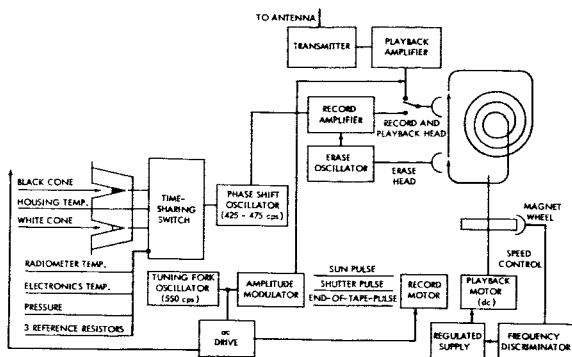


FIGURE 29.—Block diagram of the satellite instrumentation used by the low resolution radiometer.

block diagram of the satellite instrumentation is shown in Figure 29. Upon interrogation of the satellite, the tape signal is telemetered to the acquisition station and recorded on magnetic tape. This tape is sent to the ground station where it is demodulated and decommutated. The demodulation and decommutation process is illustrated in Figure 30. The data are now in analog form in terms of frequencies. A typical analog record is presented in Figure 31.

The frequencies are read directly from the analog record with a Gerber Semi-Automatic Analog to Digital Conversion System. The analog-to-digital conversion is made by a digital voltmeter whose output is fed through a Flexewriter to produce a punched tape output. This tape serves as an input for an LGP-30 computer program. Variable calibration is provided by the Gerber System through the use of zero offset and scale factor controls. As long as the analog frequency output remains linear, within the fre-

quency interval encompassed by the fixed resistors, the analog readings can be made in terms of frequency. The frequency readings are taken as a function of time which is also entered on the punched tape. The reduction of frequencies to T_e and A is illustrated in the flow diagram of Figure 26. Under the assumption that the frequency vs. resistance curves (with T_E as a parameter) are linear, the conversion from frequency to resistance is made by the following equation:

$$R = (R_1 - R_3) \frac{f - f_3}{f_1 - f_3} + R_3$$

The portion of the flow diagram used for conversion of resistances to temperatures has already been discussed. The temperatures can now be introduced into the heat transfer equations for the black and white detectors and the simultaneous solution of these two equations yields T_e and A . The computer output gives T_e and A as a function of time. The expressions for A and T_e are

$$A = \frac{k_{1w} T_b^4 - k_{1b} T_w^4 + (k_{2w} k_{1b} - k_{2b} k_{1w}) T_s^4}{(k_{3b} k_{1w} - k_{3w} k_{1b}) S \sin B} \quad (13)$$

$$T_e^4 = \frac{k_{3b} T_w^4 - k_{3w} T_b^4 - (k_{3b} k_{2w} - k_{3w} k_{2b}) T_s^4}{k_{3b} k_{1w} - k_{3w} k_{1b}} \quad (14)$$

As can be seen from the channel 6 sequencing, four commutated measurements exist between an individual sensor temperature sampling, so that T_e and A can be determined every 30 sec. The albedo must be corrected for the angle of solar incidence for every satellite position.

The orbital data provide the time-correlated geographic location of the intersection of the spin axis with the earth. Since this intersection is also, practically speaking, the center of the low resolution radiometer field of view, the radiometer data are geographically located as a function of time.

Because of instrumental and calibration difficulties, only selected orbits of Tiros II low resolution data have been reduced. Since both the low resolution radiometer and the University of Wisconsin experiment are recorded on the commutated channel 6 of Tiros III and IV, it has been found convenient to program the reduction of their data simultaneously on a CDC-1604 com-

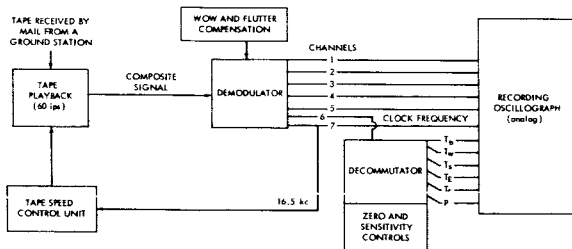


FIGURE 30.—Block diagram of the information flow at the data processing center in producing the radiation data of the low resolution radiometer and of the five medium resolution channels in analog form. The output of the decommutator consists of temperatures of the black and white detectors, three environmental temperatures, and the canister pressure P.

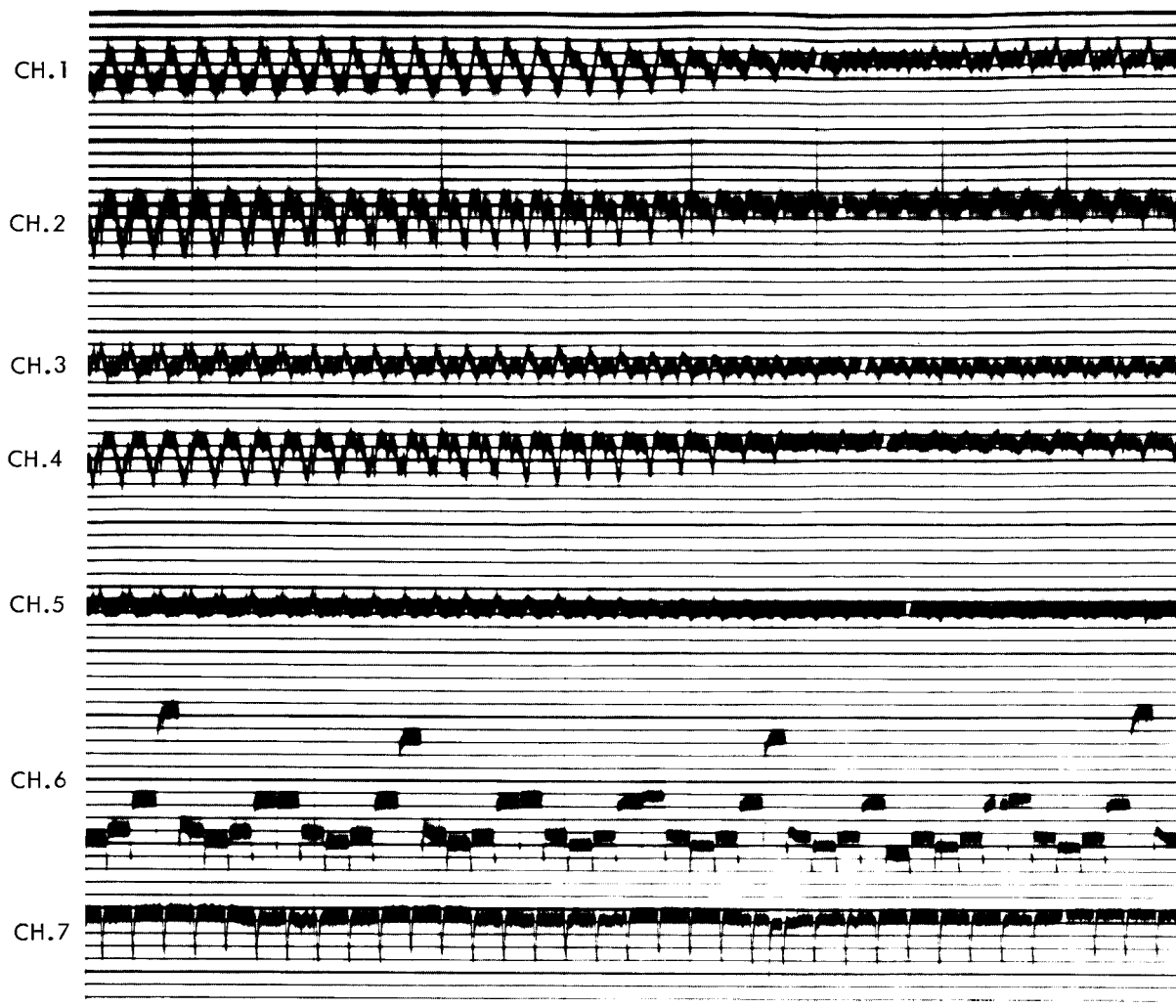


FIGURE 31.—Oscillogram showing all channels of the radiometer experiment. They are, from top to bottom, the five medium resolution channels in numerical order, the commutated channel 6, and the envelope of the clock frequency showing sun sensor pulses.

puter at the University of Wisconsin. This program is now in the final stages of checkout, and further results and conclusions must await its completion.

Selected orbits of Tiros III low resolution data have been reduced by the method described in the last section and presented in the literature (Reference 9). One of these orbits (43) will be presented here as an example of the reduced data. The geographic location of the orbit considered is over North Africa with the radiometer viewing area passing from the cloudless Mediterranean Sea over the Libyan Desert to the tropical regions

of the Sudan. Figure 32 illustrates the observed equilibrium detector temperatures, the observed radiometer housing temperature, and the resultant apparent earth temperature and albedo as a function of time. The values of T_e and A were calculated from the detector and housing temperatures indicated by the individual points and they show considerable scatter. The broad spectral response and relatively long time constant of the radiometer would indicate that the curves for albedo and earth temperature are slowly changing functions similar to the curves for the detector temperatures. In general, the predicted apparent

earth temperature curve would approach the mirror image of the albedo curve. However, it is evident from Figure 32 that sharp fluctuations do exist in the values for T_e and A . These fluctuations are introduced by the radiometer, as will be shown below. In an effort to smooth the T_e and

A fluctuations, curves were drawn to smooth the data points of detector temperature and housing temperature; these curves are also shown in Figure 32. The results for T_e and A shown in Figure 33 were calculated from the smoothed curves of detector and housing temperatures.

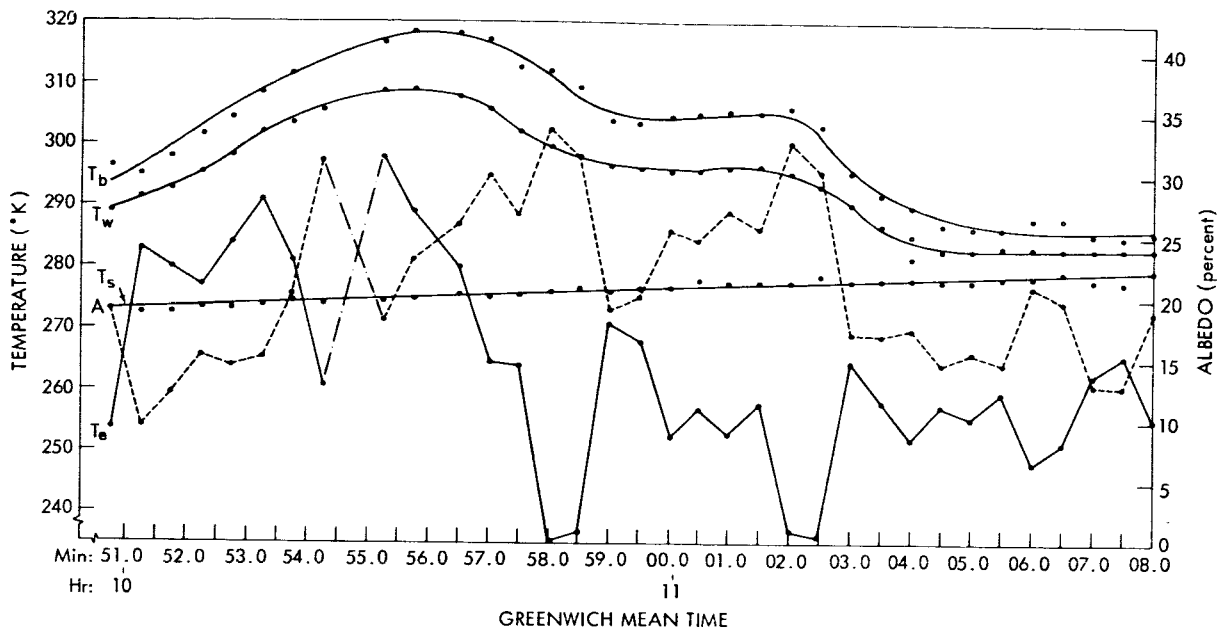


FIGURE 32.—Reduced channel 6 data for orbit 43, Tiros III, July 15, 1961. The equilibrium detector temperatures (T_b , T_w), radiometer housing temperature (T_s), apparent earth blackbody temperature (T_e), and albedo (A) are shown as a function of time. T_e and A have been calculated from the individual points of T_b , T_w , and T_s .

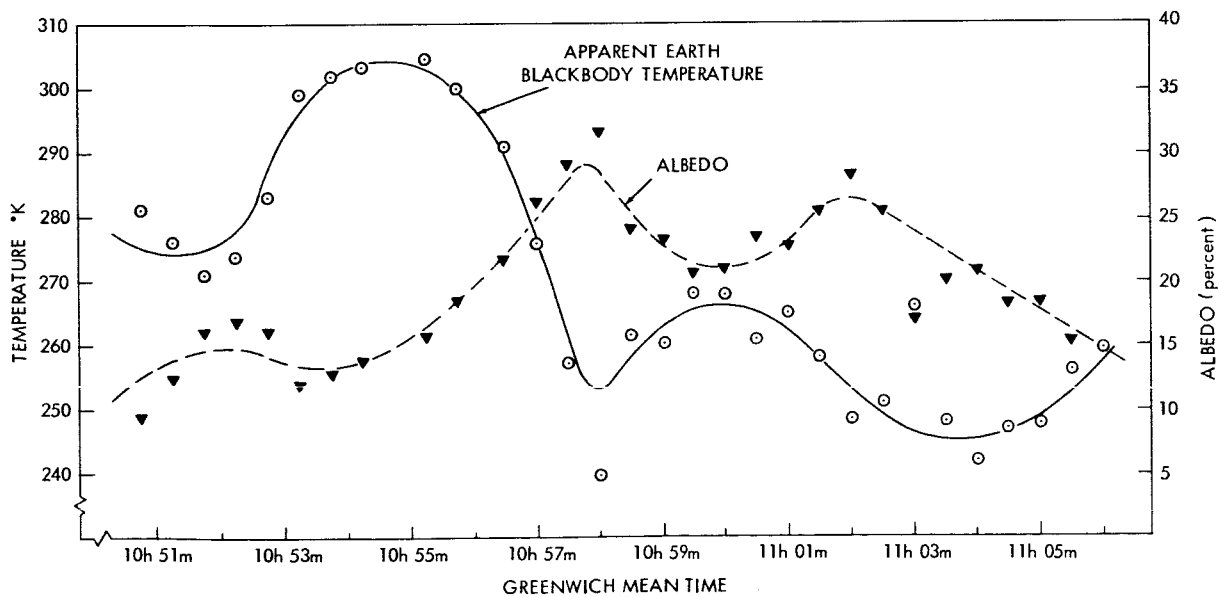


FIGURE 33.—Smoothed results of T_e and A . T_e and A have been calculated from the curves of T_b , T_w , and T_s shown in Figure 32.

In order to gain some insight into the fluctuations of T_e and A, their total differentials and

probable errors have been calculated. The total differentials for A and T_e are

$$dA = \frac{\partial A}{\partial k_{1b}} dk_{1b} + \frac{\partial A}{\partial k_{2b}} dk_{2b} + \dots + \frac{\partial A}{\partial T_s} dT_s \dots \quad (15)$$

$$dT_e = \frac{1}{4T_e^3} \left(\frac{\partial T_e^4}{\partial k_{1b}} dk_{1b} + \frac{\partial T_e^4}{\partial k_{2b}} dk_{2b} + \dots + \frac{\partial T_e^4}{\partial T_s} dT_s + \dots \right) \quad (16)$$

The evaluation of the total differentials of T_e and A was made for several different cases and the results are listed in Table 3. Mathematically, the large magnitude of the coefficients in the differentials is the reason that the small changes in detector temperatures are magnified into the resultant large changes in A and T_e . Physically, the large magnitudes of the coefficients are due to the values of the optical characteristics of the

radiometer. If the k's were to meet design specifications, the magnitudes of the coefficients in the differentials would decrease and thus would tend to decrease the large fluctuations of T_e and A. Any remaining unreasonable fluctuations of T_e and A would then be inherent in the nature of the physical process.

The probable errors for A and T_e were calculated from

$$[\text{p.e.}(A)]^2 = \left(\frac{\partial A}{\partial k_{1b}} \right)^2 [\text{p.e.}(k_{1b})]^2 + \dots \quad (17)$$

$$[\text{p.e.}(T_e)]^2 = \left(\frac{1}{4T_e^3} \frac{\partial T_e^4}{\partial k_{1b}} \right)^2 [\text{p.e.}(k_{1b})]^2 + \dots \quad (18)$$

and are also included for the cases in Table 3. The probable errors for the various parameters in Equations 17 and 18 were assumed to be

$$\left. \begin{array}{l} \text{p.e.}(k_{1b}) \\ \text{p.e.}(k_{2b}) \\ \text{p.e.}(k_{1w}) \\ \text{p.e.}(k_{2w}) \end{array} \right\} = \pm 0.02$$

$$\text{p.e.}(k_{3b}) = \pm 63 \times 10^8 \frac{^\circ\text{K}^4 \text{ cm}^2}{\text{watt}} (\pm 5.2 \text{ percent})$$

$$\text{p.e.}(k_{3w}) = \pm 48 \times 10^8 \frac{^\circ\text{K}^4 \text{ cm}^2}{\text{watt}} (\pm 5.9 \text{ percent})$$

$$\left. \begin{array}{l} \text{p.e.}(T_b) \\ \text{p.e.}(T_w) \end{array} \right\} = \pm 1^\circ\text{K}$$

$$\text{p.e.}(T_s) = \pm 0.5^\circ\text{K}$$

In Table 3 the probable error in the albedo can be seen to increase from ± 8 to ± 14 percent from case I to III with increasing $(T_b - T_w)$. For case

I the contribution of any one parameter to the resultant probable error in A is equal to the contribution of any other parameter. As the $(T_b - T_w)$ interval increases, the error contributions from k_{3b} and k_{3w} become the dominant factor in the determination of the error in A. Only one case was evaluated for the probable error in T_e , which was $\pm 24^\circ\text{K}$. As the probable error for T_e depends inversely on T_e^3 , the $\pm 24^\circ\text{K}$ is a minimum error since T_e was equal to 302°K for case IV.

The resultant probable errors for T_e and A can be reduced by: (1) decreasing the magnitude of the coefficients in the expressions for the probable error (for example $\partial A / \partial k_{1b}$); or (2) reducing the probable errors of the individual parameters for example, $[\text{p.e.}(k_{1b})]$ which are introduced by errors in data processing (errors in temperature) or calibration (errors in k's). The magnitude of the coefficients can be decreased as discussed in the previous paragraph. The errors in data processing ultimately appear in the values of T_b , T_w , and T_s and it is felt that these errors have been minimized. The errors in calibration appear

Table 3
Total Differentials and Probable Errors for the Albedo
and Earth Temperature, Tiros III.

Case	Temperatures (°K)			Total Differentials for Albedo and Earth Temperature	Albedo (percent)	Earth Temperature (°K)	Probable Error (percent or °K)
	T _b	T _w	T _s				
I	285	280	273	$dA = -120 dk_{1b} + 180 dk_{2w} + 120 dk_{1w} - 150 dk_{2b}$ $+ (580 \times 10^{-12}) dk_{3w} - (490 \times 10^{-12}) dk_{3b} - 2.8 dT_w$ $+ 2.5 dT_b + 0.28 dT_s$	13	-	±8
II	300	290	273	$dA = -100 dk_{1b} + 180 dk_{2w} + 100 dk_{1w} - 150 dk_{2b}$ $+ (1100 \times 10^{-12}) dk_{3w} - (860 \times 10^{-12}) dk_{3b} - 3.1 dT_w$ $+ 3.0 dT_b + 0.28 dT_s$	24	-	±10
III	315	300	273	$dA = -69 dk_{1b} + 180 dk_{2w} + 58 dk_{1w} - 150 dk_{2b}$ $+ (1740 \times 10^{-12}) dk_{3w} - (1480 \times 10^{-12}) dk_{3b}$ $- 3.5 dT_w + 3.4 dT_b + 0.28 dT_s$	39	-	±14
IV	318	309	275	$dT_e = 455 dk_{1b} + 319 dk_{2b} + (0.16 \times 10^{-8}) dk_{3b}$ $- 686 dk_{1w} - 473 dk_{2w} - (0.21 \times 10^{-8}) dk_{3w}$ $- 7.1 dT_b + 9.8 dT_w - 1.2 dT_s$	-	302	±24

in the various k constants and the reduction of these errors involves refinement of the calibration techniques. This refinement should be centered around improvement of the values for the reflected solar radiation coefficients. Specific areas in the reflected solar radiation calibration that can be improved are: solar simulator source, spectral calibration, control of intensity variation, the diffusely reflecting surface (uniformly white), and thermopile measurements. With implementation of improved procedures, it is reasonable to expect that the percentage errors for the k₃'s should be reduced to the order of ±2 percent. Then, for case III the probable error in albedo would be reduced from ±14 to ±8 percent and the contribution of the individual parameters to the resultant probable error would be equal. For case I the probable error in albedo would be reduced from ±8 to ±7 percent, and for case IV,

the probable error in T_e would be reduced from ±24°K to ±21°K. To further reduce the probable errors in A and T_e an overall improvement in spectral response and a reduction in conductive coupling are needed. Therefore, it is felt that the percentage errors of ±7 percent, and ±21°K in A and T_e, obtained using the improved k₃ constants, are optimum values within the state of the art limitations.

DATA COVERAGE LIMITATIONS

The value of the low resolution radiometer in measuring the heat balance of the earth is limited because of the data acquisition limitations of Tiros. Since the satellite is space oriented it is launched so that the favorable picture taking time (minimum nadir angle) occurs at local noon. Through magnetic attitude control, this relationship is maintained as nearly as possible. The

optical axes of the Tiros camera and the low resolution radiometer are coincident, the low resolution radiometer field of view approximating half the angular field of view of the cameras. Therefore, in general, the low resolution radiometer data are centered, in time, within 2 hr. of local noon. Initial launch conditions limit the latitudinal data coverage to the range of 48° North to 48° South. A latitudinal and longitudinal limitation is also imposed on the data coverage by precession of the satellite orbital plane and location of the interrogation stations.

The limitations imposed on low resolution radiometer data acquisition by spin stabilization, initial launch conditions, precession, and the location of interrogation stations reduce the potential of the system in obtaining short and long range planetary heat balance. Finding the value of short range heat balance, which would include investigation of diurnal effects, cannot be accomplished because the data are centered within 2 hr of local noon, and on a latitudinal and longitudinal basis only a small number of data points are available (Reference 10). The long range heat balance cannot be obtained, except as a longitudinal effect, because of insufficient earth coverage in the polar regions. Statistically, even for a long range longitudinal heat balance the number of data points available is not considered to be sufficient.

POST-LAUNCH OBSERVATIONS

The black detector network of the Tiros II and IV low resolution radiometers failed to respond to incident radiation at orbits 1373 and 1150, respectively. Radiation data were recorded for Tiros III for 2975 orbits and during this period no low resolution radiometer failures were observed.

In order to explain the analog pattern produced by the black detector network, after its failure in Tiros II and IV, it was necessary that the detector leg of the resistance network contain an open circuit. The detector leg consists of the black detector and a sensistor. The black detector is considered to be more susceptible to failure; its shape and composition allow warping and flaking of the detector element. A reasonable explanation for the failure is that thermal stress, due to expansion and contraction of the bolometer element, caused an open circuit in the detector leg. During the early stages of failure, the open circuit

occurred when the thermistor element expanded because of the heating by absorption of terrestrial and reflected solar radiation. As the viewing area returned to outer space the thermistor contracted and the circuit again closed to normal. Shortly after this initial phase, the open circuit became permanent because of repeated expansion and contraction of the detector element, and the network resistance then consisted of the temperature sensor and the fixed resistance in series.

SUMMARY OF THE TIROS LOW RESOLUTION RADIOMETER PROGRAM

Certain deficiencies of the radiometer became evident from the laboratory calibration data. Perhaps the most important was the excessive thermal coupling between the radiometer detectors and the satellite. The white detector did not adequately reflect the solar spectrum nor did it adequately absorb the terrestrial infrared emission. The cone reflectance was considered to be too low in value, thereby decreasing the optical efficiency. These two factors contribute to the excess thermal coupling with the satellite.

The above deficiencies manifested themselves in the reduced data by contributing to the large errors in the resultant values of T_e and A as shown in the orbit 43 data. Some of the resultant error is inherent in the physical nature of the process, since the solution of the heat transfer equations for T_e and A involves temperatures to the fourth power and differences of large numbers. Additional sources of error in T_e and A are errors in data processing and calibration.

A considerable amount of radiation data has been obtained from the Tiros III and IV low resolution radiometers. To date, only a few orbits of these data have been studied extensively because the automatic data reduction system is not yet operational. Because of large uncertainties, the absolute values of T_e and A have limited application in studies utilizing the radiation data. However, the relative measurements of T_e and A should still provide useful information for the areas on the earth where data are available.

IMPROVEMENT EFFORTS

Further improvement of a low resolution radiometer was attempted for the second generation weather satellite, Nimbus. The objectives were

to improve the radiometer performance by:

1. Reducing the excessive thermal coupling between the radiometer detectors and the satellite.
2. Improving the reflection coatings on the reflecting cones so that they would adequately reflect all radiation onto the detectors.
3. Improving the infrared absorptance and solar reflectance of the white detector. Also, since Nimbus will be earth oriented, the deterioration of the white coating from ultraviolet radiation exposure becomes more important than it was in the Tiros radiometer.

The initial effort was centered on reducing the coupling effects. The approach to the investigation of this problem was to employ a different method of mounting the detector at the terminal base of the cone, use different schemes for the electrical leads from the detector, and try to minimize the applied bias power. These investigations indicated that the degree of coupling could not be significantly improved over that attained with the Tiros radiometer.

A white coating (Naval Research Laboratory Type II) was found which did not show much deterioration from ultraviolet exposure and which had a solar-weighted reflectance of 75–80 percent. However, this paint was somewhat deficient in absorption in the infrared. The cone reflectance was improved by a better aluminizing technique.

With the present state-of-the-art limitations, further improvements of the deficiencies could not be made, since the radiometer was to be a low cost experiment. Therefore, development of a low resolution radiometer for Nimbus was discontinued.

CONCLUSION

The evaluation of the calibration data indicated that a strong thermal coupling existed between the radiometer detectors and the satellite and that the white detector coating and the cone optics were inadequate in spectral response. Improvement of these deficiencies would aid in reducing the probable errors of A and T_e , which were found to be approximately ± 8 to ± 14 percent and $\pm 24^\circ\text{K}$, respectively. With the present experiment, it is expected that improvement in the reflected solar radiation calibration could reduce the probable error in A to approximately ± 7 percent and in T_e to approximately $\pm 21^\circ\text{K}$.

For the Tiros low resolution radiometer, limitations on data coverage reduce the value of the data for heat balance studies. However, the relative measurements of T_e and A still should provide useful information for the areas on the earth where data are available.

Efforts to improve the performance of the low resolution radiometer were only partially successful because of limitations of the state-of-the-art. Development was discontinued because of the disproportionate cost in achieving an instrument which was to be of low cost.

ACKNOWLEDGMENTS

The authors wish to thank Dr. W. Nordberg and Mr. W. R. Bandeen for their critical review of this work. The assistance of Mr. R. Carter in performing the calibrations is also acknowledged.

REFERENCES

1. STROUD, W. G., "Initial Results of the Tiros I Meteorological Satellite," *J. Geophys. Res.* 65(5):1643–1644, May 1960.
2. STERNBERG, S., and STROUD, W. G., "Tiros I-Meteorological Satellite," *Astronautics* 5(6):32–34, 84–86, June 1960.
3. BANDEEN, W. R., HANEL, R. A., LICHT, J., STAMPFL, R. A., and STROUD, W. G., "Infrared and Reflected Solar Radiation Measurements from the Tiros II Meteorological Satellite," NASA Technical Note D-1096, November 1961; also *J. Geophys. Res.* 66(10):3169–3185, October 1961.
4. ASTHEIMER, R. W., DE WAARD, R., and JACKSON, E. A., "Infrared Radiometric Instruments on Tiros II," *J. Opt. Soc. Amer.* 51(12):1386–1393, December 1961.
5. HANEL, R. A., "A Low-Resolution Unchopped Radiometer for Satellites," NASA Technical Note D-485, 1961; also *Amer. Rocket Soc. J.* 31(2):246–250, February 1961.
6. JOHNSON, F. S., "The Solar Constant," *J. Meteorol.* 11(6):431–439, December 1954.
7. REID, C. D., and MCALISTER, E. D., "Measurement of Special Emissivity from 2μ to 15μ ," *J. Opt. Soc. Amer.* 49(1):78–82, January 1959.
8. DAVIS, J. F., HANEL, R. A., STAMPFL, R. A., STRANGE, M. G., and TOWNSEND, M. R., "Telemetry Infrared Data from the Tiros Meteorological Satellites," NASA Technical Note D-1293, August 1962.
9. NORDBERG, W., BANDEEN, W. R., CONRATH, B. J., KUNDE, V., and PERSANO, I., "Preliminary Results of Radiation Measurements from the Tiros III Meteorological Satellite," *J. Atmos. Sci.* 19(1):20–30, January 1962.
10. BIGNELL, K. J., "Heat-Balance Measurements from an Earth Satellite—an Analysis of some Possibilities," *Quart. J. Roy. Meteorol. Soc.* 87:231–244, April 1961.

ABSOLUTE CALIBRATION AND USE OF A SOFT X-RAY SOURCE OF CHARACTERISTIC CARBON RADIATION*

ANTHONY J. CARUSO AND WERNER M. NEUPERT

A measurement of the absolute photon yield of characteristic carbon K radiation produced by direct electron bombardment of a target is described. The carbon K emission is produced by a simple and economical source. The radiation is detected with a commercially available proportional counter, using both commercially available windows and thinner windows made at this laboratory. A method for determining the gas absorption efficiency of the counter and transmission of the counter window is described. The pulse height distributions obtained from the proportional counter vary in shape with target voltage indicating the presence of a bremsstrahlung continuum as well as characteristic K radiation. The continuum contribution is subtracted from the total counting rate to obtain the photon yield of characteristic K radiation as a function of target voltage. The results of this determination are used in the calibration of a grazing incidence grating spectrometer at 44A.

INTRODUCTION

The determination of the absolute photon yield of a soft x ray source requires a detector with a known quantum efficiency or one for which the quantum efficiency can be determined. Several types of detectors are available such as the flow type Geiger and proportional counters and the open end photomultipliers. The selection of the proportional counter is appropriate because the amplitude of its output pulses can be analyzed to obtain information on the spectral characteristics of a radiating source.

The purpose of this paper is to describe a method, utilizing a proportional counter, for determining the absolute photon yield of carbon K radiation from a simple and economical source, designed, built, and used at this laboratory for calibrating soft x ray instrumentation. The results of this experiment have been applied to the calibration of a grazing incidence spectrometer at 44A.

DESCRIPTION OF APPARATUS

A schematic diagram of the system employed in this experiment is shown in Figure 1. The x ray source has a simple configuration consisting of an elliptically shaped anode (disc) and v-shaped

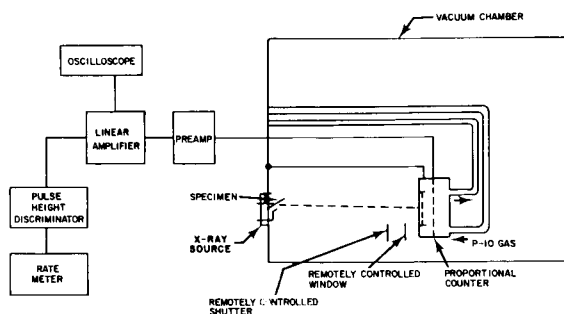


FIGURE 1.—Schematic diagram of apparatus used to determine the absolute photon yield of carbon K radiation (44A).

wire cathode. Figure 2 is a photograph of the x ray source showing the placement of the cathode and anode relative to the mounting plate. The anode, made of tantalum, has dimensions of $\frac{3}{4}$ " for the major axis, $\frac{9}{16}$ " for the minor axis and 2 mil in thickness. It is spot welded at its center to a nickel electrode and positioned at an angle of 45° with respect to its mounting plate. The v-shaped cathode (5-mil tungsten wire) with its vertex at the center of the anode, is $\frac{1}{4}$ " from the anode and parallel to it. This configuration results in an electron beam, approximately $\frac{1}{8}$ " in diameter directed normally to the anode surface of the x ray source. This source was attached to the main chamber with an o-ring seal and

*Published as *Goddard Space Flight Center Document X-614-64-23*, January 1964.

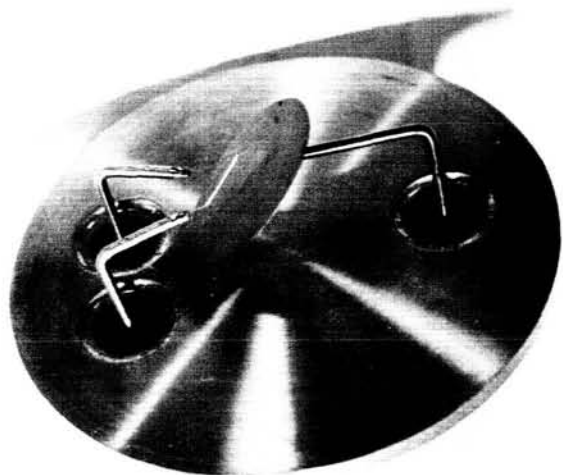


FIGURE 2.—X-ray source showing the mounting plate, anode and filament.

operated at beam currents as high as 1.0 ma at a chamber pressure of 3.0×10^{-6} mm Hg. It was found that the photon emission was stable to within 2.0 percent over long periods of operation and that the target required no cooling at these low currents. Carbon K band radiation, obtained by painting the target with aquadag, was detected within a solid angle of 3.67×10^{-5} steradians at a takeoff angle of 45° to the target face.

The detector* is a side window proportional counter 6.0 cm. long and 2.52 cm inside diameter with an anode wire diameter of 0.0466 mm. The commercially available counter window material supplied with the counter is mylar 6μ thick with one side covered with a thin layer of aluminum to provide conduction for electrical charges. An aluminum disc with an aperture of 0.252 mm wide and 16.16 mm high is placed over the counter window. P-10 gas (90 percent Argon and 10 percent Methane) flows through the counter at atmospheric pressure.

The counter is connected to a preamplifier which drives a linear amplifier having a rise time of about 0.2μ sec and a minimum input sensitivity of about 5 mv. The output of the linear amplifier is fed into a single channel pulse height analyser.

*This is a commercially available flow type counter made by the Siemens Co. in Germany and distributed in the U.S.A. by the Eastern Scientific Co., Cherry Hill, N.J.

The output of the analyser is connected into a ratemeter and scaler circuit.

EFFICIENCY OF COUNTER

The optimum gas pressure at which one can operate a counter is dependent on the strength of available window material, the counting stability as affected, for example, by the shortening of the plateau characteristic as the counter pressure is reduced, and the pressure at which the radiation is no longer completely absorbed in the gas path. Figure 3-F is a curve showing counting rate as a function of counter voltage, with the counter gas flowing at atmospheric pressure. This curve was obtained with an amplifier gain setting such that non x radiation background pulses were just below counting threshold. The proportional region is from about 1950 v to 2100 v and the Geiger region from about 2200 v to 2400 v. Beyond approximately 2430 v the counter goes into a continuous discharge. At atmospheric pressure the counter performs with excellent stability, having a good plateau in the Geiger region.

Other curves shown in Figure 3 illustrate the shortening of the plateau characteristic in the Geiger region as the counter pressure is reduced. At a counter pressure of 76 cm. of Hg one observes the plateau region extending from a counter voltage of approximately 2200 v to 2400 v while at a counter pressure of 40 cm. of Hg the plateau region is from approximately 1700 v to 1750 v.

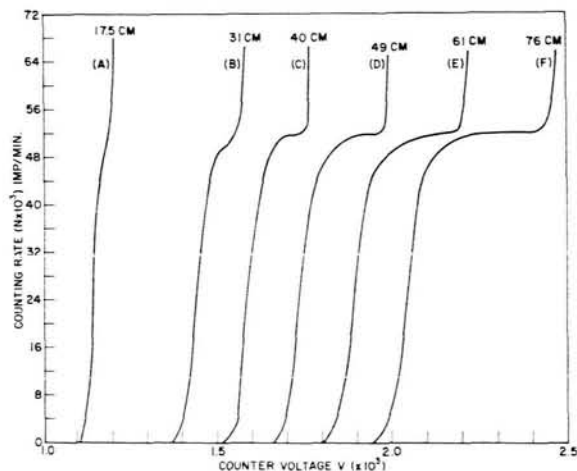


FIGURE 3.—Family of curves showing the decrease in width of the plateau region as the counter pressure is reduced at constant intensity of radiation.

At 17.5 cm. of Hg the counter goes from proportional operation almost directly into a continuous discharge. These data indicate that if one wishes to operate at low counter pressures a very stable high voltage power supply is required. This requirement is even more stringent if one wishes to operate the counter in the proportional region.

Since the primary object of the experiment is to measure the absolute photon yield of characteristic carbon *K* radiation it is necessary to know, among other factors, the gas absorption efficiency of the counter. This can be obtained from a study of the counting rate as a function of counter gas pressure. The results are shown in Figure 4.

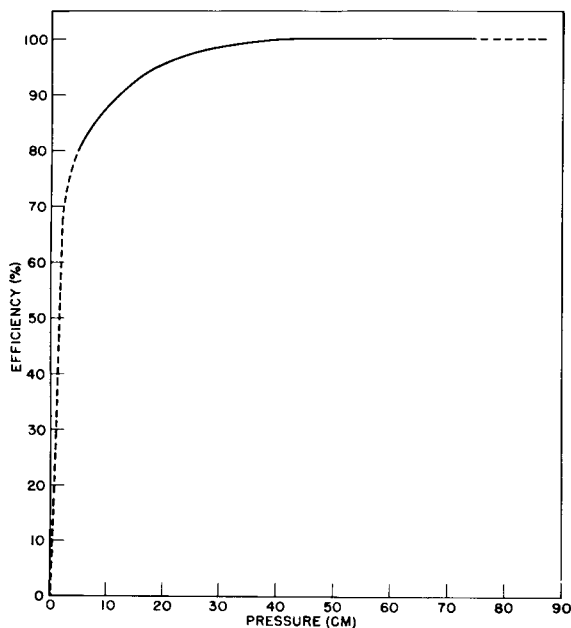


FIGURE 4.—Counter gas absorption efficiency as a function of counter pressure for constant intensity of radiation.

The fact that the counting rate does not increase for pressures above 40 cm. of Hg indicates that complete absorption of the radiation has taken place at the higher pressures. The absence of a decrease at higher pressures indicates the absence of a dead region in the counter.⁷ Below 40 cm. of Hg the counting rate does decrease, indicating that the gas path is not optically thick and hence, that the gas absorption efficiency of the counter is decreasing from 100 percent.

In order to identify the origin of soft x rays produced by the source, it was necessary to operate the counter in its proportional region thereby obtaining a degree of energy resolution. The extent to which proportional operation was achieved was checked by comparing the pulse height distributions of counts obtained from a carbon and from an aluminum target. Using a counter voltage of 2100 v, the electronics were set such that the peak for characteristic carbon *K* radiation occurred for 10 v pulses. Under these same electronic conditions, a peak was observed for characteristic aluminum *K* radiation at 57 v. The values obtained for carbon and aluminum were plotted against their respective *K* band x ray energies. These points lay on a straight line passing through the origin thus showing that the counter, at 2100 v, was operating in the proportional region.

Figure 5 shows the pulse height distributions of carbon *K* band radiation with constant current and various target voltages using the commercially available counter window material. The curves show a secondary maximum which increases rapidly as the target voltage is increased. This secondary peak has been attributed to bremsstrahlung radiation. The pulse height distribution of carbon *K* band photons occurs between 5 v and about 35 v while that due to continuum photons ranges from 35 v to beyond 105 v depending on the target voltage. In Figure 5 one observes that at a target voltage of 1000 v no observable continuum is present. As a target voltage is increased to 1500 v the continuum begins to be detected in the pulse voltage range from 40 to 80 v. At 1500 v the continuum radiation was 9.3 percent of the total count rate (carbon+continuum) whereas at 3000 v it was 44.0 percent of the total count rate. The percentages stated are valid only for the particular counter window which was used since the transmission of the window is the dominant factor. The results above clearly show that the continuum must be taken into consideration when determining the absolute photon yield of carbon *K* band radiation, especially when using thick counter windows.

The transmission of the counter window to be used for intensity measurements in the carbon *K* band wavelength region was directly measured

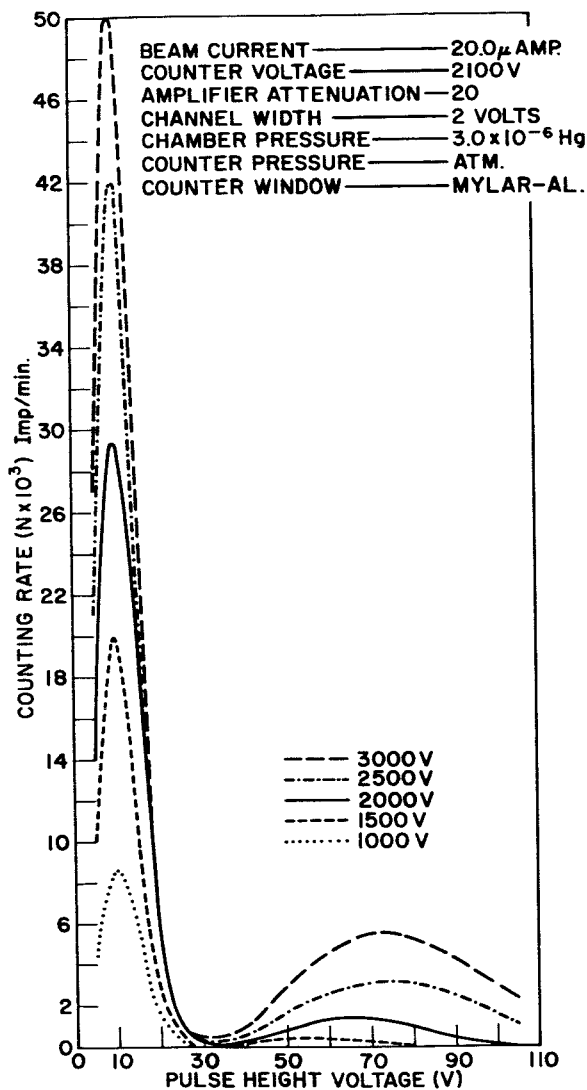


FIGURE 5.—Pulse height distribution curves for various target voltages with constant beam current using an aluminum-mylar window.

using a remotely controlled solenoid which placed the window between the x ray source and the counter having a window of unknown transmission. A solenoid controlled shutter could also be introduced in the x radiation path to check that the counter was counting only photons from the source. The transmission of the window was measured using only counter pulses produced by *K* band photons. This was done by first recording the total x ray counting ray (continuum and *K* band) with the window in and out of the radiation path. The same procedure was fol-

lowed, with the appropriate changes in electronic setting, to determine the counting rate due to the continuum alone. The difference between the total counting rate and that due to the continuum gave the counting rate of the carbon *K* band radiation. The ratio of the carbon *K* band counting rates with the window in and out of the x radiation path gave the transmission of the window. This procedure was followed for target voltages of 1.0, 1.5, 2.0, and 3.0 KV. The average transmission of the window at the carbon *K* band was calculated to be 2.43 percent ± 0.05 percent. This window was then placed in the counter and the intensity measurements were made.

INTENSITY MEASUREMENT

It was necessary to discriminate between counter pulses produced by characteristic *K* radiation and pulses produced by continuum photons in measuring the absolute *K* band photon yield of the source. The electronics were first set so that the counting rate due to the carbon *K* band plus continuum radiation was recorded. Then the electronics were set so that only the counting rate due to the continuum radiation was recorded. The difference between the total counting rate and that due to the continuum gave the counting rate of the carbon *K* band. This procedure was followed for target voltages in steps of 0.5, 1.0, 1.5, 2.0, 2.5 and 3.0 KV.

The absolute number of carbon *K* band quanta was obtained with a knowledge of the counter gas absorption efficiency, count rate, beam current, transmission of the counter window and the solid angle. Curve A in Figure 6 shows the absolute quanta per electron per unit solid angle as a function of target voltage for an aquadag target.

Another experiment was carried out utilizing spectroscopically pure carbon as the target material. The source configuration was the same as the one for aquadag previously described (the target in this case being a circular carbon disc one half inch in diameter and one-eighth inch thick.) The target was filed to smooth its surface to some extent. Absolute photon yield measurements were made under the same conditions as those in the experiment with aquadag. The results, shown in Figure 6, curve B, indicate that

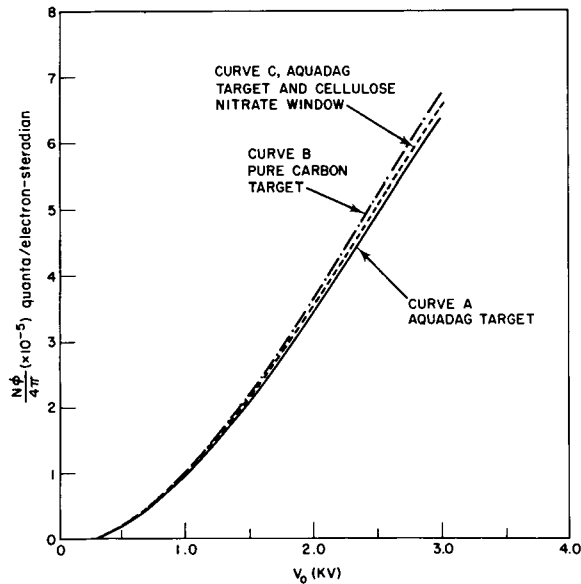


FIGURE 6.—The absolute photon yield of Carbon K band radiation as a function of target voltage V_0 .

the photon yield for pure carbon is slightly higher than for aquadag.

At a later date the photon yield measurement was repeated using a thin window (cellulose nitrate) made at this laboratory. Employing the same geometry and source of radiation, the transmission of the cellulose nitrate window in series with an electroformed mesh was determined to be 44.0 percent \pm 2.0 percent. This is an increase in transmission by a factor of 18 over the 6 μ thick mylar-aluminum window.

Figure 7 shows the pulse height distribution curves obtained using the cellulose nitrate window. It is evident from figures 5 and 7 that the counting rate is lower with a mylar-aluminum window and beam current of 20.0 μ a than with the cellulose nitrate window and a beam current of only 2.0 μ a. The shapes of the pulse height distribution curves in Figures 5 and 7 are similar except that the secondary maximum, attributed to the continuum, is more pronounced with the thicker counter window. The reason for this is that the thin cellulose nitrate window has practically the same transmission for the carbon K band as it has for the continuum radiation whereas the mylar-aluminum window has a much lower transmission for the carbon K photons than it has for the continuum.

Curve C in Figure 6 shows the absolute photon yields measured with the cellulose nitrate counter window. The results are within 5 percent of those photon yields obtained with the 6 μ thick

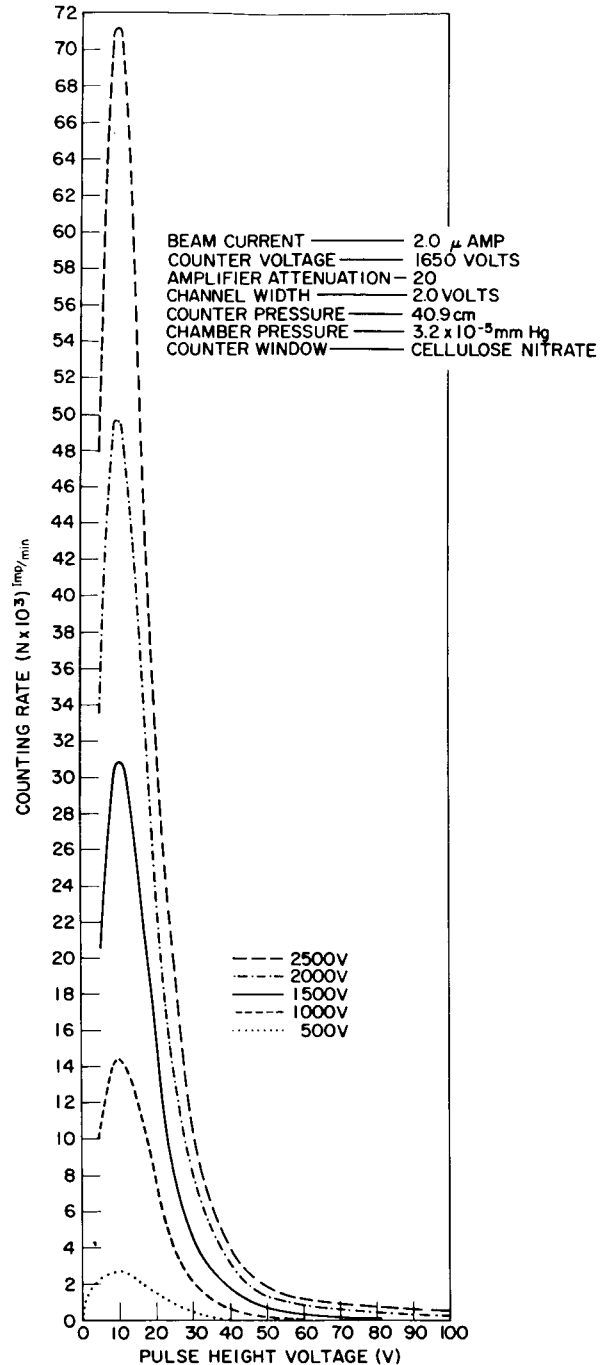


FIGURE 7.—Pulse height distribution curves for various target voltages with constant beam current using a cellulose nitrate window.

mylar-aluminum window. This check indicates that the measurements made with both window materials agree within experimental error.

DISCUSSION

The following corrections have been made in the observed photon yield measurements: Continuum radiation, counting rate in the low energy tail of the pulse height distribution and transmission of the counter window.

No corrections have been made for the spurious effects due to back scattering electrons because the target has a low back scattering coefficient and only a small proportion of back scattered electrons have energies sufficiently high for carbon K excitation.² No correction was made for surface condition of the target material. Errors due to analysis and apparatus have been calculated to be less than 10 percent. The results of this experiment may be compared with experimental work done by Dolby.² His absolute photon yield measurements are higher by a factor of about 1.7. For example, at a target voltage of 1.0 KV Dolby's measurement is 1.7×10^{-5} quanta per electron per unit solid angle, before a surface condition correction was made as compared to 0.98×10^{-5} obtained in this experiment. Brown and Ogilvie¹⁴ found Dolby's measurements for aluminum to be higher by a factor of 2. It is not clear why Dolby's measurements are higher for carbon or aluminum; however, Campbell¹³ discusses possible experimental errors.

CALIBRATION OF SCANNING SPECTROMETER AT 44A

The foregoing calibration technique has been applied in the calibration of a grazing incidence grating spectrometer designed to measure solar soft x ray and extreme ultraviolet radiation (Behring, Neupert and Nichols).¹¹ Radiation is incident on the grating at an angle of 88° . The diffracted rays continue to the exit slit in front of the detector, which is a Bendix Magnetic Electron Multiplier (M-306). This detector is mounted on a carriage which moves on a circular rail so that the exit slit follows along the Rowland circle where the spectrum is in focus. The plane of the exit slit stays approximately perpendicular to the diffracted ray at all positions along the rail, there-

by keeping the spectral passband nearly constant (to within one percent) for all angles of diffraction. Read out from the spectrometer was in analog-digital form, with the accumulation of each eight counts in the scaler circuit being reported as a change in voltage at the output of the electronics. Accumulation of multiples of eight counts (128 and 2048 counts) was also reported by appropriate algebraic combinations of voltages from two additional taps on the scaler chain.

The source described in the first portion of this paper was used for operational tests of the instrument, the distance from source to entrance slit being adjusted to simulate, as closely as possible, the actual appearance of the sun in the direction parallel to the Rowland Plane. This requirement was compromised to some extent to obtain increased fluxes at the entrance slit. A distance of about twelve inches from source to entrance slit was chosen, sufficient to guarantee that radiation from all points on the source, passing through the entrance slit, would be intercepted by some exposed portion of the ruled surface of the grating. During observation of this source with the spectrometer a separate Bendix M-306 electron multiplier monitored the constancy of the beam which, as has been pointed out, remained stable to within 2 percent over a period of several months.

The average of four scans through the carbon band is presented in Figure 8 together with mean

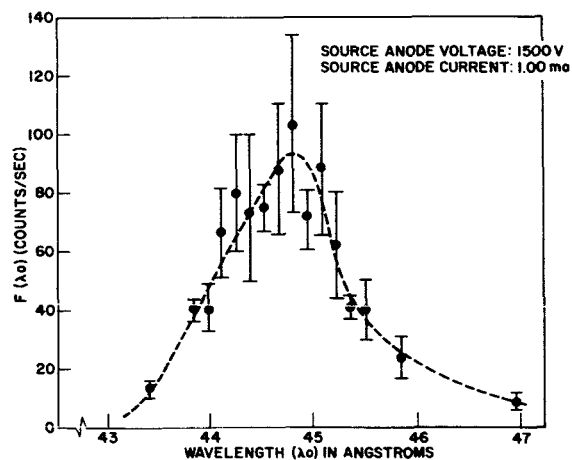


FIGURE 8.—First order carbon K emission band as observed with a grazing incidence spectrometer using an X-ray target voltage of 1500 V and a beam current of 1.00 ma.

deviations for each of the points which represent averages over uniform intervals of the spectrum. The large deviations are attributable to statistical fluctuations and also to error in measuring the time intervals over which the groups of eight counts were accumulated. The probable error of each point is of the order of 10 percent. A smooth curve (dashed line), compatible with the spectral passband, has been drawn through these points. This curve is also consistent with a previous observation of the carbon band made by Siegbahn.¹² The band was also observed in second order, appearing with reduced intensity, but still confirming the asymmetry which appears in first order. The entire nominal wavelength range of the spectrometer (10A–400A) was scanned and showed no regions of increased counting rates other than those associated with three orders of the carbon band.

A calibration for the wavelengths of the band may be obtained even though the exact shape of the band is not known. One need only assume that the efficiency of the instrument, $\epsilon(\lambda)$, defined as the probability of recording one count for one photon having a wavelength λ entering the entrance slit, when the spectrometer is set on wavelength λ_0 , be the same for all wavelengths in the band. In general, for a flux $I(\lambda)$ entering the instrument, one obtains, as an output from the spectrometer, a counting rate $F(\lambda_0)$ given by:

$$F(\lambda_0) = \int_{\lambda_1}^{\lambda_2} \epsilon(\lambda) I(\lambda) S(\lambda, \lambda_0) d\lambda$$

where $S(\lambda, \lambda_0)$, the instrumental transmission function, is defined as that fraction of the radiation of wavelength λ incident on the focal plane of the spectrometer which is admitted to the detector when the spectrometer is set on wavelength λ_0 . Neglecting aberrations of the system, the function $S(\lambda, \lambda_0)$ in the present case may be taken to be an isosceles triangle with a half width specified as $\Delta\lambda$. The limits of integration are chosen so as to include all of the spectral output of the source.

The problem of estimating ϵ is made more manageable, especially in the present case, by integrating $F(\lambda_0)$ over λ_0 :

$$F' = \int_{\lambda_1 - \Delta\lambda}^{\lambda_2 + \Delta\lambda} F(\lambda_0) d\lambda_0 = \int_{\lambda_1 - \Delta\lambda}^{\lambda_2 + \Delta\lambda} \int_{\lambda_1}^{\lambda_2} \epsilon(\lambda) I(\lambda) S(\lambda, \lambda_0) d\lambda d\lambda_0$$

where F' is now the area under the output curve of the spectrometer (Fig. 8) with units of $A \cdot \text{counts/sec}$. The limits chosen for this integration are appropriate for the symmetric transmission function just described. If the source of radiation is monochromatic, or nearly so, one may assume that the efficiency ϵ is constant over the range of integration:

$$F' = \epsilon \int_{\lambda_1 - \Delta\lambda}^{\lambda_2 + \Delta\lambda} \int_{\lambda_1}^{\lambda_2} I(\lambda) S(\lambda, \lambda_0) d\lambda d\lambda_0$$

The geometry of the spectrometer being calibrated is such that the double integral may be reduced to two measurable quantities, thereby making possible an estimate of ϵ . We must make use of the fact that, for a scanning spectrometer in which the plane of the exit slit remains always perpendicular to the diffracted ray, the integral $\int_{\lambda} S(\lambda, \lambda_0) d\lambda$ is independent of λ_0 , the wavelength setting of the spectrometer. Indeed, we have the further result

$$\int_{\lambda} S(\lambda, \lambda_0) d\lambda = \int_{\lambda_0} S(\lambda, \lambda_0) d\lambda_0 = \text{constant} = \Delta\lambda$$

where $\Delta\lambda$ is, as before, the half width of the instrumental transmission function. This half width is constant, to within 1 percent, over the entire range (10A to 400A) of the present spectrometer.

One may change the order of integration of the double integral, the mathematical requirements for this operation being obviously satisfied, to obtain:

$$\begin{aligned} F' &= \epsilon \int_{\lambda_1}^{\lambda_2} I(\lambda) \int_{\lambda_1 - \Delta\lambda}^{\lambda_2 + \Delta\lambda} S(\lambda, \lambda_0) d\lambda_0 d\lambda \\ &= \epsilon \int_{\lambda_1}^{\lambda_2} I(\lambda) d\lambda \int_{\lambda_1 - \Delta\lambda}^{\lambda_2 + \Delta\lambda} S(\lambda, \lambda_0) d\lambda_0 \end{aligned}$$

Hence

$$\begin{aligned} \epsilon &= \frac{F'}{\int_{\lambda_1}^{\lambda_2} I(\lambda) d\lambda \int_{\lambda_1 - \Delta\lambda}^{\lambda_2 + \Delta\lambda} S(\lambda, \lambda_0) d\lambda_0} \\ &= \frac{F'}{\Delta\lambda \int_{\lambda_1}^{\lambda_2} I(\lambda) d\lambda} \end{aligned}$$

The integral over the input flux in the denominator must be obtained by some means independent

of the spectrometer. The proportional counter work supplied this measurement.

The result which has been obtained using a band of radiation is applicable to monochromatic inputs as well if one uses the peak output counting rate instead of the area under the output curve. For a monochromatic input $I(\lambda_1)$, one has

$$F_{\max}(\lambda_0) = \epsilon(\lambda_1)I(\lambda_1)S_{\max}(\lambda, \lambda_0).$$

Both $F(\lambda_0)$ and $S(\lambda, \lambda_0)$ are maximized by taking $\lambda_0 = \lambda_1$. For the spectrometer being calibrated (exit slit equal in width and length to the entrance slit)

$$S_{\max}(\lambda, \lambda_0) = 1,$$

so that we obtain

$$\epsilon(\lambda_1) = \frac{F(\lambda_1)}{I(\lambda_1)}.$$

In applying this calibration procedure to the calibration of the spectrometer we take

$$I(\lambda) = \omega\varphi(\lambda)$$

where $\varphi(\lambda)$ is the flux from the source, per unit solid angle, and ω is the effective solid angle subtended by the spectrometer. The limiting apertures used in defining this angle were the width of the entrance slit and the length of the exit slit. Taking W and L as the width and height of the respective slits, and r_1 and r_2 their respective distances from the source, we obtain

$$I(\lambda) = \frac{WL}{r_1 r_2} \varphi(\lambda)$$

$$\epsilon = \frac{F}{\omega \Delta \lambda \int \varphi(\lambda) d\lambda} = \frac{r_1 r_2 F}{WL \Delta \lambda \int \varphi(\lambda) d\lambda}$$

In examining a distant source such as the sun, one wishes to compute the average number of incident photons, per square centimeter, which will result in one output pulse from the spectrometer. This quantity, which we shall denote by S , is then given by

$$S = \frac{1}{\epsilon WL} = \Delta \lambda \frac{\int \varphi(\lambda) d\lambda}{r_1 r_2 F}.$$

During calibration of the spectrometer the target voltage of the source was 1500 V and the current 1.00 ma. yielding $\int \varphi(\lambda) d\lambda = 12.5 \times 10^{10}$ photons/sec per steradian. Under these circumstances we obtain

$$F = 147 \pm 15 \text{ A} \cdot \text{counts/sec.}$$

Using $r_1 = 31.8$ cm, $r_2 = 43.2$ cm, and $\Delta \lambda = 0.85 \text{ \AA}$, we obtain:

$$S = 5.2 \times 10^5 \text{ photon cm}^{-2} \text{ count}^{-1}$$

in the wavelength range from 43A to 46A. The probable error of this result is estimated to be 15 percent.

CONCLUSION

This investigation of a commercially available proportional counter shows that it can be used as a detector for the absolute measurement of quanta in the soft x-ray region. Intensity measurements made with the mylar-aluminum and cellulose nitrate windows show that they agree well within experimental error. The counter can be used at wavelengths longer than 44A by making use of suitably thin windows having high transmission. Using X-ray sources having targets made of light elements, it is possible to obtain a calibration for a spectrometer, such as the one described here, at a sequence of discrete wavelengths in the soft x-ray region.

ACKNOWLEDGMENT

The authors wish to thank Mr. Louis Kirouac and Mr. Walter Gates for their design and building of the X-ray source and Dr. Henry Horstman for valuable discussions.

REFERENCES

1. KORF, S. A., "Electron and Nuclear Counters," D. Van Nostrand Company, 2nd Edition, 1957.
2. DOLBY, R. M., "Absolute Intensity Measurements of Carbon and Aluminum X-Ray K lines with a Proportional Counter," Brit. J. Appl. Phys., Vol. 11, Feb. 1960.
3. ROSE, M. E., and KORF, S. A., "An Investigation of the Properties of Proportional Counters," Phys. Rev., Vol. 59, June 1, 1949.

4. HOLLIDAY, J. E., "A Soft X-Ray Spectrometer using a Flow Proportional Counter," *The Review of Scientific Instruments*, Vol. 31; No. 8, August 1960.
5. REIDY, W., "A Study of the Dead Time and Resolving Time in a Proportional Counter," Department of Physics, College of Engineering, New York University, University Heights, New York 53, N.Y. AFOSR Document Number-AFOSR-TN-59-501, AD Document No. AD-215-926.
6. CULLITY, B. D., "Elements of X-Ray Diffraction," Addison-Wesley Publishing Company, Reading, Mass. 1956.
7. LUKIRSHIL, A. P., RUMSH, M. A., and SMIRNOU, L. S., X-Ray Monochromator for Ultrasoft Radiation with the Recording of the Absolute Number of Quanta," *Optics and Spectroscopy*, Vol. IX, No. 4, October 1960.
8. WORTHINGTON, C. R., and TOMLIN, S. G., "The Intensity of Emission of Characteristic X-Radiation," *Proc. Phys. Soc.* 69, pg. 401 (1956).
9. ARCHARD, G. D., "Proceedings of the Second Symposium on X-Ray Microscopy and X-Ray Microanalysis."
10. COMPTON and ALLISON, "X-Rays in Theory and Experiment," D. Van Nostrand Company, 2nd Edition, Princeton, New Jersey.
11. BEHRING, W. E., NEUPERT, W. M., and NICHOLS, W. A., "Far Ultraviolet Spectrometer for Solar Observations," *J. Opt. Soc. Amer.* 52(S), pg. 597, May 1962 (Abstract).
12. SIEGBAHN, M., "Studies in the Extreme Ultraviolet and the Very Soft X-Ray Region," *Proc. Phys. Soc.* Vol. 45, pg. 689 (1933).
13. CAMPBELL, A. J., "K X-Ray yields from Elements of Low Atomic Number; Research Lab. Report No. A1289," Associated Electrical Industries, Ltd. Aldermaston, Berks. July 1962.
14. BROWN, D. B., and OGILVIE, "Efficiency of Production of Characteristic X-Radiation from Pure Elements Bombarded by Electrons," *J. of Applied Phy.*, Vol. 35, No. 2, February 1964.

A MULTIPLE ELEMENT SOFT X-RAY SOURCE*

A. J. CARUSO AND W. M. NEUPERT

A multiple element soft x-ray source has been developed to study the soft x-ray region of the electro-magnetic spectrum. The design of the source is such as to allow one to obtain radiation from six different elements without breaking vacuum.

Figure 1 shows a cross-sectional diagram and top view of the soft x-ray source. The entire target assembly may be rotated, after removing the high voltage lead (A) and target positioning pin (B), to any one of the six target elements under high vacuum. The targets (G) are rectangles $\frac{5}{8}$ by $\frac{7}{8}$ inches of pure stock elements or vacuum deposited elements on substrates which are easily removed and replaced by target mount screws (C). The present target mount (H) has been designed for photon detection at 45 degrees with respect to the target surface. However, target mounts can be designed for photon detection from 10 to 90 degrees with a modification of the filament position.

The V-shaped filament, made of 5 mil. tungsten wire, is placed in the filament holder (E) and secured into position by knurled nuts so that the V portion is approximately $\frac{1}{4}$ inch from the target surface and parallel to it. This configuration does not interfere with the rotation of the target assembly and produces an electron beam normal to the target surface with a diameter of approximately $\frac{1}{8}$ inch. The length of the source from (A) to (F) is $2\frac{3}{4}$ inches and the target assembly, including targets, has a diameter of 2.0 inches. If space is available a target mount of greater diameter can be designed to accommodate more targets.

The test and evaluation was carried out using a flow type proportional counter with a stretched polypropylene counter window having a thickness of approximately 5000 Å. The geometry of the

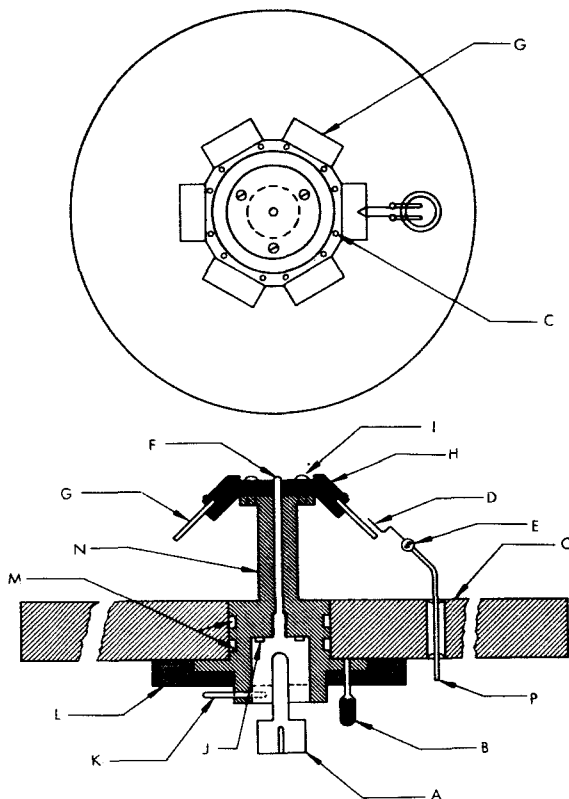


FIGURE 1.—Cross-sectional diagram of multiple element soft x-ray source: (a) phone-type jack high voltage connection, (b) target positioning pin, (c) target mounting screws, (d) filament, (e) filament holder, (f) anode conductor, (g) targets, (h) target mount, (i) target mount screws to insulator, (j) O-ring, (k) target rotating pin, (l) insulator retainer, (m) O-ring, (n) insulator, (o) flange, (p) filament connection.

test apparatus is such that radiation is detected within a solid angle of 2.34×10^{-5} steradians.

The source was attached to the vacuum test chamber using the O-ring seal arrangement shown in Figure 1. It was operated at beam currents as high as 1.0 ma. and at pressures as low as 8.0×10^{-7} mm Hg. The intensity was linear with

*Published as *Goddard Space Flight Center Document X-614-64-344*, November 1964.

respect to beam current to within 1.0 percent and photon emission for a given beam current was stable to within 2.0 percent over long periods of

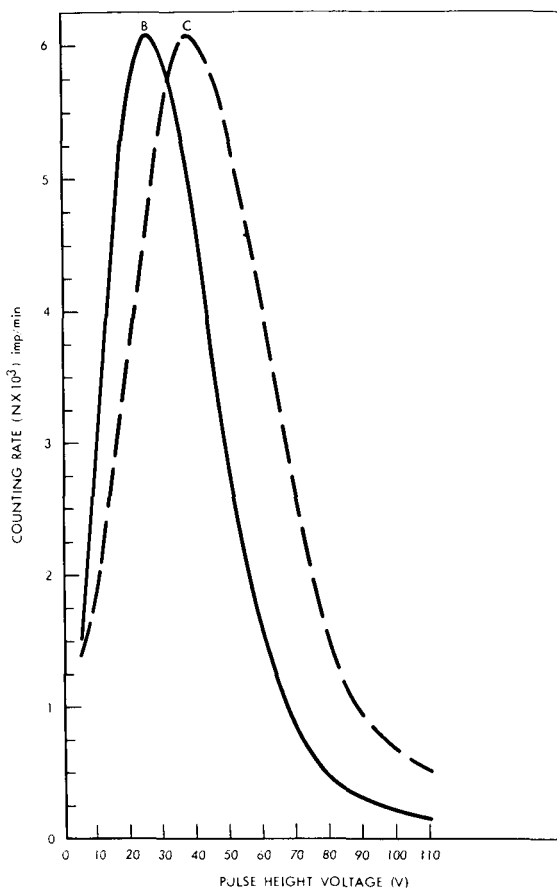


FIGURE 2.—Pulse height distribution curves: (b) Boron K and (c) carbon K radiation

operation. The target required no cooling at these low beam currents. Pulse height distributions of carbon K and boron K radiation are shown in Figure. 2.

The absolute photon yield of carbon K and boron K radiation was determined using this source. The results are shown in Table 1. The

TABLE 1.—X-Ray Photons Yields in Units of $10^4 N \phi$
($N \phi$ = photons per electron per 4π steradian)

Accelerating Voltage (KV)	Caruso & Neupert		A. J. Campbell	
	C	B	C	B
0.5	0.21	0.12	0.21	0.24
1.0	1.35	1.22	1.23	1.42
1.5	2.60	2.70	2.20	3.25
2.0	4.20	4.30	4.20	4.50
2.5	5.65	5.75		
3.0	7.55	7.24	7.60	7.5
3.5	8.70	8.60		
4.0			9.20	9.95

estimated error for both absolute measurements are ± 10.0 percent. The table compares our measurements with Campbell's.¹

The authors wish to thank Mr. Walter Gates for suggestions in the design of this source.

1. A. J. CAMPBELL, Proc. Roy. Soc. (London) 274, 319, 1963.

ULTRAVIOLET PHOTOGRAPHY AND SPECTROSCOPY USING A SPECTRALLY SELECTIVE IMAGE CONVERTER*

L. DUNKELMAN† AND J. P. HENNES

Laboratory and space research have both benefited from the new spectrally selective ultraviolet photodetectors. Large band-gap photoemitters, such as cesium telluride, having high quantum efficiencies in the ultraviolet and very low yields at longer wavelengths are now available. An image tube with such a cathode combines the advantages of spectrally selective photoemission with the flux-integrating property of photographic emulsion. Observations can thus be made at selected ultraviolet wavelength regions, either photographically or visually, even in the presence of strong long wavelength backgrounds. Visual measurements with a Cs-Te image converter, having a sapphire faceplate and a green phosphor, were made with the following results: (a) cathode quantum efficiency at 2537 Å is 0.016, which is expected to improve in future tubes; (b) minimal detectable converted 2537 Å irradiance is 4×10^{-5} ergs $\text{cm}^{-2} \text{sec}^{-1}$; and (c) tube background is undetectable. Mercury spectra obtained with this tube demonstrate the expected high rejection of visible light. Tables are included which list the brightnesses of several extended ultraviolet and visible sources of interest in the space sciences. The luminance of visible sources; the radiance of ultraviolet sources; and the relative brightness, on a photon basis, of spectrally converted ultraviolet sources along with visible sources are tabulated.

There is a continuing need in spectroscopy and photography, in both laboratory and space research, for spectrally selective ultraviolet photodetection techniques. Photomultiplier detectors containing large band-gap photoemitting cathodes now exist which effectively discriminate against long wavelengths.^{1,2} An image tube with such a cathode can combine the advantages of spectrally selective photoemission with either the flux-integrating property of photographic emulsion or the synthesizing ability of a human observer. This paper includes a brief discussion of some of the considerations to be given to the choice of window and photocathode materials, extent of image intensification, etc. in converter design. Typical examples are given of the kinds of extended UV sources of astrophysical or geophysical origin which an astronaut might observe and photograph in various ultraviolet regions.

IMAGE TUBE CONSIDERATIONS

McGee³ and Morton⁴ have recently provided excellent surveys of image tube development; accordingly, here we will include discussions of only those image tube parameters unique to ultraviolet detection.

Figure 1 shows spectral response curves of some of the photocathodes used for ultraviolet photo-detection. These emissive materials have been used in phototubes with appropriate ultraviolet transmitting faceplates such as LiF, CaF₂, sapphire, fused silica, etc. The conventional Cs-Sb

photoemitter is sensitive to visible light and is frequently not useful as an ultraviolet detector because of the presence of long wavelength

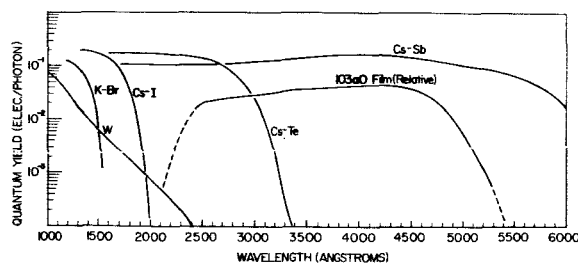


FIGURE 1.—Spectral response of typical UV photocathodes. These curves are taken from reference 1. The relative response of type 103a-0 film is from reference 5.

*Published as *Goddard Space Flight Center Document X-613-64-222*, September, 1964.

†Goddard Space Flight Center and Institute for Defense Analyses, Arlington, Virginia.

straylight. The spectrally selective cathodes Cs-Te and Cs-I are both relatively insensitive to longer wavelengths and make excellent UV detectors. Metal photocathodes, such as tungsten, are insensitive in the visible and have low yield in the middle ultraviolet (3000-2000A), but reach about 10 percent quantum yield at 1000A. Also shown is the relative quantum yield of Eastman 103a-0 spectroscopic film adapted from the Kodak handbook.⁵ The response curves show that if a weak ultraviolet source must be observed in the presence of a strong visible background, then one must employ either spectral filtering or one of the spectrally selective cathodes, and occasionally both.

Figure 2 shows a variety of available ultraviolet transmitting windows whose short wavelength cutoffs range from $\lambda\lambda 1050A$ (lithium fluoride) to $\lambda\lambda 3000A$ (Corning 0-54, glass). In general, these windows may be employed as the actual image tube faceplate. However, in practice it is usually easier to employ as the faceplate either an alkali halide, sapphire, fused silica or vycor material and provide for additional short wavelength cutoffs by using either the appropriate material shown in Figure 2 or some other crystal, glass, gas cell, predisperser, etc. as desired by the user.

In this paper, we report on an ITT ultraviolet image tube with a cesium-telluride cathode and a sapphire faceplate. However, in Figure 3 there is shown a few of many other combinations which

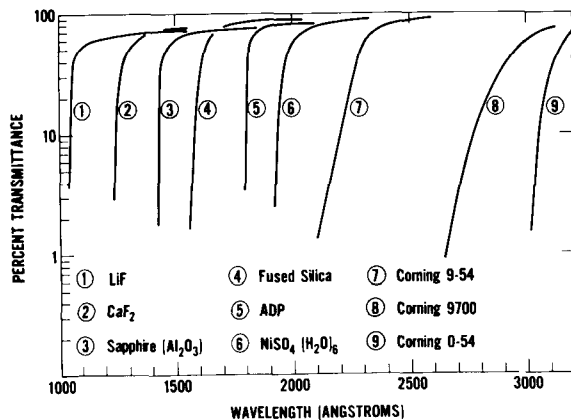


FIGURE 2.—UV optical materials. Transmittance curves, uncorrected for surface reflections, are shown for materials of thickness: 1-1.2 mm; 2-1 mm; 3-1 mm; 4-1 mm; 5-2.5 mm; 6-3 mm; 7, 8, and 9 are each about 2 mm.

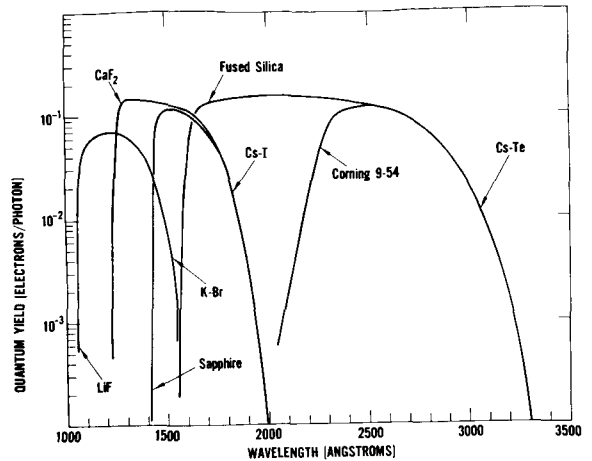


FIGURE 3.—Spectrally selective UV photodetectors. These curves are produced by combining some of the cathodes and filters from Figures 1 and 2.

are practicable, the spectral sensitivity curves being formed by combining the cathodes of Figure 1 with the windows of Figure 2. Note the relatively narrow spectral response of Cs-I with sapphire window; this region of $\lambda\lambda 1400$ to $2000A$ might be of interest in the study of the Schumann-Runge O_2 bands in the upper atmosphere. Note also the K-Br cathode with LiF window yielding a response from $\lambda\lambda 1050$ to $1500A$ for observing, for example, stellar or nebular fluxes. In many cases one would surely wish to exclude the strong

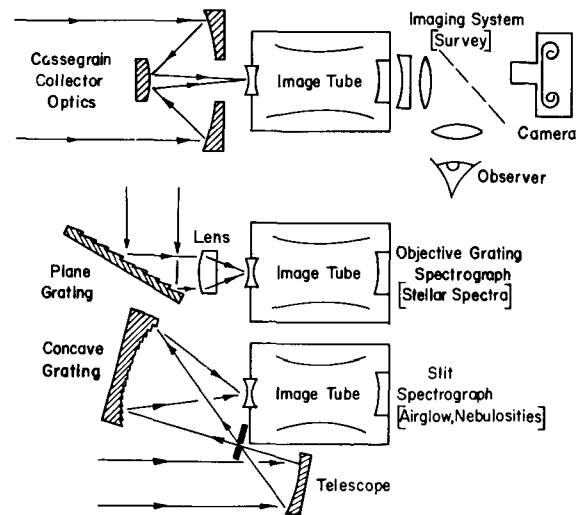


FIGURE 4.—UV imaging systems. Various possible optical systems for use with an image converter are shown. Either visual or photographic observation are represented in the top picture.

Lyman α hydrogen emission in the upper atmosphere; here one could employ a CaF_2 window or filter.

Figure 4 indicates three of many types of imaging systems. The top sketch, for example, shows a system which might be used by an astronaut for making both photographic and visual surveys of geophysical or astronomical objects such as air-

energy distribution of airglow, nebulosities and aurorae. Since so little is known of the dynamic properties of ultraviolet aurorae, a movie camera might be substituted for the still camera depending on whether one were interested in maximizing the observations of spatial or spectral distributions. Additional discussion on the expected radiances of celestial objects is given in a later section.

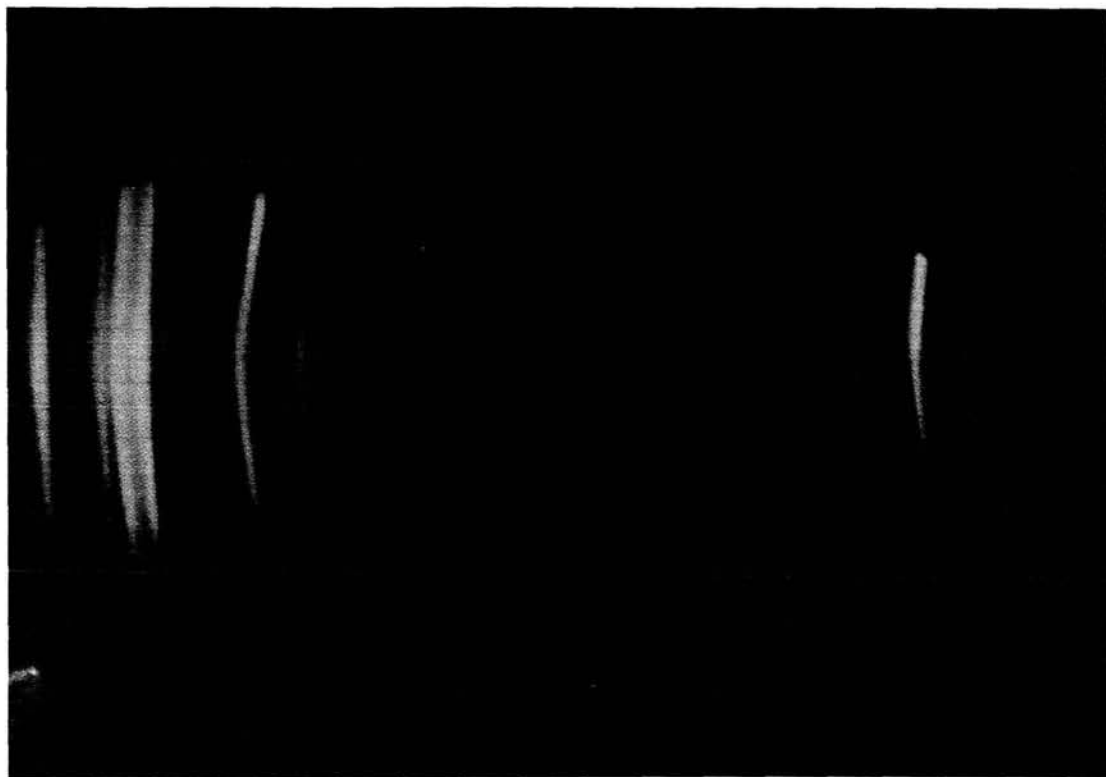


FIGURE 5A.—The spectrum of the Hg lamp is shown on both tubes. The extreme left line is 3650Å, the extreme right line is the green 5461Å. Next to the first order green line is the second order 2537Å line. Between the 2537Å and the first order 3650Å line are various blue and ultraviolet lines. Note that with the UV tube only the 2537Å line and a very faint 3650Å trace are recorded. The visible and near UV lines do not appear at all.

glow, nebulosities, stellar and planetary sources, earth albedo, etc.⁶⁻¹¹ In this case, the image tube acts as a spectral filter and a wavelength converter. Conventional photographic film would be used rather than fragile ultraviolet sensitive spectroscopic films. The middle sketch demonstrates an objective grating spectrograph to observe and obtain, for example, stellar or planetary spectra; whereas the bottom sketch would be a useful scheme for observing a spectral

The spectral selectivity of an ultraviolet image tube having a cesium telluride photocathode, is demonstrated by the photographs in Figure 5. A tube manufactured by the ITT Industrial Laboratories in Fort Wayne, Indiana has been compared with a standard 7404 image tube with an S-13 cathode. Both are electrostatic, monovoltage tubes with identical imaging structures. The tubes were used to photograph a mercury arc spectrum in an arrangement similar to that illus-

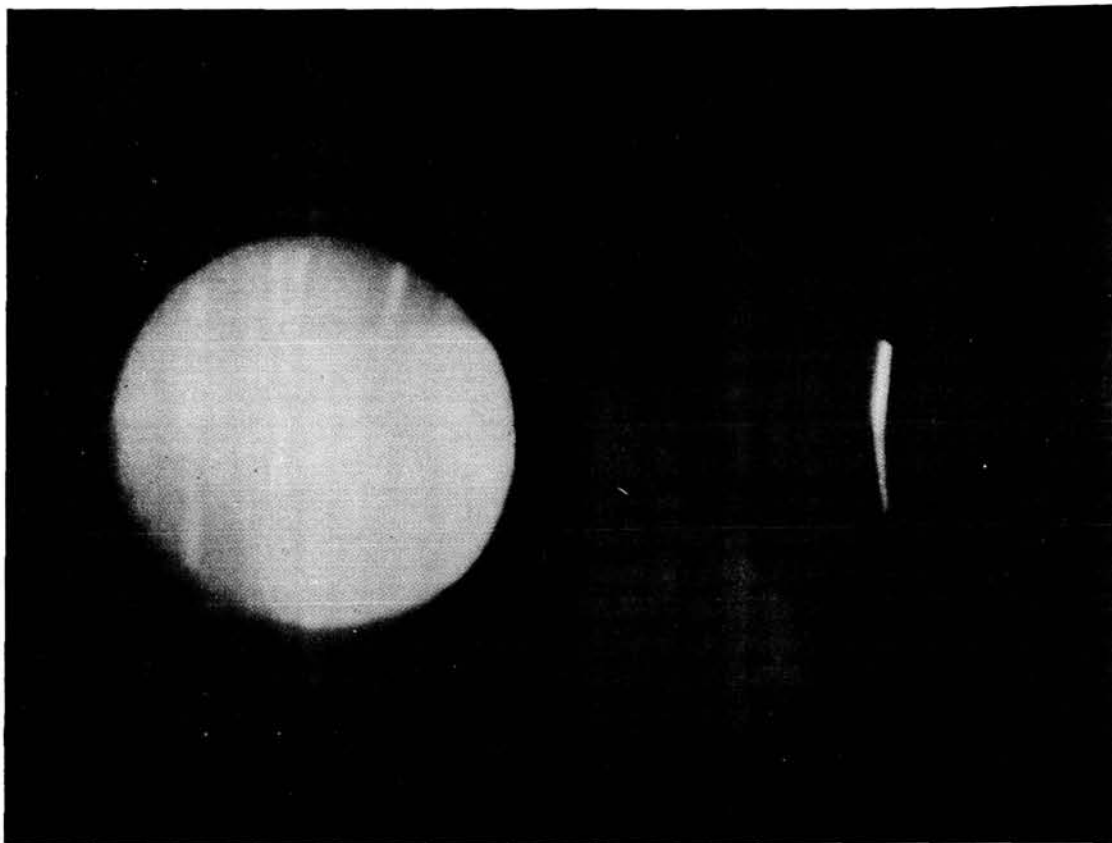


FIGURE 5B.—The spectrum of a Hg lamp in the presence of a strong continuum background (tungsten lamp) are recorded on both tubes. The visible image converter spectrum is almost washed out while the UV tube spectrum is the same as in part A.

trated in Figure 4 (bottom). The photographs show that although both tubes are equally sensitive to ultraviolet light, the Cs-Te ultraviolet tube is not affected at all by visible lines or stray light.

Figure 6 is a photograph of a display of electro-optical binoculars and the image tubes described in this paper. The upper right hand corner shows a pair of infrared binoculars which use standard infrared image tubes with cesium-silver-oxygen photocathodes. The image tubes shown are from left to right an RCA 7404 (Cs-Sb cathode with vycor window); an ITT UV tube (Cs-Te cathode with sapphire window); and an RCA far-UV tube (Cs-I cathode with LiF window). The image tube binocular, shown on the left with oculars and rubber shields removed, is equipped with quartz-CaF₂ doublet objective lenses with a 10.8 cm focal length shown in the detached barrels. The

image tubes, housed in the central cylinders of the main body of the binocular, are powered by a 12 kilovolt miniature power supply. Monoculars rather than binoculars would be used in manned space observations since extra-terrestrial observations would not appreciably benefit by double vision, certainly stereopsis is not involved here. Future developments of ultraviolet image tubes can include better collection optics and more sophisticated features such as multi-stage image intensification, magnetic focussing, simultaneous electronic readout, color translation, etc. One or more additional stages of image intensification (over the usual 30 or 40 gain in these single stage tubes) will be necessary in the cases of some of the interesting dim phenomena discussed in the next section.

The quantum efficiency of the photocathode of the Cs-Te tube reported here is 0.016. In 1958,

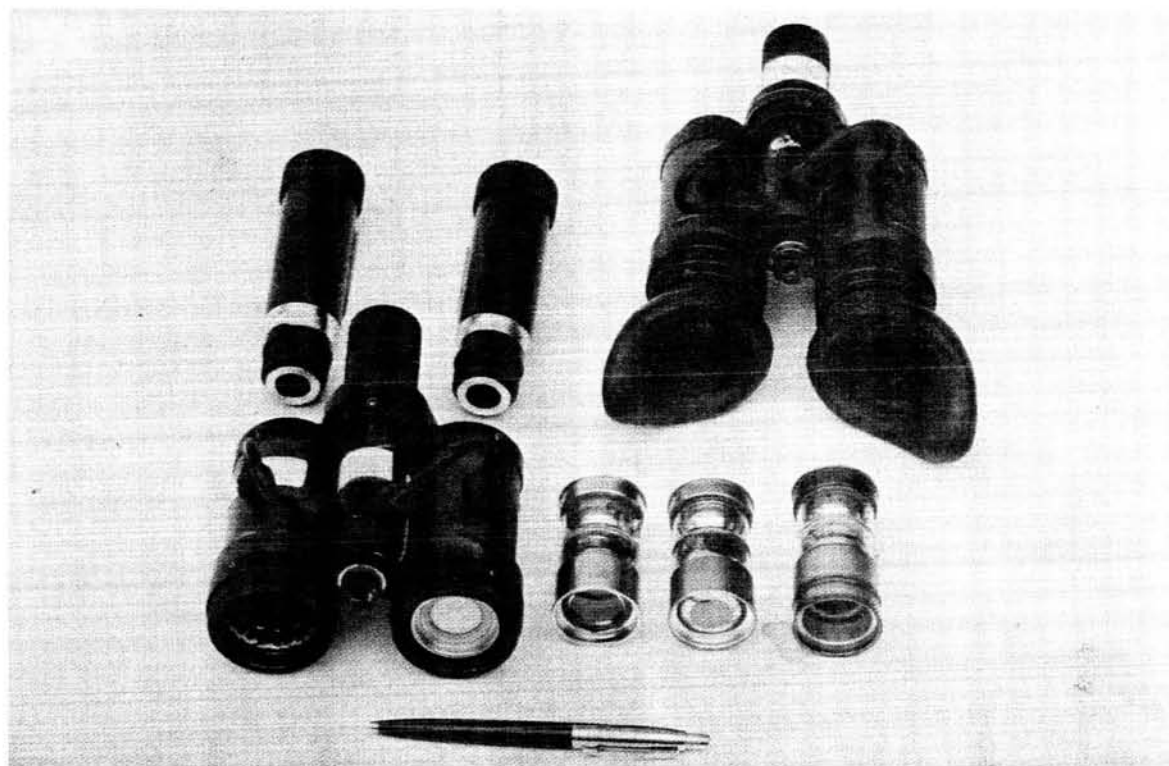


FIGURE 6.—UV image tubes and binoculars. (See text.)

when the first rubidium or cesium telluride photomultipliers were assembled the quantum efficiencies were even lower. Currently produced UV photomultipliers have yields of 0.05 or better and it is expected that this will be achieved with the image converters. Using our present tube, which has a green phosphor, as a converter we found that the minimal 2537A irradiance detectable by the eye was 4×10^{-5} ergs $\text{cm}^{-2} \text{sec}^{-1}$. With the f/5 objective the minimal detectable 2537A radiance of an extended source was 0.01 erg $\text{cm}^{-2} \text{sec}^{-1} \text{ster}^{-1}$. With this tube, the non-irradiated portion of the phosphor (the background) was black. This is what would be expected from large band-gap photoemitters which have negligible thermionic emission.

THE ULTRAVIOLET SCENE FROM SPACE

With reference to Table I, one can note the large variety and intensity range of extended (rather than point) source celestial objects of interest. These can be observed only from a vantage point above the ozonosphere (which absorbs the entire middle ultraviolet $\lambda 2000-$

3000A); or better still, at or above manned space altitudes (160 Km or more) to minimize absorption by oxygen. The table is divided into 3 sections, the first relates to phenomena in the visible region and is intended to provide (1) a convenient reference to objects fairly well known and whose observations are within our normal

TABLE I

Phenomenon	Luminance (candles cm^{-2})
Milky way, dimmest region, near Perseus	1×10^{-8}
Gegenschein.....	1.6×10^{-8}
Visible night glow (zenith).....	2×10^{-8}
Milky way brightest region, near Carina	4×10^{-8}
Zodiacal light, (30° elongation)....	1.2×10^{-7}
Visible night glow (edge-on).....	6×10^{-7}
Great Orion nebula M42.....	5.6×10^{-6}
Full Moon.....	4×10^{-1}
Fluorescent lamp 4500 white.....	4×10^{-1}

TABLE I (continued)

Phenomenon	Radiance (ergs sec ⁻¹ cm ⁻² ster ⁻¹)
Middle UV night glow (nadir), 2600A	$2 \times 10^{-5}/100A$
Orion nebulosity (mean), 1225- 1350A	$6 \times 10^{-5}/100A$
Orion nebulosity (most intense part), 1225-1350A	$4 \times 10^{-4}/100A$
Middle UV nightglow (edge-on), 2600A	$5 \times 10^{-4}/100A$
Hydrogen lyman-alpha glow (mean), 1216A	3×10^{-3}
UV Aurora, Wallops Is. 22 Nov. 1960 (zenith), 1700-1800A	10^{-2}
UV Aurora, Wallops Is. 22 Nov. 1960 (zenith), 1300-1800A	10^{-1}
Sunlit earth albedo (nadir), 2600A	$5 \times 10^{-1}/100A$
Phenomenon	Relative Brightness (Photon flux)
Middle UV nightglow (nadir), 2600A	0.03
Orion UV nebulosity (bright por- tion), 1225-1350A	0.3
Milky way (mean).....	0.7
Gegenschein.....	0.8
Middle UV nightglow (edge-on), 2600A	0.8
Visible nightglow (zenith).....	1
Hydrogen lyman-alpha glow (mean), 1216A	2
Aurora IBC-I.....	3
Zodiacal light (30° elongation)....	6
Visible nightglow (edge-on).....	30
UV Aurora, Wallops Is. 22 Nov. 1960 (zenith), 1300-1800A	100
Great Orion nebula M42.....	300
Sunlit earth UV albedo (nadir), 2600A	800
Aurora IBC-IV.....	3000
Full moon.....	2×10^7

experience and (2) a means with which one can compare ultraviolet with visible sources. In the first section the luminance (brightness) of each source is given in cd cm⁻². In the second section of the table, there is listed again a wide variety of celestial UV sources all of which have been meas-

ured from above the ozonosphere. Here the radiance (brightness) is given in ergs cm⁻² sec⁻¹ ster⁻¹. Finally, in the third section, some of these sources are listed relatively on a brightness scale. The ultraviolet fluxes have been converted to the corresponding 5500A fluxes on a one-for-one photon basis.

With fast collection optics and higher cathode efficiencies, ultraviolet sources of radiance as low as 10^{-4} ergs cm⁻² sec⁻¹ ster⁻¹ should be visually detectable with only single stage image converters. This radiance level includes most of the sources listed in Table I. Less bright sources will require further amplification stages to be detectable. High amplification of very dim sources will not necessarily increase the information content for visual detection since photon statistical fluctuations at low flux levels limit image recognition.⁴ Intensification of moderate brightness sources will, of course, make visual observation more comfortable. Intensification of both weak and strong sources will reduce exposure times for photographic recording. This latter effect can be an important consideration in rocket as well as in manned space observations.

REFERENCES

1. L. DUNKELMAN, W. B. FOWLER, and J. P. HENNES, *Appl. Optics* **1**, (1962) 695.
2. L. DUNKELMAN, *J. Quant. Spectrosc. Rad. Transfer* **2**, (1962) 533.
3. J. D. MCGEE, Chapter 4.1, pp. 81-103, in *Space Age Astronomy*, Academic Press, New York, 1962.
4. G. S. MORTON, *Appl. Optics* **3**, (1964) 651.
5. *Kodak Plates and Films for Science and Industry*, Eastman Kodak Co., Rochester, New York, 1962.
6. J. A. O'KEEFE, L. DUNKLEMAN, S. D. SOULES, W. F. HUCH, and P. D. LOWMAN, Jr., Chapter 19, pp. 327-347, in *Mercury Project Summary*, NASA SP-45 U.S. Govt. Print. Off., Washington, D.C. 1963.
7. L. DUNKELMAN, A. BOGGESS III, and R. SCOLNIK, Private communication.
8. J. E. KUPPERIAN, Jr., A. BOGGESS III, and J. E. MILLIGAN, *Astrophys. J.* **128**, (1958) 453.
9. T. P. STECHER and J. E. MILLIGAN, pp. 1061-67 in *Space Research III* North-Holland Publ. Co., Amsterdam, 1963.
10. A. BOGGESS III, Chapter 7 in *Space Astrophysics*, McGraw-Hill, New York, 1961.
11. J. P. HENNES, W. B. FOWLER, and L. DUNKELMAN, *J. Geophys. Res.* **69**, (1964) 2835.

LOW ENERGY CHARGED PARTICLE DETECTION USING THE CONTINUOUS CHANNEL ELECTRON MULTIPLIER*

DAVID S. EVANS

It has been found to be possible to operate the continuous channel electron multiplier, a recently developed type of windowless electron multiplier, in a gain saturated mode such that single, charged particles entering the input mouth of the channel will initiate output pulses whose amplitude and shape are both uniform and independent of the character of the excitation radiation. By suitable design of the channel this saturated pulse can be made stable against changes in such operating conditions as ambient pressure and applied voltage. When operated in this mode the efficiency of the channel for the detection of electrons over the energy range 250 eV to 10 keV is estimated to be greater than 50 percent.

INTRODUCTION

The important contribution of low energy charged particles to outer radiation belt and auroral phenomena has become apparent in the course of their investigation. In the case of visual aurora in the auroral zone, it has been found that the major portion of the total energy input is supplied by charged particles having energies below 40 keV. In one example (McIlwain, 1960)¹, a virtually monoenergetic beam of 6 keV electrons was observed in association with a visual auroral display.

Comparison of data obtained from different detectors on Injun I in the outer belt has shown that the energy flux of electrons having energies between 1 keV and 40 keV was often tenfold greater than that of electrons having energies above 40 keV (O'Brien, et al., 1962)². Very little, however, could be determined about the particle intensity or energy spectrum of this low energy component.

The efficient detection and energy analysis of these low energy particles has, in practice, posed a number of problems. It is critically important that the amount of material surrounding the active volume of the detector be minimized or, ideally, eliminated altogether. As an illustration, the Anton type 213 Geiger-Müller counter, having

an entrance window of 1.2 mg/cm² thickness, requires a minimum energy of 40 keV for electrons and 500 keV for protons to trigger a count.

Scintillation counters, having only a very thin coating of material over the scintillator to reduce light sensitivity, have been used to detect electrons with energies as low as 3 keV. In the absence of refrigeration, however, the light output from such a small energy loss in a scintillator is too low to be resolved as a pulse above the thermal noise of a photomultiplier tube. Moreover, photomultiplier tube anode pulses, arising from the liberation of only one or a few electrons from a photocathode, have a very broad pulse height distribution and so are unsuitable for counting techniques. Under these circumstances, the anode output current, rather than individual pulses, is often monitored as a measure of the total energy flux incident upon the scintillator. This method, however, is too insensitive to detect weak particle fluxes and gives no direct information about the energy distribution of the incident particles.

Bare electron multipliers with a dynode structure similar to a photomultiplier tube have been used to detect directly charged particles allowed to strike the cathode. The absence of a photon-electron conversion process at the cathode permits the use of a cathode material of relatively high work function, thereby reducing thermal noise. The detector response extends to extremely low

*Published as *Goddard Space Flight Center Document X-611-64-154*, May 1964.

energy particles because of the complete elimination of an entrance window. Either the anode current or anode pulses from such a detector may provide a measure of the incident particle flux. As with the scintillation counter, however, the output pulse height distribution is very broad. In a typical case the pulses observed at the multiplier anode, due to a beam of monoenergetic electrons striking the cathode, ranged over a factor of 100 in amplitude (D. Mathews, private communication).

A recent development along the lines of a windowless electron multiplier is the Bendix continuous channel electron multiplier described by Wiley and Hendec (1962).³ The simplicity, small size, and ruggedness of this device make it ideal for use in sounding rocket and satellite research. Consequently, an investigation has begun at Goddard Space Flight Center to determine its characteristics as a detector for low energy charged particle radiation.

DESCRIPTION OF THE CHANNEL MULTIPLIER

In its simplest form, the channel multiplier is a straight glass capillary tube having dimensions such that the length-to-inside-diameter ratio is between 50 and 100. The inside diameter itself may be from less than a tenth of a millimeter to more than a millimeter. A layer of special semiconducting material, having secondary electron emission characteristics suitable for an electron-multiplication process, is deposited over the interior surface of the tube.

When a potential difference is applied between the ends of the tube, an axial electric field is established down its length. Any electron ejected from the inside surface, for example by means of photoelectric or secondary emission processes, will thus be accelerated down the tube. Simultaneously, the electron will drift *across* the tube with whatever lateral velocity was acquired in the

ejection process. The electron multiplication occurs when the potential difference and tube dimensions are such that these free electrons gain enough energy from the electric field between encounters with the surface that, on the average, more than one secondary electron is generated at each encounter.

In this manner (Figure 1) a single electron ejected at the low potential or input end of the tube can result in an electron cascade of some magnitude at the high potential or output end of the channel. A typical experimental channel multiplier gain curve obtained by Wiley and Hendec is shown in Figure 2.

The exact nature of the semiconducting surface has not been published, although it has been shown by Angel, et al. (1961)⁴ that the relative spectral response of the material to the photoelectric effect is similar to that of tungsten, implying a work function of about 4 eV. Thus the channel has virtually no response to light of wavelengths longer than 1500A and essentially no background noise due to thermionic emission of electrons. The surface is, moreover, quite stable with respect to changes in the secondary emission ratio after extended and repeated exposure to the atmosphere. Hence, no special handling is necessary.

In Wiley and Hendec's investigations beta radiation, extreme ultraviolet light, and low energy photoelectrons from an illuminated photocathode were among the sources used to excite the channel. In all cases, the output *current* collected by a separate biased anode at the output end was compared to the input excitation current to obtain the electron gain data. In general, the channel was operated in these investigations at a gain which did not permit observation of the individual anode pulses arising from single events at the front end of the channel; hence, no work was done on this aspect of the channel multiplier.

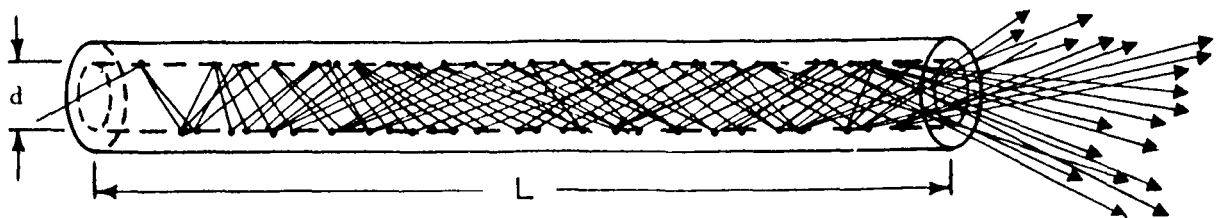


FIGURE 1.—Illustration of the electron multiplication process within the channel electron multiplier.

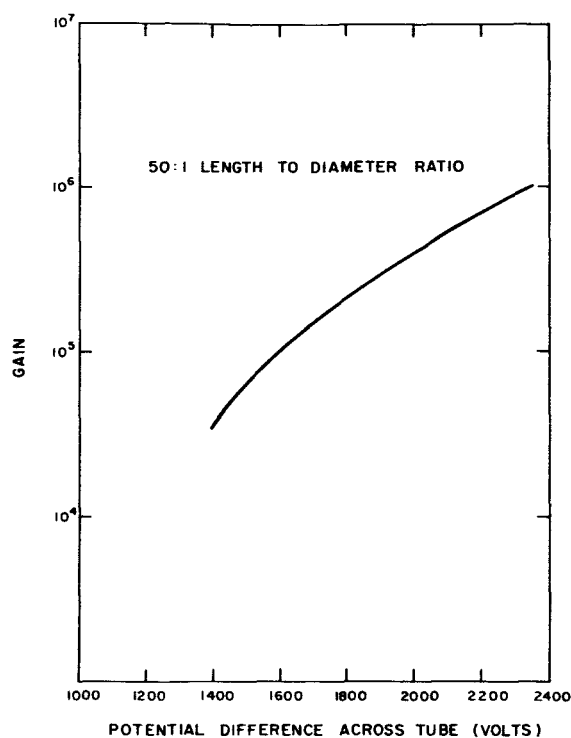


FIGURE 2.—An experimentally determined electron gain curve (after Wiley and Hendee).

The work reported here, on the other hand, concentrated upon the characteristics of the pulses appearing at the output end of the channel.

DESCRIPTION OF THE PRESENT WORK

For the initial, exploratory phase of the investigation, a number of straight channel electron multipliers were obtained. Most of these channels had a length-to-bore ratio of 50 and an inside diameter of 1 mm. For comparison, however, two channels with inside diameters of 1 mm and length-to-diameter ratios of 100, and two channels having inside diameters of approximately .5 mm and length-to-diameter ratios of 50 were also obtained. The measured ohmic resistance of the semiconducting layer was generally in the range 10^9 to 10^{10} Ohms. This resistance is the determining factor for both the power dissipated in the channel and any effects associated with its RC relaxation time.

The experimental scheme most often employed to investigate the channel's pulse output is illustrated in Figure 3. The output end of the channel

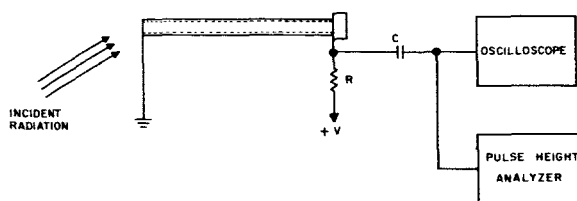


FIGURE 3.—The experimental setup employed for the study of output pulses.

was essentially covered by a foil to collect the total charge of the electron cascade. The voltage pulse developed across the shunt capacity of the collection circuit was displayed on both an oscilloscope and the output of a pulse height analyzer. With a knowledge of the collection circuit shunt capacity an electron gain could be determined from the voltage pulse amplitude on the assumption that the cascade was initiated by a single secondary electron.

Normally the collection end of the tube was operated at a potential between 1000 V and 4000 V with respect to both the input end and the surrounding stainless steel vacuum system. This procedure lessened the chance of stray positive ions being attracted into the mouth of the channel causing anomalous counts. To expose any pressure sensitivity in the multiplication process, the channel multiplier was operated at ambient pressures ranging from 10^{-3} to 10^{-6} mm Hg. The excitation radiation was provided by Co^{60} , Fe^{55} , or S^{35} radioactive sources, or by the low energy charged particles produced by a Bayard-Alpert pressure gauge placed near the channel.

The first examination of the channel's output pulses was made at an applied potential of 1400 V and a pressure of less than 10^{-5} mm Hg. Under these conditions the channel gain was about 10^5 and a charge sensitive amplifier proved necessary in order to observe the output pulses at all. The pulse height distribution of the output pulses was found to be very broad, the differential pulse height distribution displaying a tendency to continue rising with decreasing pulse height. It may be noted that the channel was operated in this case with a gain similar to the gains employed in the *current* output measurements of Wiley and Hendee.

Further investigation showed, however, that when the channel's electron gain was increased to

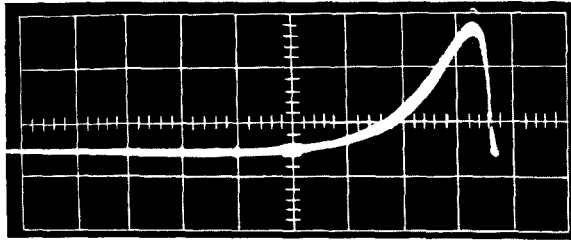


FIGURE 4.—Saturated pulse output from a straight channel multiplier. The pulse amplitude is 100 mV and the pulse width is about 4μ sec.

more than about 10^7 , the pulse height distribution of the output pulses became peaked; i.e., all of the output pulses had nearly the same amplitude which was independent of the character of the exciting radiation. Figure 4 shows an oscilloscope photograph of this effect. The appearance of this uniform output pulse is attributed to a saturation in the electron multiplication process basically caused by the magnitude of the electron cascade.

The high gain necessary for the appearance of the saturation effect was achieved through two separate but interdependent processes. The first was simply an increase in the potential difference applied across the tube. Up to a point, this procedure will increase the gain by increasing the average energy acquired by the individual electrons between collisions, and so increasing the secondary electron emission ratio.

The second mechanism depends on positive ion production by interaction between cascade electrons and residual gas molecules. These free ions are accelerated by the electric field toward the input end of the tube where they may interact with the wall to initiate a new cascade. This second, regenerative feedback effect dominates the behavior of the straight channel multiplier in the high electron-gain mode of operation.

The clearest evidence of the role of ion feedback in controlling the multiplication process was the existence of a strong pressure dependence in both the amplitude and the character of the output pulse. If, for example, the potential difference was adjusted to yield a satisfactory output pulse at 10^{-4} mm Hg, a decrease in pressure to 10^{-6} mm Hg would cause the output pulse height distribution to revert to the low amplitude, broad, "washed out" type of distribution previously

associated with electron gains on the order of 10^5 . When the procedure was reversed and the applied potential set at a value sufficient to give a saturated output at 10^{-6} mm Hg, an increase in pressure to 10^{-4} mm Hg produced distortions in the output pulse shape, and instabilities in the form of spurious counts and often continuous, uncontrollable pulsing.

The contribution of ion feedback was also evidenced in the tube's response to penetrating radiation. Under exposure to gamma rays from Co^{60} , the channel output pulses exhibited the same peaked amplitude distribution as was found when the excitation was by low energy charged particles unable to initiate cascades except at the very mouth of the tube. The unsaturated pulses that would have resulted from gamma ray interactions liberating electrons well away from the input end of the tube were not seen because of the efficient ion feedback mechanism.

Perhaps, however, the most significant manifestation of the ion feedback mechanism was revealed in its effect on the leading edge of the output pulse. The rise time of the output pulse was observed to be typically 0.5 microsecond, which compares poorly with the few tens of nanoseconds estimated from the transit time down the tube of individual cascade electrons. Close examination of the leading edge of the pulse revealed a "staircase" appearance. This strongly suggested that the observed output pulse was the sum of numerous, successive electron cascades, each initiated by positive ions accelerated down the tube and striking the inner surface near the input end. The fact that this "staircase" effect was observed at ambient pressures less than 10^{-6} mm Hg indicates both the large number of electrons involved and the consequences of permitting relatively few ions to drift unobstructed to the input end of the tube to restart a cascade. Indeed, the appearance of the pulse's leading edge is felt to be the most sensitive guide as to the presence and extent of ion feedback in the multiplication process.

DISCUSSION OF THE SATURATION EFFECT

The tentative explanation of the gain saturation in the channel multiplier, which accounts for the appearance of output pulses of uniform amplitude, is based upon the alteration of the potential

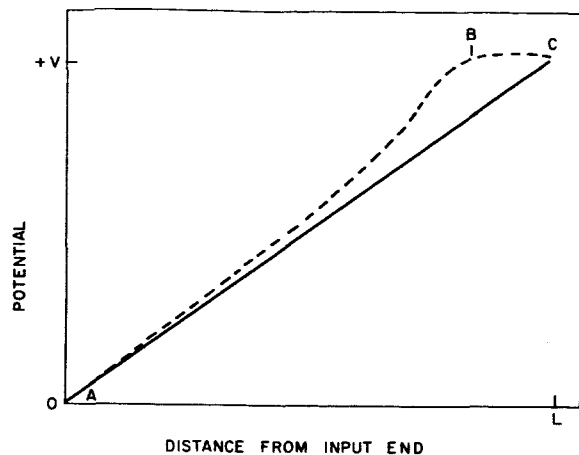


FIGURE 5.—The potential plotted against position along the channel multiplier. The solid line is in the absence of a cascade. The dashed line illustrates the change to be expected at the height of the cascade.

gradient within the tube by surface charges deposited on the interior walls of the tube during the multiplication process. The solid line in Figure 5 displays the voltage as a function of position within the tube in the absence of a pulse, the semiconducting surface being assumed to be uniform down the tube. During the electron multiplication process, a positive charge accumulates upon the walls of the tube because of a net loss of electrons from the surface. The potential vs. position curve may then be altered to the one depicted qualitatively by the dashed line in Figure 5. One will note that in the region BC (near the output end) the potential gradient has flattened, and the cascade electrons can no longer acquire the energy necessary to support the multiplication process. On the contrary, this portion of the tube will tend to become an electron sink, thereby limiting the gain.

Although no quantitative work has been done using this saturation model, observations on the nature of the multiplication process as a function of position in the tube provide a strong basis for believing that the gain limiting is indeed due to such a mechanism. In order to investigate this, a positively biased anode plate was placed at the output end of the tube to collect the cascade space charge and so insure that alterations of potential in the tube itself were due to only surface charges. The potential changes within the tube were observed by utilizing the capacitive coupling of a wire looped tightly around the outside of the tube.

It was found that when the pickup loop was positioned toward the input end of the channel a positive pulse was observed, indicating that electron multiplication was occurring. As the loop was moved toward the output end, a region was reached where the pulse was characterized by an initial positive phase followed by a negative phase. At a position still further toward the output, the pulse was entirely negative, indicating a net accumulation of electrons and a lack of multiplication. Whenever the channel was operated in the non-saturating mode, a positive pulse was observed at all points down the length of the channel, the entire tube apparently being utilized for electron multiplication.

The model predicts the existence of a period, while the surface charge is being redistributed, during which reduced amplitude or no pulses will be observed. Such a dead time, lasting about 100 microseconds, was observed at high count rates. This dead time is, in principle, connected with the RC relaxation time of the channel, and hence may be reduced by decreasing the ohmic resistance of the semiconductor strip. Such a procedure would be limited by the maximum allowable power dissipation of the device.

DISCUSSION OF THE INSTABILITY

The investigation of the channel multiplier was primarily directed towards its use as a radiation detector in satellite and sounding rocket experiments. The pressure environment is uncontrollable in both systems; in the case of the sounding rocket changes in pressure over many orders of magnitude must be tolerated. Consequently, much of the investigation concerned itself with eliminating the dependence of the electron gain upon pressure and understanding the nature of the instability mentioned earlier.

If the channel multiplier gain was adjusted to yield a saturated, stable output pulse at 10^{-6} mm Hg and the pressure was then allowed to increase slowly as was described earlier, the instability first appeared as secondary after-pulses occurring 30–100 microseconds after the initial primary pulse (induced by a radioactive source, for example). Further pressure increases caused continuous pulsing, even in the absence of excitation radiation. The transition from after-pulsing to

continuous pulsing was a smooth and reversible one, and both are believed to be aspects of the same fundamental instability.

The relatively long time delay before the appearance of an after-pulse suggested at first that, if charged particles were directly responsible, the process giving rise to the instability was taking place over a volume much larger than the interior volume of the channel itself. However, experiments performed with the open ends of the tube effectively blocked by aluminum, stainless steel, or copper foils demonstrated that the unstable after-pulsing arose from within the capillary itself.

Prospective models to explain the instability divided themselves into two crude categories. Models in the first class were basically dependent upon the magnitude of the electron cascade. Among these was, for example, the thin-film field emission of electrons from the semiconducting layer which had acquired a large surface charge density as a direct result of the electron multiplication process. The possibility exists that field emission of electrons could persist long after the termination of the primary pulse and thus cause after-pulsing. The effect of any ambient pressure variation would simply be to change the magnitude of the electron gain and hence the surface charge densities involved.

The second set of explanations is fundamentally connected with the existence of significant ion feedback in the channel. Referring to Figure 5, one observes that, although the potential alteration in the region BC effectively prevents an electron cascade developed in the region AB from arriving at the output, electron multiplication is still possible in the front portion. Hence, in this region of the tube, the regenerative ion feedback-electron multiplication process could constitute a means of self-perpetuating the presence of free charged particles until the back portion had recovered to the point of again allowing cascades to pass to the output.

A third "combination" model, based on the production of neutral metastable molecules by the electron cascade, depends upon both the presence of residual gas in the tube and the creation of a large electron cascade. However, attempts to accentuate such a process through the introduction of helium rather than air as the residual gas proved fruitless.

The strong interdependence between the magnitude of the electron cascade and extent of ion feedback in the straight channels did not allow a clear demonstration as to which of the alternative explanations of instabilities was valid.

The dominating influence of regenerative ion feedback in the straight channel is due essentially to the fact that ions, which are mostly produced at the back end of the tube where the electron cascade is largest, are permitted to move to the input end where the subsequent liberation of electrons has the most effect. Clearly a channel electron multiplier in the form of a curved capillary tube would contain these ions in the back end where their effect on the overall electron gain would be negligible. Such a channel multiplier would have the additional benefit of minimizing the regenerative ion-feedback-electron-multiplication process and thus give an indication of the basic cause of the instability.

DESIGN OF THE CURVED CHANNEL

In order to design a curved channel multiplier a crude calculation of the loop gain of the ion feedback-electron multiplication mechanism was performed. The radius of curvature of the capillary tube was then chosen to reduce the loop gain enough to insure quenching of free charged particles well within the observed dead time.

Figure 6 shows a section of channel multiplier with a radius of curvature R_0 . The path taken by an ion produced at the point A and moving under the influence of the axial electric field until striking the wall at B is shown schematically. The equations of motion governing the ion are:

$$m \frac{d^2 r}{dt^2} - mr \left(\frac{d\theta}{dt} \right)^2 = 0 \quad (1)$$

$$mr \frac{d^2 \theta}{dt^2} + 2m \frac{dr d\theta}{dt dt} = eE \quad (2)$$

By neglecting the coriolis force term $\frac{dr dt}{dt dt}$ and approximating r by R_0 in Eq. 1, one may integrate the equations to obtain

$$t = \theta R_0 = \sqrt{3R_0 d} \quad (3)$$

for the maximum possible distance through which an ion produced at A will pass before striking the wall at B.

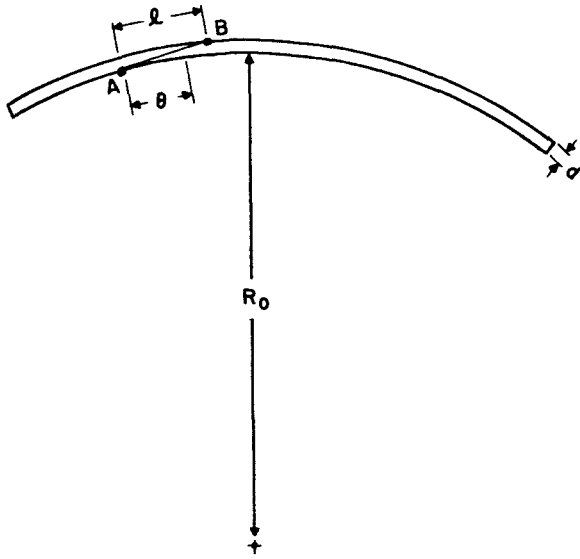


FIGURE 6.—A section of a curved channel multiplier displaying the quantities used in the calculation.

Because the electron multiplication may be considered uniform along the length of the tube, it follows that the electron gain over a length x of the tube is related to the electron gain over the full tube length L by the expression

$$G(x) = G(L)^{x/L} \tag{4}$$

$G(L)$ being approximately 10^8 for a saturated pulse.

Given a number of electrons N passing a point P in the tube, the number of ions created as a function of the distance x further down the tube is given by

$$N_{ion}(x) = \alpha N G(L)^{x/L} dx \tag{5}$$

where α is the first Townsend coefficient.

We take the worst case and assume that all these ions move the maximum permissible distance l before striking the interior wall of the tube and, with a probability of 1.0, eject a secondary electron. This electron is itself multiplied up to give a net gain at the point P of

$$dN' = \alpha N_0 G(L)^{x/L} G(L)^{(l-x)/L} dx \tag{6}$$

which, when integrated to obtain the contribution of all the ions of interest, gives

$$N' = \alpha N G(L)^{l/L} l \tag{7}$$

as the total number of electrons arising through the feedback mechanism.

Recalling that $l = \sqrt{3R_0d}$ and using the maximum α at a design pressure of 10^{-3} mm Hg, one obtains for the loop gain

$$\frac{N'}{N} = 0.012 \sqrt{3R_0d} G(L) \sqrt{3R_0d/L}$$

The time necessary for each oscillation of the ion-electron process is no more than a few tenths of a microsecond. Hence, assuming the peak number of electrons to be 10^8 , a loop gain of .01 will insure that stray charged particles are completely eliminated within a few microseconds. Further, a loop gain of .01 would indicate that the contribution of ion feedback to the overall electron gain of the tube is no more than 1 percent at a pressure of 10^{-3} mm Hg. Hence, Eq. 8 becomes

$$\sqrt{3R_0d} \cdot 10^{8\sqrt{3R_0d/L}} = .83$$

By setting the inside diameter d equal to 1 mm and the length L equal to 100 mm, the equation may be solved numerically to obtain

$$R_0 \approx 2.4 \text{ cm.}$$

A channel multiplier designed to such specifications would be in the form of a circular arc extending over about $\frac{2}{3}$ of a circle having a radius of 2.4 cm.

In addition to the approximations made specifically in the calculation, the effect of gas adsorbed upon the inner walls has been neglected. The ionization of this adsorbed gas is probably most important at relatively low pressures and accounts for ion feedback observed below 10^{-6} mm Hg. The effect upon the gain function $G(x)$ of altering the potential gradient inside the tube during a cascade has also been neglected. However, the choice of the loop gain is conservative to the point that if changes in $G(x)$ occur, the electron cascade will in all probability still be quenched.

CHARACTERISTICS OF THE CURVED CHANNEL MULTIPLIER

The curved channel multiplier obtained from the Bendix Corporation, shown in Figure 7, had an inside diameter of 1 mm and a length-to-diameter ratio of 100. The capillary tube had the form of an arc extending over $\frac{3}{4}$ of a circle having

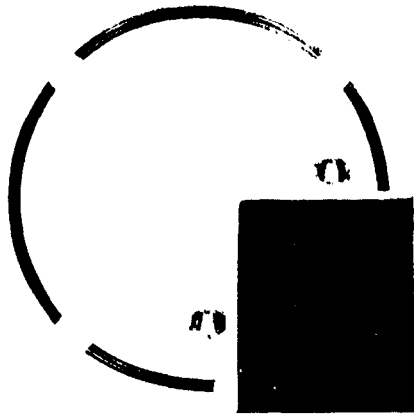


FIGURE 7.—The curved channel multiplier used in the experimentation. The channel is mounted in a Kel-F block for ease of handling.

a radius of curvature of 2.0 cm. The experimental arrangement was identical to that used with the straight channel multipliers.

It was quickly established in the initial trial of the channel that the saturated pulse mode of operation was induced at an electron gain of about 5×10^7 , roughly the same gain as was found necessary in the straight channel. In order to achieve this gain, an applied potential of 3000 V was required.

To determine whether ion feedback was interfering with the electron gain mechanism, a careful examination was made of the leading edge and rise time of the output pulse. It was found that at all pressures below 10^{-3} mm Hg the rise time of the output pulse was less than 20 nanoseconds; moreover, the leading edge gave no evidence of distortion by ion feedback effects. Subsequent observations showed no variation in the electron gain of the channel multiplier during a change of more than two orders of magnitude in pressure, thus substantiating that there was no effective contribution from regenerative ion feedback to the electron multiplication process.

Determined efforts were made to excite the instability by increasing both the ambient pressure and the applied potential to the point where arcing became a danger. However, no instability was noted, either in the form of spurious or continuous pulsing or in the form of after-pulsing. The sole observable effect was a distortion in the shape of the output pulse evident at pressures

above a few microns. This pressure, however, is well above the maximum pressure expected during the operation of the channel in sounding rocket or satellite experiments.

It is clear that either the ion feedback-electron multiplication process described earlier or another mechanism dependent upon ion feedback effects was the basic cause of the instabilities observed during investigations using the straight channel.

Figure 8 displays the output pulses from the curved channel in a case where the excitation was by 200 eV ions produced by a Bayard-Alpert type gauge. A comparison with the output pulses from a straight channel (Fig. 3) shows an apparent loss of resolution or "sharpness" in the pulses from the curved channel. The peaked pulse height distribution remains, however, entirely compatible with straightforward pulse-counting techniques. It should be pointed out that the rise time of the pulse shown in Figure 8 has been lengthened to about 0.1 microsecond by an emitter-follower circuit used at the output in order to drive the shunt capacity and thus obtain greater voltage amplitude.

Once a channel multiplier having a stable pulse mode of operation had been obtained, an examination of the efficiency of the channel for the detection of very low energy charged particles was begun. This efficiency, in principle, depends first upon the probability that the incoming primary particle will produce one or more secondary electrons when striking the input end of the channel, and secondly upon the probability that the initial secondary electron will result, after multiplication, in a saturated output pulse (collection efficiency).

Massey and Burhop⁵ have pointed out that the energy distribution of secondary electrons emitted

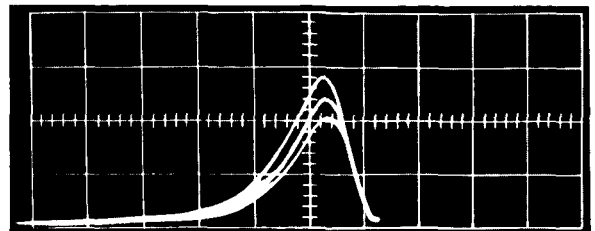


FIGURE 8.—The saturated output pulses from the curved channel multiplier. .21 V/cm, .2 μ sec/cm.

from a surface under impact from primary charged particles, is not strongly dependent upon the type or energy of the primaries, the secondaries in general being emitted with energies of only a few eV. Hence, it may be expected that the "collection efficiency" of the channel will remain more or less independent of the character of the incident radiation. One may conclude, then, that changes observed in the channel's counting efficiency due to changes in the incident radiation essentially reflect the variation in the secondary electron emission ratio with the energy or type of the incident particles. Therefore, by considering the general behavior of the secondary emission ratio exhibited by all materials, one may obtain a crude picture of the energy dependence of the detection efficiency of the channel.

It would be expected, for example, that the channel responds initially to electrons having energies of about 10 eV. The relative efficiency should then increase with increasing primary electron energy, reaching a maximum for some energy in the range from a few hundred to a few thousand eV. As the primary electron energy is increased still further, the secondary emission ratio displays—in all materials—a long, very slow decline which will be mirrored in a similar decline in counting efficiency. In the case of tungsten, of interest because its photoelectric spectral response is similar to that of the material used in the channel, the secondary electron emission ratio has a maximum of 1.4 for an incident energy of 700 eV and declines to .12 for primary electron energies of 200 keV.

When ions (protons) are allowed to excite the channel, the counting response should first be observed with ions having an energy of a few hundred electron volts. Above this threshold the secondary emission ratio, and, therefore, the counting efficiency, rapidly increases with increasing ion energy, and, judging from typical secondary emission curves, this relative efficiency remains large for ion energies up to several hundred keV.

It should be pointed out that particles of these energies may directly initiate a count only by entering the channel through the open mouth. It is estimated that to penetrate the tube's glass walls and initiate a cascade at a point well down the tube from the input end, an electron must have about 0.5 MeV and a proton about 5 MeV.

Excitation by x-ray bremsstrahlung from lower energy electrons is a possibility, of course.

The collection efficiency, acting in a sense as the normalizing factor for the relative response curve, is believed to be quite high; probably greater than 50 percent and perhaps as high as 80–90 percent. This conclusion is based on a measurement of the channel multiplier's absolute efficiency for the detection of beta rays from S^{35} . This particular beta spectrum has an end point at 169 keV and an average electron energy of about 60 keV. For incident electrons of this average energy, the secondary emission ratio for *true*, low energy secondary electrons is about 0.35, based upon the data for tungsten. The absolute efficiency for the detection of betas of the S^{35} spectrum was observed to be about 25 percent, from which one may infer a collection efficiency of around 75 percent. Uncertainties in the calibration of the radioactive source and possible error in using the properties of tungsten rather than those of the actual channel multiplier surface leave some doubt as to the exactness of this figure. However, the figure of 75 percent does not seem to be in great error. The fact that relatively few cascades are terminated during the first stages of multiplication by the complete loss of an electron into a wall is probably due to the cascade electrons being constrained to intersect the wall at a glancing angle. This insures that the secondary electrons are produced very near the wall surface and are thus able to leave the wall with very little possibility of absorption in the semiconducting material.

In addition to the efficiency measurement made with the aid of a radioactive source, attempts were made to obtain similar data by using a beam of low energy, monoenergetic particles from a Van de Graaff generator to excite the channel. These experiments were unsuccessful because of the inability to measure the very weak input particle flux that was necessary to avoid dead time effects in the channel. On the other hand, experiments of this kind are critically important in determining accurate detection efficiency vs. particle energy characteristics, and work has begun to solve the technical problems involved.

Of great interest also insofar as practical application is concerned, is the relative variation in detection efficiency with changes in either the

ambient pressure of the applied potential. An investigation of such effects was easily carried out by using a stable radioactive source, in most cases S^{35} , and the results were very encouraging. It was noted, for example, that the count rate of the channel due to the radioactive source displayed no systematic variation that could be attributed to a change of two decades in ambient pressure. This result might have been expected from the pressure independence of the electron gain in the curved channel.

When the count rate was plotted against applied potential (Fig. 9), no significant change in efficiency was evident as long as the electron gain was sufficient to maintain the saturated mode of operation. During this phase of the experiment it was noted that if the applied voltage was increased rapidly by several hundred volts it required several minutes for the counting rate to stabilize to the nominal value from an initial somewhat higher value. The net variation in count rate during the stabilization amounted to only one to two percent and would not, in practice, give rise to any trouble.

The long term count rate stability of the channel multiplier was demonstrated in the course of these experiments in which no gross changes in count rate were observed during continuous operation lasting over several days. However, short term decreases in the electron gain of the channel were induced by operating the channel at very high count rates. This fatiguing would result in gain changes ranging as high as 50 per-

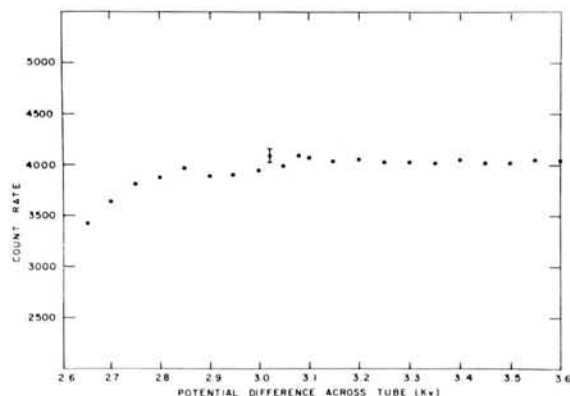


FIGURE 9.—The count rate of the channel is a function of applied potential. The excitation was provided by a S^{35} source.

cent and persisting for up to five hours. Fortunately, the fatigue in gain was not accompanied by any apparent effect upon the detection efficiency of the channel—provided, of course, that the degraded gain remained sufficient to result in a saturated output pulse.

The conclusion on the basis of these preliminary studies is that as long as the channel is operated with an electron gain of five to tenfold higher than that needed to insure a saturated pulse, it will provide a suitable and reliable low energy charged particle detector for space physics experiments. Of particular importance, the curved channel multiplier appears to exhibit excellent short and long term stability when operated in the counting mode, even in the presence of relatively large changes in applied potential, wide variations in ambient pressure, and gain fatigue effects occurring at excessively high counting rates.

Among the numerous further investigations that are contemplated are the determination of counting efficiency as a function of incident particle energy and a study to ascertain the effect, if any, of gas adsorbed upon the interior surface of the tube. This second phenomenon could cause

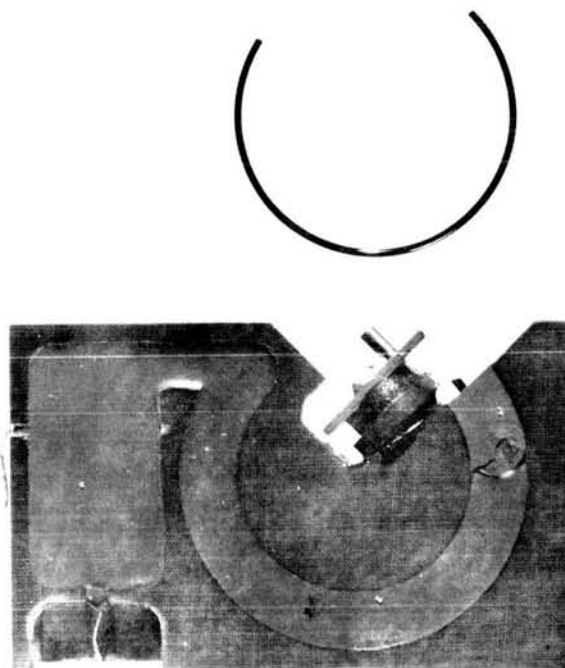


FIGURE 10.—A flight model of a channel multiplier detector unit, with an unpotted channel beside for comparison.

very long term drifts in the efficiency of a channel multiplier in a satellite.

Because the output pulse from the channel multiplier carries no information concerning the type or energy of the primary radiation, knowledge of these parameters must be obtained by magnetic or electrostatic analysis of the incident radiation prior to its entering the channel. Fig. 10 displays one simple approach to the solution of this problem. The channel multiplier, its output end sealed with a protective cap, has been potted into a groove cut in a phenolic block. The rectangular area in the block contains the potted load resistor and output coupling capacitor. A small permanent magnet positioned behind the collimeter tube provides analysis of the type and energy of the radiation entering through the collimeter.

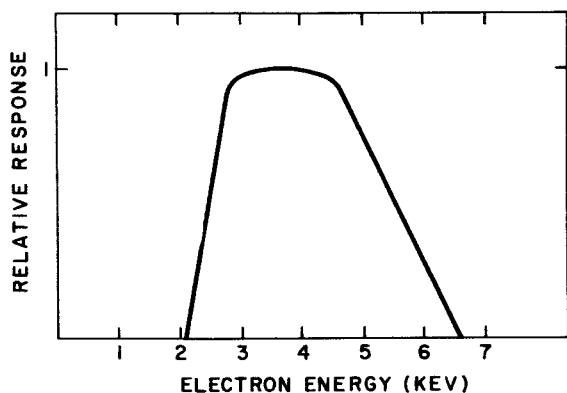


FIGURE 11.—The relative response of the detector system as a function of energy.

The quality of the energy discrimination for electrons is displayed in Fig. 11 where the relative response of the detector system has been plotted against the incident electrons. The magnet strength and position were chosen to obtain maximum response at 3.5 keV, and in this case a very constant response from 2.75 to 4.5 keV was obtained.

A number of units of this design, utilizing magnets of varying strengths to obtain a particle energy spectrum, were successfully flown on sounding rockets during the spring of 1964 to obtain information on the low energy portion of the auroral particle spectrum.

REFERENCES

1. C. E. McILWAIN, "Direct Measurement of Particles Producing Visible Auroras," *J. Geophys. Res.* 65, 2727-47. (1960)
2. B. J. O'BRIEN, C. D. LAUGHLIN, J. A. VAN ALLEN, and L. A. FRANK, "Measurements of The Intensity and Spectrum of Electrons at 1000-Kilometer Altitude and High Latitudes," *J. Geophys. Res.* 67, 1209-1225. (1962)
3. W. C. WILEY and C. F. HENDEE, "Electron Multipliers Utilizing Continuous Strip Surfaces." Paper presented at The Eight Scintillation and Semiconductor Conference, Washington, D.C., March 1-3, 1962.
4. D. W. ANGEL, H. W. COOPER, W. R. HUNTER, and R. TONSEY, "Extreme Ultraviolet Detecting with The Bendix Single Channel Photomultiplier." Paper presented at The Image Intensifier Symposium, Fort Belvoir, Virginia, October 24-26, 1961.
5. H. S. W. MASSEY and E. H. S. BURLOP, "Electronic and Ionic Impact Phenomena," Oxford, Clarendon Press, 1952.

AN ULTRAVIOLET SPECTROPHOTOMETER FOR SATELLITE ASTRONOMY*

KENNETH HALLAM AND JOHN MANGUS

An astronomical spectrophotometer is described which was designed for use on OSO-B, the second U.S. Orbiting Solar Observatory. The spectrophotometer is intended to analyze the ultraviolet emission from stars and nebulae. The entrance slit of the spectrophotometer defines the field at the primary focus of a 6-inch Gregorian reflecting telescope. The secondary serves as the collimator for a rotating 1200 lines/mm plane reflection grating. Diffracted light from the grating is collected by an off-axis paraboloid and focused through the exit slit to the photomultiplier. The signal from the photomultiplier is pulse-counted and fed into the magnetic tape storage of the space-craft. The instrument utilizes the spin of the gyroscopically-stabilized satellite to sweep the $1^\circ \times 0.4^\circ$ field over the celestial sphere. Slow precession of the spin axis allows eventual coverage of the whole sky. Ten adjacent wavelength bands of 180 Å width between 1500 Å and 3300 Å are defined by successive steps of the grating. An input flux of 4×10^{-9} erg/sec. cm^2 Å or over results in a statistical error in the output of 2 percent, while a flux 8×10^{-11} erg/sec. cm^2 Å may be determined within 10 percent mean error. This sensitivity will permit useful observations of stars 5 stellar magnitudes fainter than the brightest ones. Details of the optical system and special requirements imposed upon it by characteristics of its special environment are reviewed.

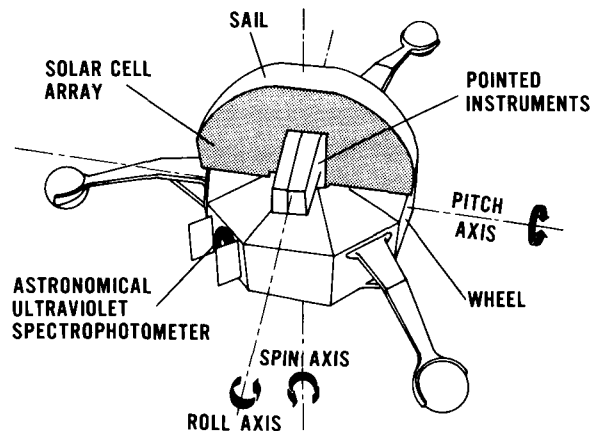
INTRODUCTION

An astronomical spectrophotometer is described which is designed to gather comprehensive data on the ultraviolet radiation from the brighter stars and nebulae. The experiment is constructed for use in a wheel compartment aboard the Orbiting Solar Observatory B spacecraft. Precession of the satellite spin axis about the solar vector combined with the earth's movement with respect to the sun are utilized in obtaining nearly full sky coverage of both the northern and southern hemisphere. Spectral scanning is accomplished by the incremental movement of 1200 line/mm original plane reflection grating blazed for 2000Å. Its spatial resolution of $1^\circ .00 \times 0^\circ .44$ is preserved by sequencing the spectrophotometer output to an azimuth sensor. Viewing of scattered solar irradiance from the Earth's upper atmosphere necessitated the incorporation of an albedo sensing system into the experiment to minimize residual saturation effects in the spectrophotometer detector.

*Published as *Goddard Space Flight Center Document X-610-64-188*, July 1964. This work was also published in the *Proceedings of the International Conference on Optics* in the section on Photographic and Spectroscopic Optics, Tokyo, Japan, September 1964, and in a supplement of the *Japanese Journal of Applied Physics*, May 1965.

SPACECRAFT

The Orbiting Solar Observatory B as shown in Figure 1 consists of a pointed experiment compartment, a solar array and a rotating wheel section. In orbit, the oriented experiment compartment is aligned parallel to the solar vector and the solar array is kept nearly perpendicular to the solar vector by sun sensors and the control system.



ORBITING SOLAR OBSERVATORY B

FIGURE 1.—Orbiting Solar Observatory B.

The spin axis of the satellite is nominally perpendicular to the solar vector throughout the useful lifetime of the spacecraft. This perpendicular relationship is maintained within ± 3.5 degrees by means of the control system. However, the spacecraft is not controlled in roll about the solar vector and will precess approximately one degree in 24 hours. This precession will cause the spin axis to sweep an arc through approximately one half of the celestial sphere during the course of the six months satellite lifetime. Similarly the plane in which the wheel rotates will sweep over the entire celestial sphere. By mounting the spectrophotometer in the wheel plane, it may therefore cover both the northern and southern hemisphere with high positional accuracy.

The spinning wheel section, which provides gyroscopic stability to the spacecraft, is divided into nine compartments. Five compartments house experiments and the other four contain telemetry, control and command systems.

SPECTROPHOTOMETER

Unlike a laboratory instrument, one must work backwards, by first determining the constraints placed on the spectrophotometer design by the spacecraft and the spacecraft environment. The high rotational rate (30 rps) of the wheel indicates that observations should be made looking parallel to the spin axis to gain sufficient time for accurate measurement of the flux gathered by the instrumental optics. Although this is quite possible, analysis of the satellite motion indicates that only an extremely limited region of the sky would be observed and it appears improbable even to optimize the viewing path through the Galaxy. Therefore, we have mounted the spectrophotometer to look along a radius of the wheel, but with a depression angle of 5° simply to prevent direct viewing of the sun.

The field angle is precisely defined by a slit located at the focus of the paraboloidal objective of a Gregorian type telescope as illustrated in Figure 2. The opposed secondary is also a paraboloid, and becomes the collimator for the spectrophotometer. A major draw back of this system is that the field angle is magnified by the secondary or collimating mirror. As the field angle is increased the angular dispersion of the monochromator must be increased to maintain a con-

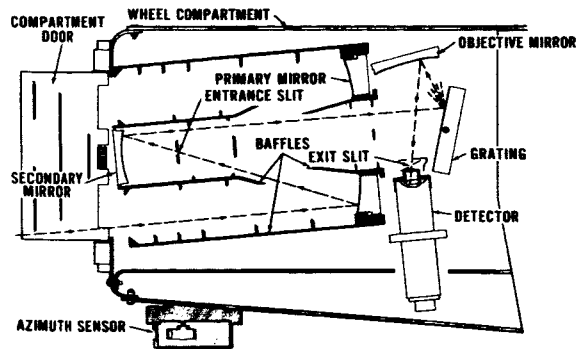


FIGURE 2.—Astronomical ultraviolet spectrophotometer.

stant spectral resolution. The diameter of the secondary mirror must also be sufficient to prevent vignetting of the field, even though this decreases the free area of the primary mirror. Because of these factors, the diameter of the collimated beam from the secondary that passes through the central hole in the primary is nearly independent of the focal length of the secondary.

An overall light rejection ratio on the order of 10^{-8} to 10^{-11} must be achieved to protect the sensitivity of the instrument to starlight while bathed in sunlight. Doors on the spacecraft rim panel increase the baffling length to the entrance pupil at the primary mirror, thus allowing useful operation closer to the sun. The incident solar radiation, which is reduced by diffuse scattering from baffles painted with Parson's Optical Black Laquer, must incur at least two reflections before passing through the hole in the primary mirror.

The primary mirror maximum diameter is limited to six inches due to tipping of the optical axis and restricted mounting space. To minimize obscuration of the primary mirror by the secondary mirror and for optimization in packaging the objective, the design ratio of their focal lengths is weighted against the effective area and focal ratio of the primary mirror.

The optimal azimuth field angle θ and elevation field angle ϕ are largely determined by the spacecraft rotational and orbital motion. The satellite performs about 15 orbits per day or on the order of 29000 usable spacecraft rotations per day. To prevent gaps in scanning successive 180° azimuthal arcs of the celestial sphere as they are shifted due to the daily precession of the spin axis, one is left with 80 looks per degree of azimuth. The 80 looks per degree are then sub-

divided into X looks per spectral bandpass $\Delta\lambda$ for each of Y adjacent bandpasses. X , Y , $\Delta\lambda$, θ , and φ are then chosen by the expected sensitivity of the total spectrophotometric system, by the anticipated flux and spectral energy distribution of stellar sources, by their average separation on the plane of the sky, and by the characteristics of the OSO data handling system.

Several factors must be considered in the monochromator design. The dispersive element, collecting mirror and detector must be packaged behind the Gregorian objective. The entrance slit must be wide enough to assure field overlap from one scanned band in the sky to the next precessed one, and the angular magnification due to the monochromator collimator must be limited to permit sufficient linear dispersion across the exit slit without spectral overlap.

At present, original reflection diffraction gratings are the most reliable spectral resolvers in space applications. Hass¹ has shown that the reflectivity of ultraviolet coatings changes less than 1 percent when irradiated by 1 MEV electrons and 5 MEV protons. In addition, Hass had made coatings with a reflectivity of 85 percent at 1400 Å. Gratings blazed in this wavelength region are available with ultraviolet coatings. An original grating was selected for use because at this time data concerning the deterioration of the more efficient replica gratings in high vacuum is inconclusive.²

The working angle of the grating was chosen to maximize the spectral purity. Because of the compactness of the dispersing system, second order spectra must be rejected by meticulous baffling as shown in Figure 2.

An off-axis $f/65$ paraboloidal collector mirror with a focal length of 9.6 cm is used to obtain a linear dispersion of 64 Å/mm. Ten adjacent bandpasses 180 Å wide are scanned from 1498 Å to 3298 Å without second order overlap. This should be sufficient resolution, to study nebular emission sources in primary spectral detail, as predicted by Daub³ and Osterbrock.⁴ With a resolution of $\frac{\lambda}{\Delta\lambda} = 10$, a bandpass of 180 Å separates many of the anticipated lines except for doublet structure.

An Ascop 541 photomultiplier tube with a cesium telluride photocathode on a sapphire win-

dow was selected as the detector because of its combined ruggedness and spectral sensitivity. The tube has a gain of 17×10^6 at 2750 volts, and a peak quantum efficiency of about 10 percent at 2500 Å. Photocathode nonuniformity is decoupled from variations of source brightness within the field by use of a calcium fluoride field lens behind the exit slit to image the primary mirror onto the photocathode. This lens also prevents vignetting by the photocathode and gains clearance between the grating and photomultiplier.

The output of the photomultiplier tube is gated into the photo-pulse counter by an azimuth phase sensor which consists of a very small auxiliary telescope mounted on and perpendicular to a rotating shaft below the spacecraft heat shield. This azimuth shaft, which is parallel to the spacecraft spin axis, is rotated back and forth about its own axis through an arc of 180 degrees, in one degree increments, by a Bendix stepping motor. The telescope has a 1.2 cm diameter entrance aperture and a 2.6 cm focal length. The 0.5° wide slit located in the telescope focal plane is mounted with its edges parallel to the rotating shaft. A Texas Instrument photoconductive silicon diode senses a solar image scanning across the telescope slit each rotation of the spacecraft. This signal is used to gate the output pulses of the spectrophotometer photomultiplier into a counter for the image scan duration of 5.5 milliseconds. Then the rotary positions of the Bendix indexing head, which has a repeatable positional accuracy of ± 5 arc-minutes, and of the diffraction grating are read out. After 80 rotations of the satellite, the azimuth shaft rotates one degree, and on the next 80 rotations, the adjacent 1.0×10^{-44} element of the sky will be observed. This scanning mode permits 8 looks per degree of azimuth for each of the ten adjacent bandpasses of the spectrophotometer. More than one look is required to build statistical accuracy of the data.

With a primary mirror area of 128 cm², the above scan rate results in a statistical error in the output of 2 percent for an input flux at 2537 Å of 4×10^{-9} erg/sec cm² Å. A flux of 8×10^{-11} erg/sec cm² Å may be determined with a 10 percent error, making a dynamic range useful in the observations of stars 5 stellar magnitudes fainter than the brighter ones.

It was found that repetitive pulses of ultraviolet radiation, of about the same spectral intensity as predicted by Green ⁵ for the Earth's albedo, saturated the photomultiplier tube causing a cumulative increase in the dark current level. Recovery times to dark current levels which were much longer than one rotational period of the satellite were most effectively reduced by reversing the bias between the photocathode and first dynode.

Sensing of the Earth's albedo is accomplished by employing a telescope optically identical to that used in the azimuth sensor. The telescope slit was rotated so that an 8° field is viewed in a plane parallel to the azimuth plane of the spectrophotometer. Before the spectrophotometer is to view excessive albedo flux levels, a sensing circuit

attached to the albedo telescope reverses the first stage bias for the scanning period. This technique is adequate to preserve the sensitivity of the experiment more than 95 percent of the time.

The gate width and/or photomultiplier gain may be checked in orbit by Cerenkov radiation induced in the field lens and photomultiplier window by a 5 microcurie strontium 90 source. This source is located on a lever arm which upon command pivots in front of the exit slit.

REFERENCES

1. G. HASS, Personal Communication.
2. E. CHIN, Personal Communication.
3. A. CODE, *Astr. J.* 65, 278 (1960).
4. D. OSTERBROCK, *Planet. Space Sci.* 11, 621 (1963).
5. A. GREEN, *Applied Optics* 3, 203 (1964).

HORIZON SENSING FOR ATTITUDE DETERMINATION*

BARBARA K. LUNDE

This report discusses the horizon scanners used in determining the angles between the direction to the center of a body in space and the axes of a spacecraft. Several satellites have used various types of horizon sensors for attitude determination. For example, Tiros satellites and Explorer XVII (1963 9A) used passive scan; Mercury capsules have used conical scan. It is planned that the Orbital Geophysical Observatory, Nimbus, Gemini, and Apollo programs will include horizon scanners. A scanner accuracy within 1-2 degrees has been obtained for earth scanners. These errors were caused by the radiation pattern of the earth. Data on that pattern for various wave bands are included.

INTRODUCTION

For attitude determination it is often necessary to determine the angles between the direction to the center of a body and the axes of a vehicle. The direction to the center of the other body may be found by calculating the direction perpendicular to the plane of its horizon. In order to determine this plane from a space vehicle, the angles between the direction of the discontinuity between space and the body of interest and the axes of the vehicle are determined at several points on the horizon.

PHYSICAL PHENOMENA WHICH MAY BE USED TO DEFINE THE HORIZON

Several physical phenomena make possible a variety of methods of sensing the discontinuity between space and the body of interest—in other words, the horizon.

Infrared Radiation from the Earth

If the earth is the body of interest, one physical phenomenon which differentiates it from space is the infrared radiation it emits. The earth's horizon may be defined as the sharp gradient of infrared radiation which exists at the limb, or border, between it and outer space.

Since the earth has a fairly uniform temperature this gradient may be used for space navigation

under a wide variety of circumstances, whether or not the limb is illuminated by the sun. In other words, it may be used under both day and night conditions. Figure 1 shows the output from an infrared sensor which was mounted in Tiros III (1961 $\rho 1$). As the satellite rotated, the field of view of the sensor scanned the sky, the earth, then the sky again. As the sensor's field of view passed over the horizon, it detected the

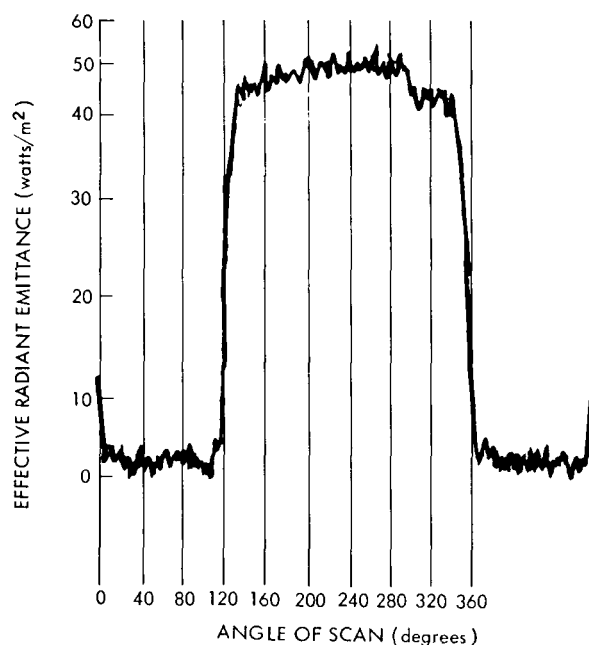


FIGURE 1.—Scan from Tiros radiometer for the spectral range from 8-12 μ .

*Published as *NASA Technical Memorandum X-956*. This report was presented at the Goddard Memorial Symposium of the American Astronautical Society, Washington, D.C., March 16 and 17, 1962.

sharp change in the level of infrared radiation which is represented by the sharp rise in the signal. Passing over the opposite horizon of the earth, it detected the sharp decrease in the level of infrared radiation. It should be emphasized that this figure represents fairly ideal conditions. The sharp gradient of infrared radiation which exists at the boundary between the earth and space is the physical phenomenon most commonly used at present for sensing the horizon in space navigation.

Albedo from the Earth

The ratio between the light received and the light reflected by a body in space such as the earth, sometimes referred to as the albedo, also has been used for horizon determination. However, it is limited in application. At night, when the earth is eclipsing the sun with respect to a space vehicle, no reflected sunlight appears. At other times the earth appears to have phases like those of the moon, complicating the calculation of the horizon plane by means of the albedo. Figure 2, a photograph taken during one of the Mercury suborbital flights, illustrates the phenomenon of albedo. The horizon is easily seen as the gradient between the apparent surface of the earth, which reflects sunlight, and black space beyond, which does not. (The Georgia-Florida coast line is visible in the picture.)

Air Glow Around the Earth

The upper portion of the earth's atmosphere radiates because of excitation by the sun. This phenomenon, known as air glow, could be used to define the horizon. It occurs in both the day and night.

The total visible intensity of air glow is approximately equal to the total amount of starlight seen on a clear, moonless night. Air glow is rarely observed on the surface of the earth because it is evenly distributed throughout the sky. Much of the illumination by air glow is in a few spectral lines. For the sodium line at 5893A, one of the more intense and better known lines, the intensity (relative magnitude) in the summer is 2000 during the day and 230 at night. In winter it is 15,000 during the day and 200 at night. Its intensity during the day is about 75 times the nighttime intensity. Auroral intensity is 1000.

This air glow radiation comes from a rather narrow band in the atmosphere occurring approximately 90 km above the surface in a layer about 10 km thick. Some scientists think that if the earth were viewed from a distance, with a sensor which is sensitive to only one particular band of air glow, a doughnut of radiation would be seen circling a dark earth and the rest of space would be dark, except for point sources. The earth's horizon then could be determined by measuring the physical position of the air glow band. Very few measurements have been made of the air glow phenomenon.

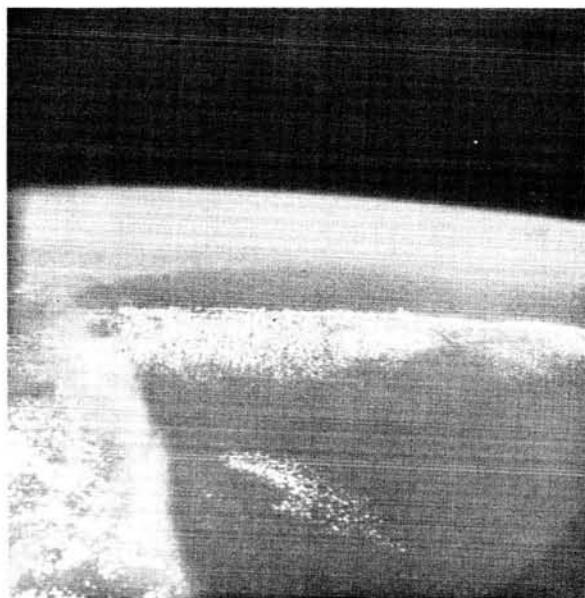


FIGURE 2.—Photograph taken from a Mercury capsule on a suborbital flight.

Some of the most recent data indicate that air glow intensity varies from point to point over the earth's surface. Its altitude appears to vary with time; altitude measurements have been taken only at White Sands, New Mexico. The picture taken during the suborbital Mercury flight (Figure 2) shows no evidence of the air glow phenomenon. The intensity of air glow in the ultraviolet is only about one ten-billionth of the thermal energy radiated by the earth. This illustrates one of the main limitations on the usefulness of air glow for space navigation purposes.

Other Bodies

The horizons of bodies other than the earth may be defined by use of the general phenomena already discussed:

1. Emitted thermal radiation;
2. Reflected radiation (albedo);
3. Air glow.

The horizons of the nearby planets, Venus and Mars, and perhaps artificial satellites, may be determined by sensing the gradient between the infrared radiation which they emit and that of space.

The visible radiation emitted by the photosphere of the sun may be used to determine its horizon. The horizons of the moon, the near planets, and perhaps artificial satellites may be determined by the gradient of reflected sunlight at certain times.

SCANNING SCHEMES

Any instrument for altitude determination needs, in addition to an intensity sensor, a means for determining the direction of the incident radiation. This means will be called a scanning scheme. Five scanning schemes have been used or proposed for determining the direction of the local vertical.

Passive Scan

The passive scan can be used only on rotating spacecraft. A sensor with a small field of view is mounted in the vehicle, its field of view perpendicular to the spin axis of the vehicle. As the vehicle rotates, the field of view of the sensor is swept through space.

The data in Figure 1 were taken with a scanning scheme similar to this. The times at which the field of view of the scanner passes the two horizons can be determined with the use of these data. From these time parameters the attitude of the vehicle at any instant can be determined. This technique was employed with infrared sensors on several Tiros weather satellites, and was used on Explorer XVII (1963 9A). It has also been employed on daytime rocket shots using sensors which detected the earth's albedo.

Conical Scan

A second type of scanning method is the conical scan. Figure 3 shows the Mercury space capsule

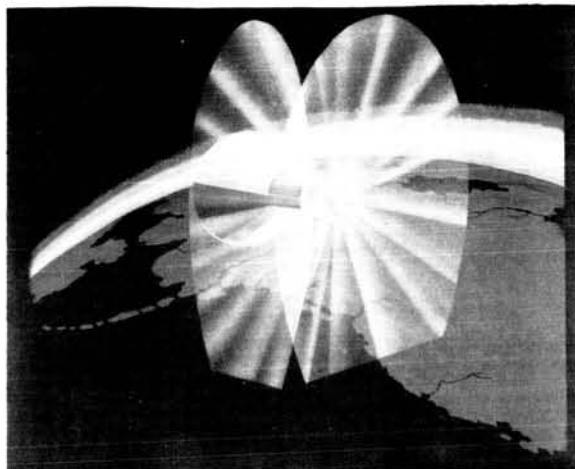


FIGURE 3.—Conical scanning pattern of the horizon scanners on the Mercury capsule.

with cones representing the fields of view which its horizon scanner sensors sweep out in space. The Mercury capsule has two scanners which are used to determine pitch and roll errors. Each scanner, again, has a small field of view, but the rotation of the field of view is done by the scanner itself. It scans a cone out of the space. This cone intersects the earth, and the output of the sensor is used to determine the location of the gradient of infrared radiation between space and the earth. The local vertical can then be computed. It is presently planned that this type of scan will be used in Nimbus, the advanced weather satellite. The scan pattern is inefficient from an information theory point of view. The information on the direction of the horizon is only a very small part of the information which enters the sensor.

Linear Scan

Figure 4 shows the Orbiting Geophysical Observatory and the linear scan pattern that the fields of view of its horizon scanner sensors sweep out in space. In this case, there are four sensors on the vehicle and they each have small fields of view. These fields of view are rotated in a plane from space until they intersect the edge of the earth. Thereafter, they oscillate at the edge of the earth. Once they have locked on the edge of the earth, the angle between the direction of the fields of view and the satellite can be determined. From this information the direction of the local vertical can be calculated. This type of scanning

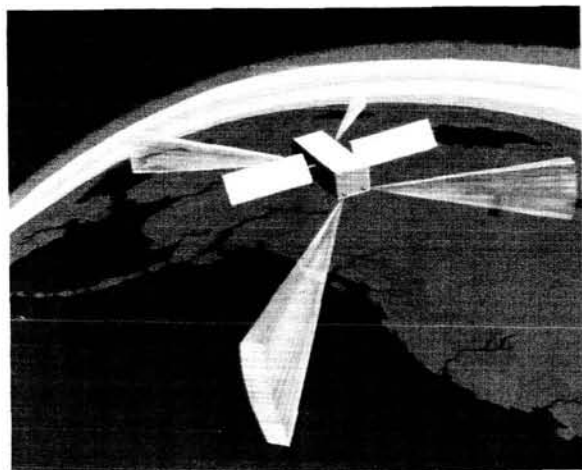


FIGURE 4.—Linear scanning pattern of the horizon scanners on the Orbiting Geophysical Observatory.

is better from an information theory point of view than the conical type scan because the sensor field of view spends a longer time at the edge of the earth. Only infrared sensors have been used with these two methods of scanning.

Nutating Scan

Nutating scan is similar to linear scan but uses only one sensor. The field of view of the sensor is again small and is initially on space. It sweeps down until it intersects the horizon. Then it oscillates about the horizon. Meanwhile the whole scanning head is rotated about an axis perpendicular to the field of view. Thus the projection of the sensor field of view on the earth is a sinusoidal pattern along the earth's horizon. The direction of the local vertical is then determined by calculations using the position of the field of view of the sensor with respect to the vehicle at each instant.

Illumination Comparison

A fifth type of horizon sensing does not use moving parts. The illumination on the scanner from one side of the earth is compared with that from the other side of the earth. This information is used to determine the local vertical. On a vehicle within 1000 miles of the earth, it is difficult to image the entire earth on the scanner since such an imaging system would require a very wide field of view. A scanner which uses

the illumination comparison method of horizon sensing has been proposed which solves this problem. It uses an infrared sensor and a unique imaging system. A scanner which employs this method has been used on the Ranger moon vehicle. It was designed to determine the direction to the earth from distances greater than 80,000 miles using sensors sensitive to the visible light of the earth's albedo.

PROGRAMS WHICH USE HORIZON SCANNING

Programs which have used horizon sensors for attitude determination include spinning satellites and rockets such as Explorer XVII and the Tiros satellites. The Mercury program has used horizon scanners; they have also been used to monitor the performance of some NASA rockets and in the Air Force Agena rocket and satellite system. The Jet Propulsion Laboratory has used an earth sensor in its Ranger vehicle.

Future programs planning to use earth horizon scanners are the Orbiting Geophysical Observatory, the Nimbus advanced weather satellite, the Gemini manned vehicle, the Apollo manned vehicle, classified programs for the Air Force, and the Saturn rocket testing program.

Future programs for which earth horizon scanners are under consideration include various communication satellites, rocket probes, and probes to Venus and Mars. Horizon scanners would determine the direction between local vertical of those planets and the vehicle axes. Another future use for horizon scanners may be in the Advanced Orbiting Solar Observatory. This satellite will orbit the earth but observe the sun. An earth horizon scanner could be used to determine the direction to the center of the earth. A sun horizon scanner or limb sensor could be used to determine precisely the direction to the center of the solar disk. A moon horizon scanner may be used on lunar missions such as Apollo.

HORIZON SCANNING ACCURACY

The accuracy of any determination has two basic limitations. One source of inaccuracy is errors developed in the instrument making the determination and the other source is variability in the physical phenomenon utilized for the determination.

Instrument Error

In the case of horizon scanners, sources of instrument error are detector noise, mechanical tolerances, and mounting rigidity. The error caused by a certain amount of detector noise varies, depending on the type of pattern employed. For example, conical scan is more susceptible to detector noise than is linear scan. Detector noise can be made to have as small a contribution to the error of the attitude determination as desired by using large optics, sophisticated scanning methods, and better detectors. Mechanical tolerances in scanners can be made better than a few seconds of arc, if necessary, by good mechanical design. In other words, the error in determining the local vertical due to mechanical tolerances of the scanner can be made smaller than a few seconds of arc. Mounting rigidity can be a problem in some space vehicles. Errors as high as 1 degree can develop if care is not taken. However, alignment of an experiment with a heavenly body can be within 0.1 second of arc if the radiation coming into the experiment is shared with the attitude sensor.

Variability in Physical Phenomena

Earth

The other class of phenomena which cause errors in detecting the local vertical are those due to the heavenly body in question. Errors dependent on the physics of the body generally are a certain magnitude of the apparent diameter of the body and decrease in absolute magnitude as the scanner gets farther away from the body, because of the apparent decrease in angular diameter. The maximum error due to the ellipticity or oblateness of the surface of the earth, in determining the direction of the horizon at one point, is not greater than 0.2 degree. This error could be made zero if the ellipticity of the earth could be taken into account in calculating the attitude. Much larger errors can arise from irregularities in the apparent surface of the earth, such as those that might be caused by high clouds. These may be as large as 0.25 degree at an altitude of 200 naut. mi. but may be mitigated somewhat by the refraction by the atmosphere of the radiation from the tops of the clouds.

Another source of error from some types of scan is cold clouds appearing on the face of the

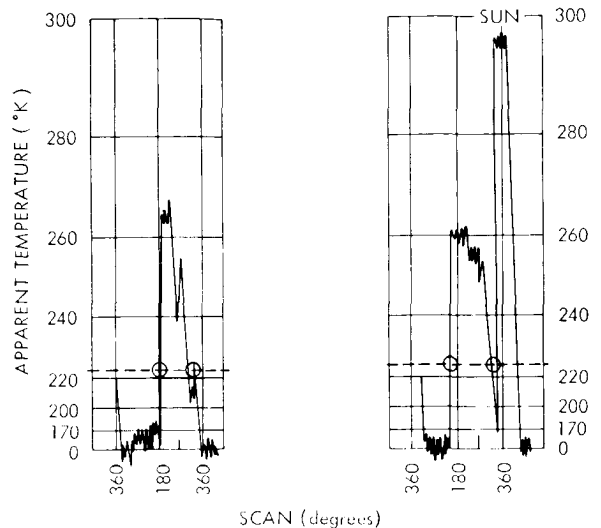


FIGURE 5.—Samples of the horizon sensor signals from the Mercury scanner for the spectral range of 2-15 μ .

earth. Figure 5 shows recordings from sensors in the horizon scanners in one of the Mercury flights, MA-5 (1961 α 1) (see Reference 1). If the output of the radiometer is above the dashed line, the presence of the earth in the field of view is indicated. The horizon is defined by the crossing of this line. In the scan on the left the field of view of the sensor passed over a very cold cloud on the face of the earth. The temperature of the cloud appeared to be about 210°K. At various times during this flight clouds colder than 200°K were observed. This type of cloud emits only about $\frac{1}{2}$ the total radiation emitted by an average area of the earth and only $\frac{1}{4}$ of the radiation with wavelengths between 8 and 12 μ that an average area of the earth radiates. These clouds could cause a great error in the attitude determined by any scanner whose field of view crossed them. The scanner might confuse the cloud with the horizon of the earth. There are scanning methods which are not bothered by the appearance of this type of cloud on the face of the earth. However, they would be bothered by such a cloud appearing at the horizon. In this case the cloud would have a smaller apparent diameter. However, its radiation would be augmented by the atmosphere through which it would be viewed.

Dr. Rudolph Hanel has suggested that by using a scanner sensitive only to wavelengths around 15 μ , which are strongly absorbed by carbon

dioxide, the phenomena of cold clouds could be avoided. Radiations at this wavelength should come from very high in the atmosphere at the top of the carbon dioxide region, above the level of cold clouds (References 2 and 3).

The trace on the right in Figure 5 was taken at a time when the sensor field of view passed near the sun. The signal from the sun appears to the right on the picture. The amplitude of the signal from the sun is clearly stronger than that of the earth, but the scanner could view the onset of the signal from the sun as the horizon of the earth and an attitude determination from this data could be in error by many degrees. It is interesting to note from this data that the apparent diameter of the sun on the trace is about 73 degrees which is much larger, of course, than the apparent diameter of the visible sun in the sky, 0.5 degree.

Several Tiros satellites carried a 5 channel radiometer which senses radiation from the earth in 5 different spectral bands. Figure 6 shows an example of the signals from the radiometer. The first channel is from a radiometer sensitive to radiation of 6.5μ wavelengths. This type of radiation is strongly absorbed by water vapor, so the signal in this channel is from the top of the water vapor in the atmosphere. The signal from this channel shows severe limb darkening. In other words, to the sensor, as it scans over the earth, the wave's shape appears sometimes to be almost triangular rather than square and there is a small gradient of radiation at the horizon. Channel 2 is the signal from the radiometer sensitive to radiation of wavelengths of 8 to 12μ . This is considered to be the water vapor window and the radiation recorded here probably came from near the surface of the earth or from ex-

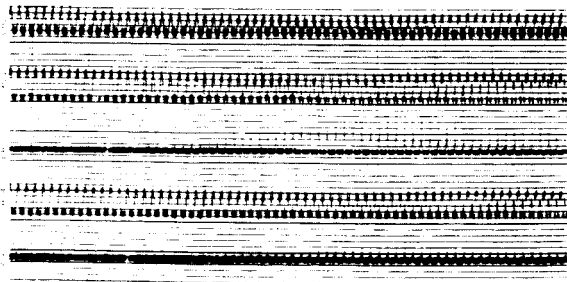


FIGURE 6.—Sample of Tiros radiometer data.

tremely opaque clouds. Figure 1 is actually an enlargement of the signal from this channel. Channels 3 and 5 were sensitive, chiefly, to visible light and were used as a check on the TV camera in the satellite. Channel 4 was sensitive to radiation wavelengths from 8 to 30μ . The signal in this channel is very similar to that of channel 2 which was sensitive to 8 to 12μ radiation. Figure 7 shows an enlarged view of a signal from channel 2. The signals from the Tiros 5 channel radiometer show the cold clouds and sharp gradients on the surface of the earth which also appear in the signals from the Mercury horizon scanner shown earlier (Reference 4).

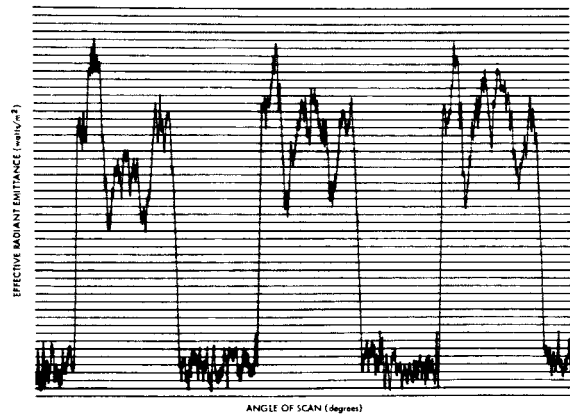


FIGURE 7.—Scans from Tiros radiometer for the spectral range of 8- 12μ .

The angular width of the signal from the earth was measured for 34 consecutive scans on channel 2. A level of about $1/5$ the maximum signal level was chosen as the "slicing level." The diameter of the earth appeared to vary from scan to scan. The rms variation was 1.8 angular degrees; it can be completely attributed to system noise. The variation in apparent diameter could cause attitude errors of this approximate magnitude.

Other Bodies

Mars would look much like the earth to an infrared horizon scanner. It probably has only a few water clouds but may have high, opaque dust clouds. In fact, one interesting and as yet unexplained fact about Mars is that its ellipticity, or oblateness, as calculated from the orbits of its satellites is $1/2$ that measured on photographs of

the planet. A reasonable hypothesis which explains this anomaly is that the solid surface has the ellipticity predicted by the motion of the satellites, and the apparent equatorial bulge is caused by dust clouds. Mars' actual ellipticity is somewhat less than the earth's but its *apparent* ellipticity is almost twice the earth's.

This or a similar phenomenon could also occur on the earth, causing errors in horizon scanners. The earth's apparent shape when it is viewed with radiation of certain wavelengths, and the consistency of that shape with time, have not been determined experimentally.

In the past, the attitude of a spacecraft with respect to the sun has been determined by sensors which use the illumination from the total solar disk. However, sunspots can have an area equal to 1/300 of the total disk of the sun. A sensor which compares radiation from the two halves of the solar disk could thus be in error by 5 seconds of arc. For a precise determination of the direction of the sun, a method of limb tracking or horizon sensing of the sun must be used.

Experimentally Determined Accuracy

Tiros satellites have also carried another infrared sensitive horizon scanner called the horizon pippier. The attitude of the satellite was calculated by reducing the horizon pippier data and also from the horizons and other landmarks appearing on the TV pictures from the satellite.

If the best data from the scanner are hand selected and smoothed with a computer, the attitude determined from these data agrees with that determined from the TV pictures within 1 or 2 degrees. The attitude determined from the TV pictures is considered more accurate.

An accuracy of 1 degree of arc was specified for the horizon scanner on the Mercury capsule. Figure 5, recorded from the Mercury scanner, indicates that the scanner failed to remain within the specified 1 degree error at certain times.

CONCLUSION

In conclusion, horizon scanners for attitude determination, up to now, have been limited by instrument design rather than by errors due to the body which they have scanned. The accuracy of scanners which have been flown has been, at best, 1 or 2 degrees. It appears that an accuracy of 0.2 degree for altitudes below 1000 miles and 0.1 degree for altitudes above 1000 miles will be attainable by techniques now in development.

The use of physical phenomena other than the total infrared radiation from the earth, such as air glow or narrow spectral band infrared radiation, may be necessary to attain better accuracy. More information must be known on these phenomena before horizon scanners can be designed to use them.

REFERENCES

1. "Report on Analysis and Correlation of Data from MA-5 Mission," Barnes Engineering Co., Report 4858-SR1, August 20, 1962.
2. HANEL, R. A., BANDEEN, W. R., and CONRATH, B. J., "The Infrared Horizon of the Planet Earth," *J. Atmos. Sci.* 20(2):73-86, March 1963.
3. BURN, J. W., "The Application of the Spectral and Spatial Characteristics of the Earth's Infrared Horizon to Horizon Scanners," in: IEEE Transactions on Aerospace-Support Conference Proceedings, August 4-9, 1963, pp. 1115-1126.
4. ALTSHULER, T. L., "Infrared Transmission and Background Radiation by Clear Atmosphere," General Electric Co., Report TIS 61 SD 199.

RESPONSE OF MODIFIED REDHEAD MAGNETRON AND BAYARD-ALPERT VACUUM GAUGES ABOARD EXPLORER XVII*

G. P. NEWTON, D. T. PELZ, G. E. MILLER, LTJG, USN, AND R. HOROWITZ

Two Redhead and two Bayard-Alpert gauges were used in the satellite Explorer XVII to measure the density, pressure, and temperature of the earth's upper atmosphere. Gauge pressures from 10^{-6} to 10^{-10} torr (a function of satellite attitude and velocity, and of certain atmospheric parameters) were measured over the altitude range from 260 to 900 kilometers. During the spin period of 0.67 second, the gauges recorded pressure changes as large as a factor of 30 (e.g. from 3×10^{-9} to 8×10^{-8} torr). Comparison of the independent gauge outputs yielded values in good agreement with kinetic theory predictions.

INTRODUCTION

Knowledge of the neutral particle concentrations is important for an understanding of the physical and chemical processes occurring in the upper atmosphere. The Explorer XVII satellite (1963 9A), known prior to launch as "S-6"†, was placed into orbit to measure directly the neutral particle density, composition, temperature, and the electron density and temperature. The orbit attained had an inclination of 58 degrees, a perigee of 257 km, and an initial apogee of 920 km. The total of the previous direct measurements of the neutral particle atmosphere in the altitude region covered by Explorer XVII was so small that the first two days' output from the satellite significantly increased the amount of available data for study of this region.

The satellite was a pressure-sealed stainless steel sphere 0.875 m in diameter. Its experiment sensors—two Bayard-Alpert gauges (BAG), two Redhead gauges (RHG), two magnetic mass spectrometers and two electrostatic probes—were located as shown in Figure 1. The satellite was spin-stabilized at $1\frac{1}{2}$ cps, about an axis through the two mass spectrometers. The pressure gauges were mounted at the vertices of an equilateral tetrahedron, two of them being located on

the satellite equator. Explorer XVII exceeded its expected 90-day operational lifetime by ten days, and considerable data were obtained from all experiments.

Some of the first observations of the gauge responses are discussed in this paper. As more data are analyzed, a more complete picture of the gauge performance will become available and will be presented in future papers.

THE GAUGE EXPERIMENT SYSTEM

To measure the neutral atmospheric density and temperature two Bayard-Alpert and two magnetron Redhead vacuum gauges were used. The reasons for these choices were:

1. These gauges are accepted as instruments for vacuum measurements.
2. The sensors complement one another and together cover the expected 10^{-6} to 10^{-11} torr pressure range.
3. The two types of detectors provided at least two decades of overlap in pressure range.
4. The gauges are being studied extensively in other laboratories as well as at GSFC.

Figure 2 shows a flight BAG (top) and a flight RHG (bottom). Both gauges have ion traps painted on the inside of the glass envelope, out of

*Published as *NASA Technical Note D-2146*, February 1964.

†Horowitz, R., "S-6, An Aeronomy Satellite," *Advances in the Astronomical Sciences*, Volume 12, 1963. pp. 21 to 39.

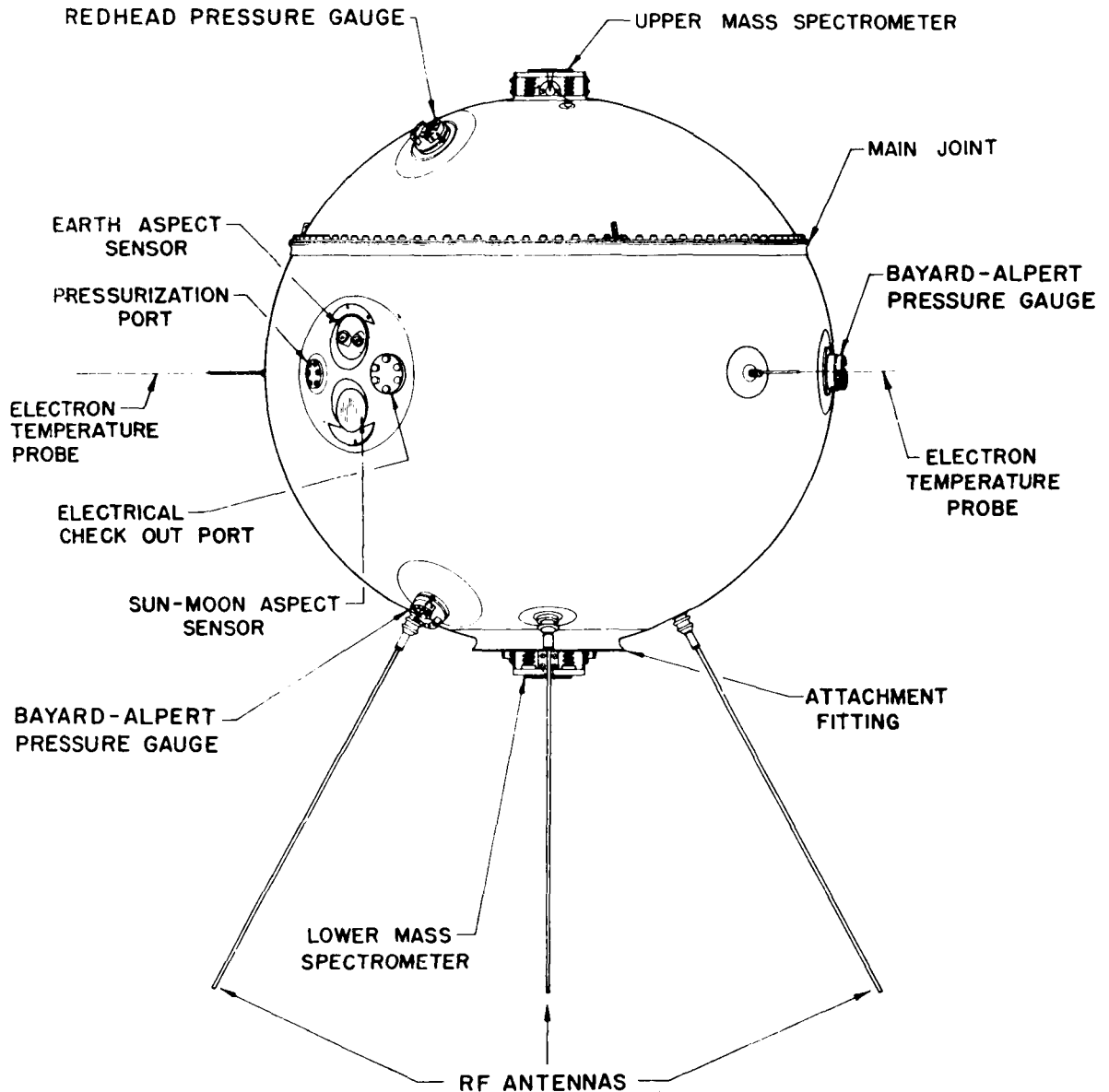


FIGURE 1.—Lateral view of Explorer XVII.

the ionization region, and have a volume of approximately 55 cm^3 .

The Bayard-Alpert gauges were made by Westinghouse Electric Corporation, Electron Tube Division, and have been assigned the Westinghouse number WX 4250. They are 2.5 cm in outside diameter. Each detector has three filaments spaced symmetrically about the gauge axis, two of which were available during flight. If the first filament opened for any reason, the second fila-

ment would be switched automatically into use. However, both Bayard-Alpert gauges survived launch and the satellite operational life of more than 700 operations, on the primary filament.

The flight gauges differ from the normal laboratory Bayard-Alpert gauge in that:

1. The gauges have been miniaturized and ruggedized.
2. Ion traps have been added.

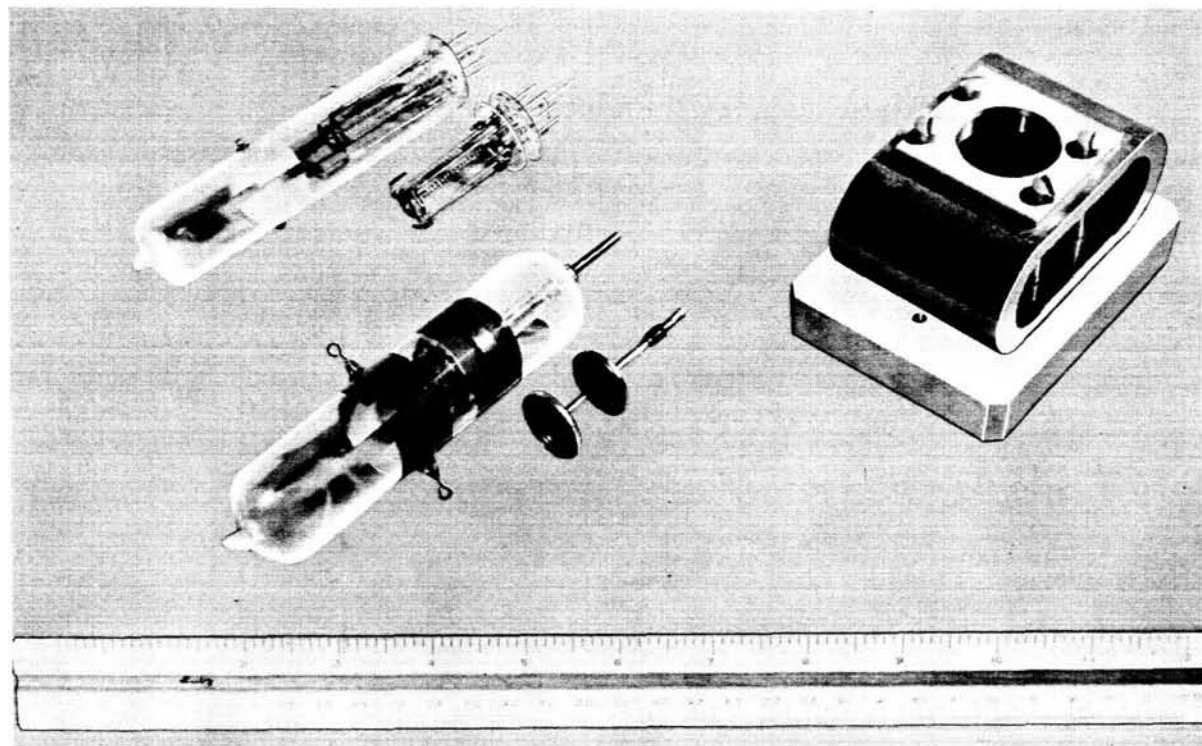


FIGURE 2.—The flight Bayard-Alpert gauge (top), the flight Redhead gauge (bottom), and the Redhead magnet (right).

3. The tapered collector is brought through the same press as the other electrode connectors.
4. Each tube has three filaments.

The Redhead gauges, of 3.0 cm outside diameter, were made by NRC Equipment Corporation and carry the designation NRC 528. The cathode spool is mounted coaxially with the envelope, and a special magnet fits over the tube to provide an axial magnetic field. These detectors differ from the NRC 552 laboratory Redhead gauge in that:

1. The auxiliary cathodes are eliminated, and the tube is miniaturized and ruggedized.
2. The magnetic field is parallel to the gauge axis.
3. The anode is painted on the inside of the envelope.
4. Ion traps have been added.

The four gauge systems, which consisted of components similar to those shown in Figure 3,

were electrically and otherwise independent of one another, and were programmed to operate automatically for 4 minutes when the satellite was commanded on. After turnoff, they were then inoperative until the next spacecraft turn-on.

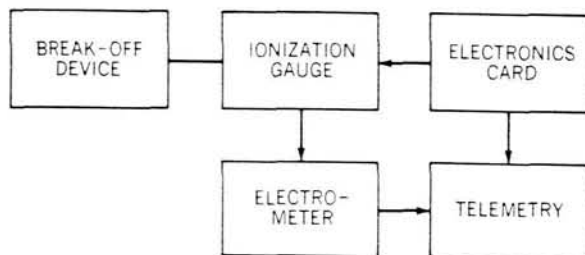


FIGURE 3.—Representation of a pressure gauge experiment.

To maintain vacuum cleanliness, the gauges were sealed under vacuum prior to installation in the satellite. After the satellite was established in orbit, the sensors were opened by break-off devices, which exposed the gauges to the atmosphere through knife-edge kovar orifices 0.938 cm in diameter. These orifices have a calculated molecular conductance of 8.2 liters/second for

nitrogen at room temperature. They protrude from the satellite shell a short distance and have 180 degree acceptance angles.

As is seen in Figure 2, the ion traps are cylindrical electrodes painted on the gauge walls between the ionization region and the orifice. For approximately 90 seconds during each 4 minute operating interval, a potential of 30 v was applied between traps to help evaluate the effect—if any—of entering, charged ambient particles. No significant effect was observed.

The electronics consisted of the power supplies to provide the gauge potentials and, for the Bayard-Alpert gauges, emission-current regulators, which held the grid currents constant to better than 2 percent. Logarithmic electrometers converted the gauge currents to voltages suitable for telemetry.

LABORATORY RESPONSES OF THE GAUGES

The laboratory responses of the gauges were obtained with the detectors energized as they would be during flight. For the Bayard-Alpert gauge the potential difference from filament to grid was 105 v, and from filament to collector, -27 v with the filament ground. The emission current was 5.0 ma.

The Redhead gauge was operated with 4800 v between the anode and cathode, the latter being grounded through the electrometer. The magnetic field strength was 1000 gauss.

Figure 4 shows the output current of a flight Bayard-Alpert and a flight Redhead gauge plotted against nitrogen pressure.* The gauge residuals have been subtracted to obtain these curves. The flight BAG calibration was obtained by comparing its response to that of a Westinghouse 5966 tube operating at 0.1 ma. emission current, which previously had been compared to a Consolidated Vacuum Corporation GM 110 McLeod gauge over the pressure range from 10^{-3} to 10^{-6} torr. The flight BAG is linear from 10^{-9} to 10^{-4} torr and, for nitrogen at the flight emission current, the gauge has a sensitivity of $4.2 \times 10^{-2} \pm 25$ percent ampere/torr. The ratio of the helium sensitivity to that for nitrogen is 0.17. The precision of the measurements was better than 10 percent even

*Hereafter all pressures are equivalent nitrogen pressures unless otherwise specified.

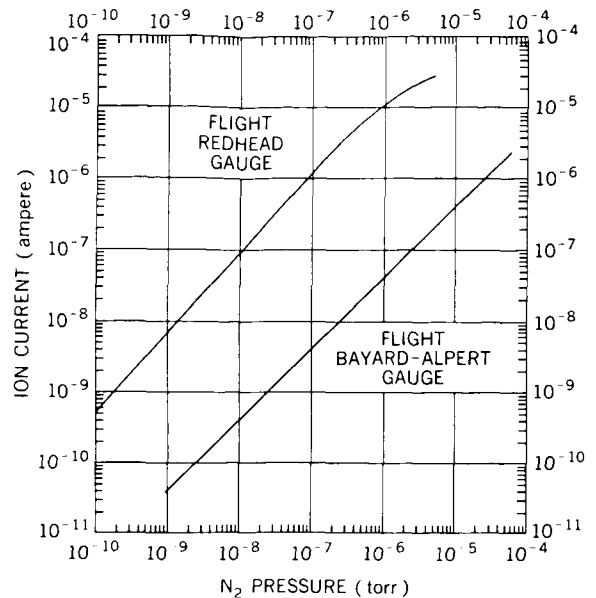


FIGURE 4.—Ion current versus pressure characteristics of the flight Bayard-Alpert and Redhead gauges.

though the data were taken on different days with intervening bake-outs of the system at 400°C and outgassings of the sensors. Most of the tubes calibrated had sensitivities within 25 percent of each other. Normally, the use of different filaments in the same gauge changed the sensitivity by a few percent.

The response of the flight RHG is nonlinear in the pressure region from 2×10^{-10} to 2×10^{-7} torr, and the collector current i_+ follows an expression of the form $i_+ = kp^n$ where n is 1.15, k is a constant and p is the pressure. The calibration data supplied with the gauge was stated as being accurate to ± 30 percent. At Goddard Space Flight Center (GSFC), the gauges were compared to a laboratory NRC 552 Redhead gauge which had been, in turn, compared to the McLeod-calibrated WH 5966, thus providing a calibration believed valid to ± 30 percent. As a result of the comparisons, the NRC and the GSFC calibrations of the flight gauges were found to agree. The best-fit curve through the overlapping region of the two sets of data is used to analyze the flight measurements. The following observations can be made about the Redhead gauge:

1. The assigned gauge sensitivity in amperes/torr for a particular gas varies with pressure. For example, in nitrogen at a pressure

between 10^{-10} and 10^{-9} torr the value is 7, and increases to 11 at the 10^{-7} to 10^{-6} torr range.

2. Preliminary analysis of laboratory measurements indicates that the ratio of the Red-head gauge sensitivity for nitrogen to that for helium at a particular pressure is the same as that of the Bayard-Alpert gauge.
3. The gauge is insensitive to pressure changes in the 10^{-5} torr region.
4. In the 10^{-9} to the 10^{-6} torr range, the gauges would generally strike a discharge in less than 30 seconds. The time to start varied from tube to tube and appeared to be a function of tube cleanliness. There was no starting problem in flight.
5. In certain regions of pressure the gauges are unstable, and the current consists of transients and noise, as well as a dc component. These pressure intervals are well defined and constitute a small part of the gauge response; thus they do not significantly interfere with the analysis of the flight data.

THE SATELLITE MEASUREMENT TECHNIQUE

In considering how the measurements were performed with the satellite, it is helpful to recall that the gauges are moving at a high speed compared to that of the atmospheric molecules.*

*Spencer, N. W. et. al., "On the Use of Ionization Gauge Devices at Very High Altitudes," *Amer. Rocket Soc. J.*, April 1959, pp. 2901—294.

Further, in the equivalent situation with the satellite at rest, the ambient molecules, in effect, are moving with the satellite speed in the opposite direction (as is illustrated in Figure 5, with the gauge facing into the "neutral beam"). In this case the gauge internal pressure will exceed the ambient. The trace on the right in Figure 5 shows the ratio of gauge pressure to ambient pressure as a function of the satellite spin angle, for a satellite speed of approximately 7.5 km/sec., with the gauge temperature equal to the ambient temperature. It is seen that the gauge pressure is a factor of 24.7 above the ambient when the gauge is in position 2. If the two temperatures were not equal, the ordinate would require adjustment by a thermal transpiration term. In position 3, the sensor does not "see" the beam and the pressure in the gauge is determined by the thermal transpiration equation or, for equal temperatures, equals the ambient pressure. In position 4, the pressure is essentially zero, since only particles in the high-energy tail of the thermal velocity distribution can overtake the satellite and enter the sensor. Position 1 is equivalent to position 3. Thus, over a spin cycle we would expect an ideal gauge response to exhibit this behavior. Since the four Explorer XVII gauges are mounted at the vertices of an equilateral tetrahedron, at least one will always experience a "ram" velocity and thus provide a detectable and known sample of the neutral-particle atmosphere.

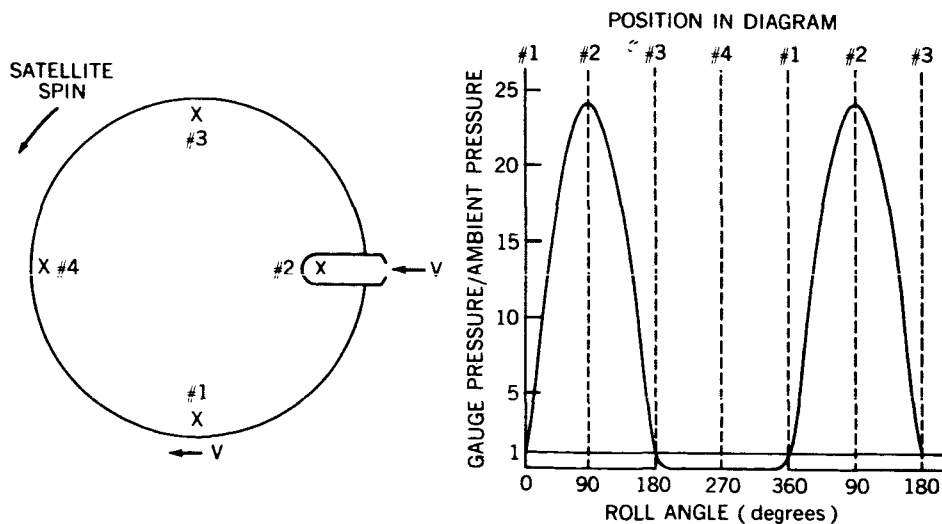


FIGURE 5—Plot of predicted gauge-pressure-to-ambient-pressure ratios versus satellite roll angle.

THE GAUGES' FLIGHT RESPONSE

The satellite was operated during the first apogee pass, and the openings of the gauges were observed. Analysis of the data indicates that the gauges survived the launch vibration, and all indicated pressures after opening in the 10^{-9} torr range or lower, thus demonstrating their ability to indicate this low pressure in orbit.

Because of the satellite's attitude during the first few days in orbit, the operation of the lower BAG could not be immediately evaluated. The present evidence, however, is to the effect that it experienced a loss of sensitivity after opening. Related telemetry information indicates that proper operation should be expected. The apparent contradiction has not been resolved.

Figure 6 is a machine plot of the pressure gauge telemetry output for the orbit 118 pass over Blossom Point, Maryland. The period is 0.67 second corresponding to one rotational period of the satellite. The output of BAG I indicates that the sensor is in the rarefaction. The trace labeled BAG II is the output of the equatorial BAG, RHG I is in turn the equatorial RHG signal and RHG II is the upper RHG response. The spin axis is 75 degrees from the velocity vector. The Redhead gauge output voltages are inverted, that is, a high voltage value corresponds to a low pressure. The peak-to-peak pressure values shown in Figure 6 are proportional to the atmospheric density at the time and location of measurement, and inversely proportional to a function of the satellite velocity. The analysis of the nonequatorial gauge data requires modified velocity considerations to obtain atmospheric density. The density obtained from these three detectors for this pass agree to better than 21 percent and is compatible with the expected densities. The phase relations between the gauges, which can easily be seen in Figure 6, are proper and in accord with the known satellite attitude.

In Figure 7, the equatorial RHG pressure for the Blossom Point pass 118 is compared with its theoretical response as predicted from kinetic-theory considerations. The peak of the theoretical curve has been normalized to fit the maximum of the measured gauge pressure. The measured pressure agrees quite well with the theoretical curve shape. The resolution of the data is such that one pressure gauge output sample is

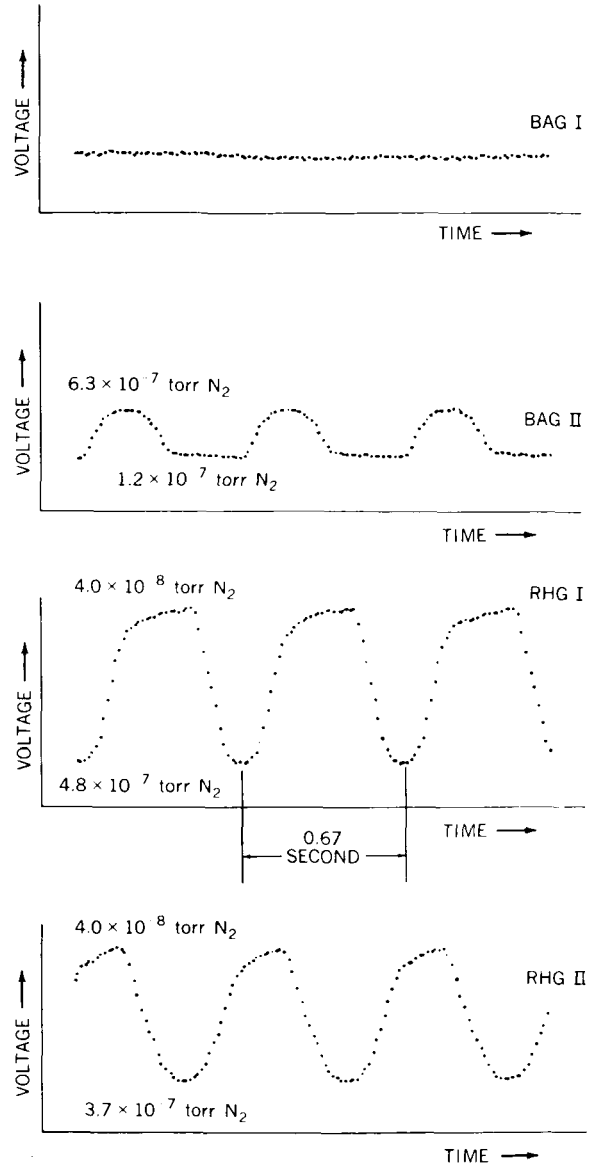


FIGURE 6.—Pressure gauge telemetry output signal plots (Orbit 118, April 10, 1963, at 18:30 E.S.T. over Blossom Point, Maryland).

obtained approximately each 9 degrees of satellite spin angle. It is seen that the measured maximum-to-minimum gauge pressure ratio for a spin cycle is 12, i.e., from 4.0×10^{-8} to 4.8×10^{-7} torr. Other passes have shown ratios as large as 30, e.g., from 3×10^{-9} to 8×10^{-8} torr. The theoretical gauge pressure, when the sensor is behind the satellite is a factor of more than 10^9 below the ambient value (0 and 180 degrees roll angle). The gauge, in this case, does not follow completely the

theoretical response into the rarefaction region, but apparently has some effects of residual outgassing. Therefore we have the opportunity to observe and investigate the outgassing-residual of the sensors and its implications, each spin cycle. With respect to data analysis, the outgassing residual can only affect the density to the order of 10 percent in this case, since the minimum gauge pressure is less than 1/10 the maximum pressure. This same RHG has indicated gauge pressures at perigee in the 10^{-6} torr range and both RHGs have indicated gauge pressures at apogee in the 10^{-10} torr range which are spin modulated. When the ram velocity is less than that experienced during Blossom Point pass 118, which is a near-maximum ram velocity condition, the gauge response follows the theoretical response to lower pressures, and thus the gauge pressure at the 180 degree point is not dominated by the outgassing-residual.

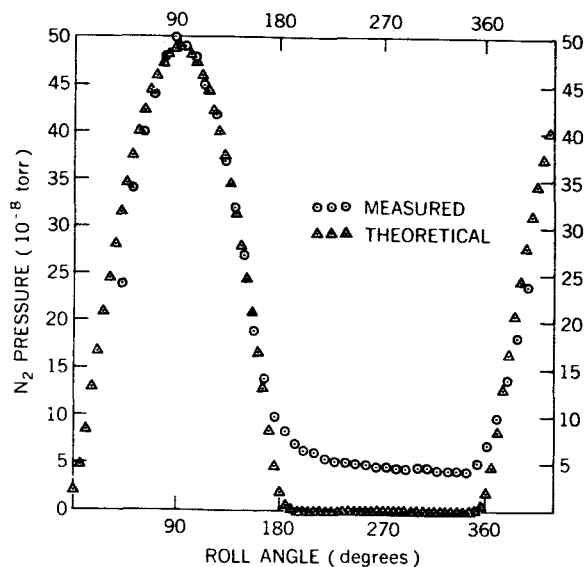


FIGURE 7.—Plot of theoretical and measured equatorial Redhead gauge pressure versus satellite roll angle.

At apogee, the equatorial BAG indicates a 10^{-5} torr pressure, which appears, in comparison with the RHG at apogee, to be due to residual gas in the BAG. The BAG as launched indicated a considerably higher pressure, i.e., in the 10^{-5} torr range, than the other sealed gauges, which were in the 10^{-9} torr range. It is not presently known whether this higher pressure contributed to the observed gauge residual.

As analysis of the data proceeds, atmospheric densities and temperatures, and in addition information regarding the following, will be obtained:

1. The behavior of the gauge residuals, both long term (day-to-day) and short term (orbit-to-orbit);
2. Gauge responses to particular satellite orientations and ram velocities;
3. Apparent outgassing rates of the gauges and their variations;
4. The effect, if any, of sustained exposure of the gauges to an atomic oxygen environment.

Sufficient data have been studied already, however, to show:

1. Very good agreement of the densities obtained from the two RHGs and the BAG, presently estimated to be generally better than 30 percent;
2. Agreement of measured densities with values obtained from theory and indirect measurements;
3. The ability of the gauges to follow large (several orders of magnitude) pressure variations, as experienced in an orbital period;
4. The ability of the gauges to follow large (factors of 12 to 30) pressure changes in a period of 0.670 second;
5. The agreement between the shape and phase of the gauges' responses to those expected from kinetic theory.

It is concluded that these gauges have made meaningful *in situ* direct measurements which can be interpreted easily in terms of the neutral atmospheric parameters—density, temperature, and pressure.

ACKNOWLEDGMENTS

The authors wish to thank Mr. H. Benton, Mr. D. Williams, Mr. L. Clark, and Mr. J. Bush for their tender loving care of the flight equipment from inception to launch.

AN APPLICATION OF PRESSURE DIVISION FOR CALIBRATING IONIZATION GAUGES

JAMES R. POYNTER AND SACHIO SAITO

INTRODUCTION

The pioneer calibration of ionization gauges has been made by Nottingham and Torney¹ and others using various methods, based upon the McLeod Gauge in most instances. Due to the limitations of the McLeod Gauge and also the difficulty in operation, a different approach to the calibration of vacuum gauges was sought.

A method of generating a pressure of known value was described by Knudsen,^{2,3} as a method of standardizing her radiometer gauge. A comprehensive study of Knudsen's method was recently conducted by Schuhmann.⁴ Both Knudsen's method and Schuhmann's study were based on the application of Boyle's law, namely, that at constant temperature, the pressure of a gas varies directly as the density, or inversely as the volume.

The method of ionization gauge calibration herein described is an application of a technique utilizing the research of Knudsen and Schuhmann. This method will serve as a link between known pressures and higher vacuum ranges. The contribution of the pressure-division experiment within the ART Program is that it defines the feasibility, and shows the limitations of the method in vacuum calibration technology.

BACKGROUND

The vacuum gauge calibration described here uses the pressure-division method, one that is accurate and directly useful. However, one must realize that the term "vacuum gauge calibration" encompasses a vast technology of complex measurements. The vacuum spectrum is comparable to the electro-magnetic spectrum in the number of orders of magnitude which it spans. Any single calibration system can accommodate only a few decades of measurement without having to appreciably change its form.

The pressure-division system covers the lower range of the ionization gauge pressure span, namely, 10^{-3} to 10^{-6} torr. This system is more accurate than the presently used McLeod Gauge which also operates in this range. The linearity of an ionization gauge can be verified in this range, thereby establishing the basis for extrapolating the gauge to higher vacuums. It is also the range in which many environmental and component test chambers are operated.

The basic approach to vacuum gauge calibration is similar to the calibration of other transducers, that is, establishing that the gauge undergoing calibration performs properly under reproducible conditions as specified by the gauge manufacturer. By using dry nitrogen gas, problems of the behavior of the gauge with different gases or combinations of gases are avoided. Added complexity would only serve to confuse the basic calibration measurement.

To clarify the position of the pressure-division system in the chain of vacuum calibration, a brief description of using this system follows. The pressure-division apparatus is in effect a generator of known pressures. A reference ionization gauge is calibrated in conjunction with it. This calibration follows a straightforward procedure but is time consuming. Therefore, the pressure-division system is reserved for the purpose of periodically re-calibrating the reference gauge and serving as the reference standard.

THEORY OF PRESSURE DIVISION

A quantity of a selected gas is injected into a comparatively small volume under a pressure high enough to be accurately measured by a displacement manometer and a cathetometer with a precision of 0.02 mm. This gas of a known pressure and known volume is then allowed to expand into

an expansion reservoir of a much larger volume. The resultant pressure can then be calculated by the application of Boyle's law if an isothermal environment is maintained.

By repeating this gas expansion cycle, with various initial pressures and volumes expanding into expansion reservoirs of different volumes, and by cyclic combinations, a significant number of calculations can be made. If an ionization gauge that is connected to the expansion reservoir is activated; and a gauge reading is made before the gas expansion and again after expansion; and these readings are compared with the calculations obtained through the application of Boyle's law; then, the accuracy and linearity of the ionization gauge can be established within the range of pressures developed.

DESCRIPTION OF APPARATUS

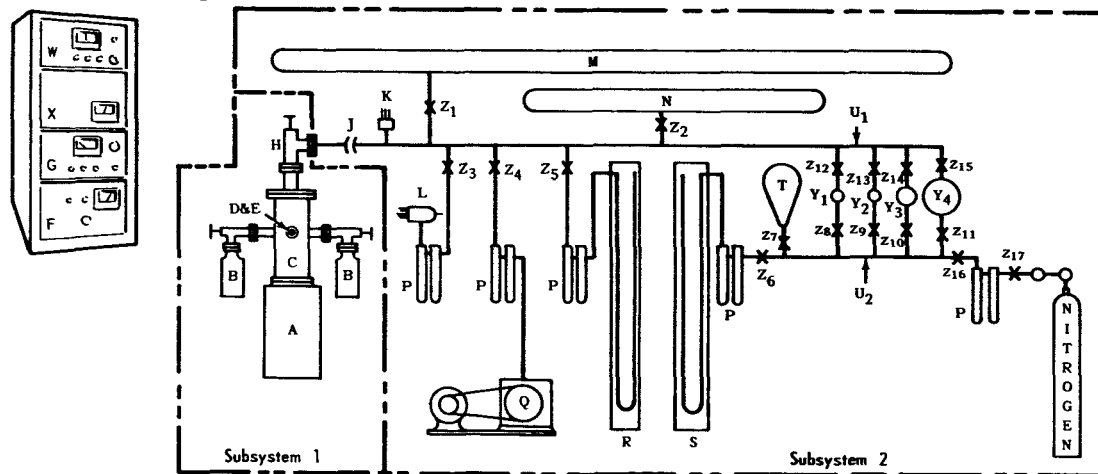
The pressure-division experiment was conducted using the apparatus shown schematically in Figure 1 and pictorially in Figures 2 and 3. The apparatus is divided into two subsystems.

Subsystem 1 is the ultra-high vacuum system and is used as a pumping and metering station for lowering the pressure to the required working vacuum of the experiment. Subsystem 2 is the high vacuum system and is used to expand the various generated pressures and to establish the calibration of a reference gauge by calculations based on Boyle's law. Subsystem 2 is fabricated of Pyrex glass components. It is to be noted that all cold traps are the double symmetrical glass type designed by P. R. Yeater;⁵ the refrigerant used is liquid nitrogen. All glass stopcocks are high vacuum type and lubricated with Apiezon "L" grease.

CALIBRATION OF APPARATUS

The volumes of controlled pressure reservoirs Y_1 , Y_2 , Y_3 and Y_4 were calibrated by mass determination prior to assembly into the system. These four reservoirs were cleaned, weighed, filled with laboratory grade triple-distilled mercury, and re-weighed. The volume of the calibration chamber T was determined by filling it with water

ELECTRONIC CONTROL PANEL



LEGEND

SUBSYSTEM 1

- A. 140 l/sec VacIon Pump
- B. VacSorb Forepump
- C. Ultra-High Vacuum Sump
- D. Nude Ionization Gauge
- E. Hastings Thermocouple Gauge Tube
- F. VacIon Pump Control Unit
- G. Ionization Gauge Control Unit (D)
- H. Kane 1" Ultra-High Vacuum Indium Seat Valve

SUBSYSTEM 2

- J. High Vacuum Glass Ball Joint
- K. Hastings Thermocouple Gauge Tube
- L. Bayard-Alpert Type Ionization Gauge
- M&N. Expansion Reservoirs
- P. Cold Traps
- Q. Mechanical Forepump
- R&S. U-Tube Mercury-filled Manometers
- T. Calibration Chamber

- U₁. 1" Manifold
- U₂. ½" Manifold
- W. Ionization Gauge Control Unit (L)
- X. Hastings Vacuum Gauge
- Y₁ through Y₄. Controlled Pressure Reservoirs
- Z₁ through Z₁₇. High Vacuum Stopcocks

FIGURE 1.—Schematic of pressure division apparatus.



FIGURE 2.—Pressure division calibration system.

from a calibrated burette. The results of these volume determinations are:

Volume of controlled pressure
reservoir $Y_2 = 0.8654$ ml

Volume of controlled pressure
reservoir $Y_1 = 1.2716$ ml

Volume of controlled pressure
reservoir $Y_3 = 10.683$ ml

Volume of controlled pressure
reservoir $Y_4 = 97.952$ ml

Volume of calibration chamber $T = 2331.1$ ml

The volume of the $\frac{1}{2}$ -inch manifold U_2 was determined in the following manner: All stopcocks were closed, then, stopcock Z_7 was opened, and the calibration chamber T , having a volume (V), was filled through stopcock Z_{17} to a controlled pressure of approximately 1 atmosphere of dry nitrogen. This pressure (P) was measured by manometer S and read with the cathetometer. Stopcock Z_7 was then closed and the $\frac{1}{2}$ -inch manifold U_2 was evacuated through stopcock Z_{17} to a pressure of a few microns; stopcocks Z_{16} and Z_{17} were then closed and stopcock Z_7 opened,

allowing the confined gas to expand into the $\frac{1}{2}$ -inch manifold U_2 . The manometer was then read, indicating the expanded pressure (P_1), and the volume was calculated by the formula $PV = P_1V_1$.

By a similar procedure, the volumes of the 1-inch manifold U_1 and the expansion reservoirs M and N , with their accessory tubulations, were calculated, with the following results:

Volume of $\frac{1}{2}$ -inch manifold $U_2 = 839.2$ ml

Volume of 1-inch manifold $U_1 = 2069.0$ ml

Volume of expansion reservoir $N = 7615.5$ ml

Volume of expansion reservoir $M = 15,756$ ml

CALCULATION OF EXPANDED PRESSURES

Figure 4 is a plot of the pressures indicated by the ionization gauge versus the calculated values obtained by the pressure-division method. The procedure for calculating the pressure values, with specific reference to the results of run number 3, is discussed here: Nitrogen, at a pressure of 1.76 mm of mercury, was injected into calibration reservoirs Y_1 , Y_2 , and Y_3 . The system was evacuated to 4.8×10^{-6} torr (P_2), as indicated by the ionization gauge control unit.

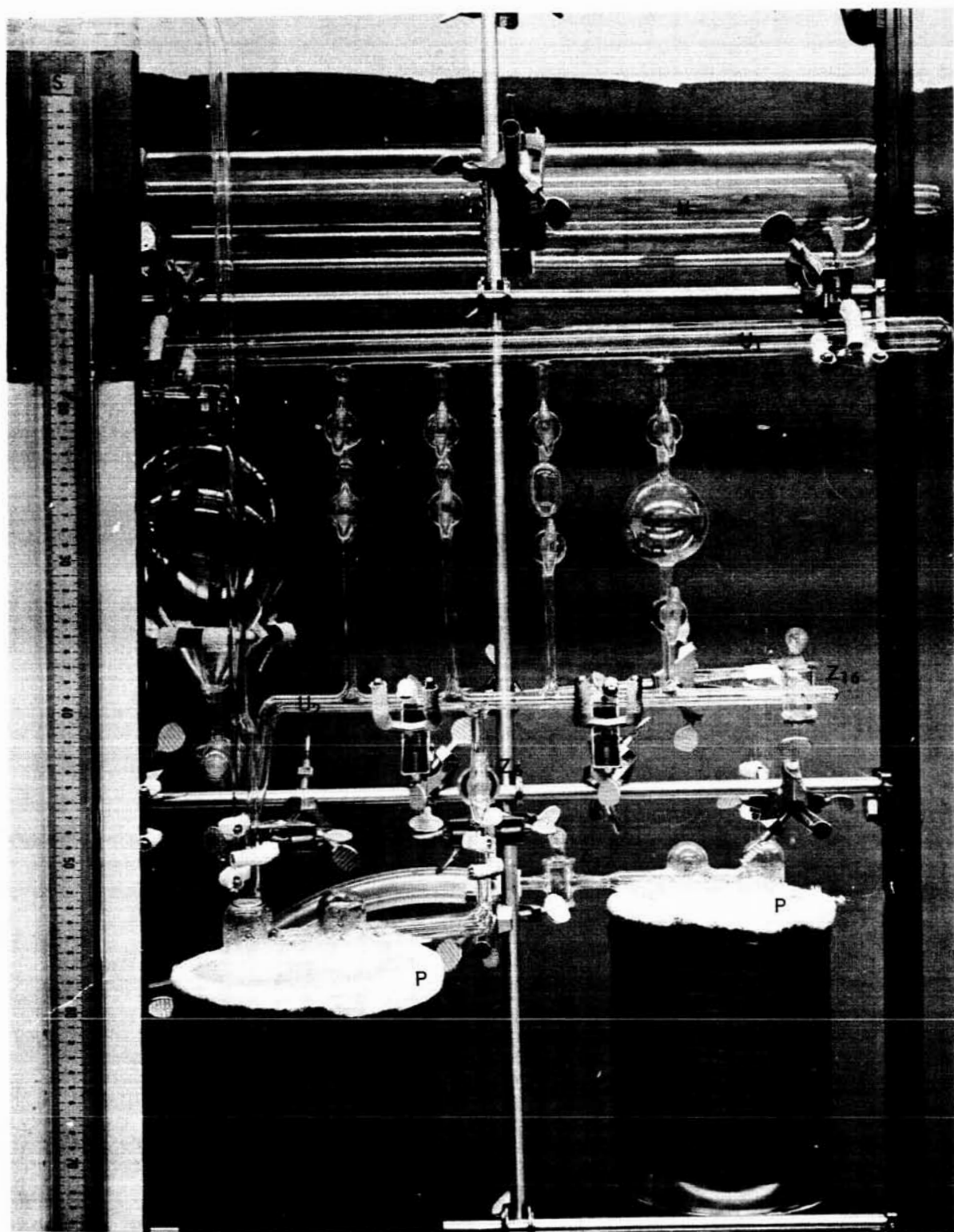


FIGURE 3.—Controlled pressure reservoirs and calibration chamber.

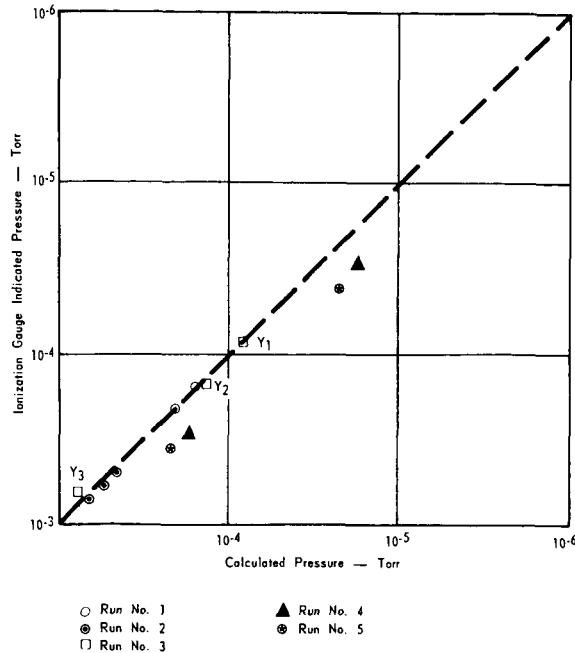


FIGURE 4.—Indicated versus calculated pressure.

Expansion of Y_1 Reservoir

The nitrogen in calibration reservoir Y_1 , having a volume of 1.2716 ml (V_1) at 1.76 mm of mercury (P_1), was allowed to expand into the evacuated system having a volume of 25441 ml (V_2), thus creating a total volume of 25442 ml (V_f).

After stabilization of the expanded gas, the ionization gauge filament was activated, and an indicated pressure of 9.6×10^{-5} torr was read on the ionization gauge control unit. Then, using Boyle's law, the pressure in the total volume (P_f) was calculated by the following formula:

$$P_f V_f = P_1 V_1 + P_2 V_2$$

$$P_f (25442) = (1.76 \times 1.2716) + (0.000048 \times 25441)$$

$$P_f = 0.0000927 \text{ or } 9.3 \times 10^{-5} \text{ torr}$$

Expansion of Y_2 Reservoir

The nitrogen, at a pressure of 1.76 mm of mercury (P_1), in calibration reservoir Y_2 , having a volume of 0.8654 ml (V_1) was allowed to expand into the system having now a volume of 25442 ml (V_2) at 9.3×10^{-5} torr (P_2), thus creating a new volume of 25443 ml (V_f).

The ionization gauge was activated, and the indicated pressure was read as 1.8×10^{-4} torr; the calculated pressure was obtained as follows:

$$P_f V_f = P_1 V_1 + P_2 V_2$$

$$P_f (25443) = (1.76 \times 0.8654) + (0.000093 \times 25442)$$

$$P_f = 0.00015 \text{ or } 1.5 \times 10^{-4} \text{ torr}$$

Expansion of Y_3 Reservoir

Finally, the nitrogen at a pressure of 1.76 mm of mercury (P_1), in calibration reservoir Y_3 , having a volume of 10.6836 ml (V_1) was allowed to expand into the system, which now had a volume of 25443 ml (V_2) and a pressure of 1.5×10^{-4} torr (P_2), thus creating a new volume of 25453 ml (V_f).

Upon activation of the ionization gauge, the new indicated pressure was read as 8.3×10^{-4} torr and the calculated pressure obtained as follows:

$$P_f V_f = P_1 V_1 + P_2 V_2$$

$$P_f (25453) = (1.76 \times 10.683) + (0.00015 \times 25443)$$

$$P_f = 0.000888 \text{ or } 8.9 \times 10^{-4} \text{ torr}$$

CALIBRATION OF IONIZATION GAUGE BY PRESSURE DIVISION

Operation of the Ionization Gauge

In order to reduce the effects of gauge pumping, an attempt was made to minimize interaction between the gauge and the system. This was accomplished by operating the grid with greatly reduced emission current. The normal operating grid current of 10 ma was reduced to 1 ma. To further minimize interaction, the filament was operated only a fraction of the time. Due to operating the gauge at reduced grid current, it was necessary to proportionately reduce its pressure indications.

Initial Pumpdown

The Kane valve H was closed, and Subsystem 1 was roughed in by means of the two VacSorb fore-pumps to a pressure of 1 to 5 microns; then, the VacSorbs were valved off, and the VacIon pump was put into operation. Subsystem 2 was evacuated to approximately 10^{-4} torr by means of the mechanical pump, with stopcocks Z_4 and Z_{17}

closed and all other stopcocks open, and with the Dewars on each of the cold traps filled to the required level with liquid nitrogen.

Uniting Subsystems

The Kane valve H was then opened uniting the two subsystems. When the VacIon pump had brought the pressure of Subsystem 2 down to approximately 10^{-6} torr, determined by activating ionization gauge L , the Kane valve H was closed, thus readying Subsystem 2 for the calibration run and placing Subsystem 1 in a standby condition.

Injection of Gas

Stopcocks Z_{12} , Z_{13} , Z_{14} , and Z_{15} were closed, stopcocks Z_8 , Z_9 , Z_{10} , Z_{11} , and Z_{17} were opened, and nitrogen at a pressure between 0.4 and 50 mm of mercury was metered into the calibration volumes Y_1 , Y_2 , Y_3 , and Y_4 . Stopcocks Z_8 , Z_9 , Z_{10} , Z_{11} , Z_{16} , and Z_{17} were then closed. The metered quantity of dry nitrogen was measured by manometer S and by the cathetometer. The filament of the ionization gauge L was activated, and the indicated pressure of the system was read on the ionization gauge control unit W . The filament was then switched off.

Gas Expansion

Next, stopcock Z_{12} was opened, allowing the confined gas from the controlled pressure reservoir Y_1 to expand into the tubulation and the expansion reservoirs. A five minute period was allowed for expansion stabilization before the filament of the ionization gauge was activated and the expanded indicated pressure read. This procedure was repeated for the expansion of the gas confined within the controlled pressure reservoirs Y_2 , Y_3 , and Y_4 . The calculated pressures were then obtained, using Boyle's law, and were compared with the indicated pressure readings obtained from the ionization gauge and its control unit. The system was then evacuated in preparation for another calibration run.

A total of 45 calibration runs was accomplished utilizing this technique. The results of five calibration runs were selected as typical and are plotted on the log-scale graph of Figure 4. On this graph the ordinate of a given point is the

logarithm of the pressure indicated by the ionization gauge; the abscissa of the given point is the logarithm of the corresponding calculated pressure in torr.

RECOMMENDATIONS AND CONCLUSIONS

The results obtained with the pressure-division system prove that it can be used as a vacuum gauge calibration system. For greater convenience in routine operation, Subsystem 2, which is used to expand various pressures, can be compacted and simplified. This can be done by reducing the volumes of the expansion reservoirs (M and N) and the controlled pressure reservoirs (Y_1 through Y_4), while maintaining their respective volume ratios. Also, a controlled leak should be incorporated to meter the nitrogen gas into the controlled pressure reservoirs, and the u-tube manometers should be replaced with ivory point index manometers.

Within the range of the present system, the use of Pyrex glass is satisfactory. To extend the operating range to vacuums greater than 10^{-6} torr, the components should be fabricated from stainless steel and the volumes of the reservoirs reduced as suggested in the preceding paragraph. High vacuum bakeable valves and provisions for bakeout should also be required.

The accuracy and the sources of error encountered in this pressure-division experiment were similar to those discussed by Schuhmann (4) who stated that the accuracy attained was ± 10 percent within the 10^{-4} to 10^{-6} torr range. The results of 45 calibration runs completed, 5 of which are presented in Figure 4, indicate a similar accuracy. As mentioned earlier, the calibration of an ionization gauge is time consuming and requires approximately 20 runs of about 3 hours each.

As an alternative, there is a less time-consuming method of gauge calibration in which the accuracy is degraded. A reference ionization gauge can be calibrated by the pressure-division method, and in turn can be used as a reference against which other ionization gauges are compared. However, this procedure has the disadvantage of the usual random variation in the response of the reference ionization gauge. This random variation, being comparable to that of the gauge to be compared, degrades the accuracy of the resulting calibration

by a factor of the order of the square root of 2 (6). This would be adequate for many tests performed in environmental chambers where such accuracy is acceptable. In applications requiring higher accuracy, the gauge should be calibrated by the pressure-division system, or, by a system of comparable accuracy.

The authors acknowledge the assistance of Mr. Philip Robling in assembling the pressure-division apparatus and in performing the many calibration runs required.

REFERENCES

1. WAYNE B. NOTTINGHAM and FRANKLIN M. TORNEY, JR., Technical Report 379, MIT Research Laboratory of Electronics, AD248376.
2. M. KNUDSEN, Ann. d. Phys. (Leipzig 44,581) (1919).
3. M. KNUDSEN, Ann. d. Phys. (Leipzig 31,633) (1910).
4. S. SCHUHMAN, "Study of Knudsen's Method of Pressure Division as a Means of Calibrating Vacuum Gauges," 1962, Transactions of the American Vacuum Society Symposium, 9th National Vacuum Symposium.
5. P. R. YEAGER, NASA, Langley Research Center.
6. J. R. ROEHRIG and J. C. SIMONS, JR., "Vacuum Instrumentation," National Research Corp.

ANALOG TO DIGITAL CONVERTER FOR THE S-57 GSFC ION-CHAMBER EXPERIMENT*

JOSEPH C. THORNWALL

INTRODUCTION

This paper describes an 8-bit analog-to-digital converter (Figure 1) using magnetic-core and solid-state components in pulse operation. All

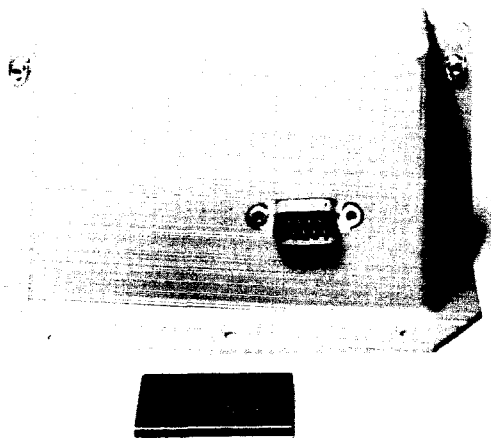


FIGURE 1.—A-D converter.

of the circuits are pulse-type with the exception of two telemetry input amplifiers, one data output gate, one transistor flip-flop and a DC comparator circuit. The input amplifiers were necessary to improve the rise and fall time of the telemetry signals and the flip-flop was required to convert the output binary number pulses to DC levels suitable for use by the data handling system. Figure 2 is a simple overall system block diagram showing how this converter fits into the system.

A pulse-circuit magnetic-core design approach was used for the following reasons:

1. To minimize the average power consumption.

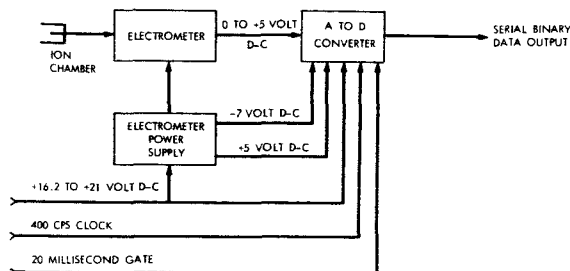


FIGURE 2.—System block diagram.

2. To provide low impedance circuits to reduce cross coupling problems in a relatively high component density package.
3. To make possible a high base current overdrive design, for good circuit operation margins with temperature, without increasing the average power consumption.
4. To take advantage of the inherent reliability offered by magnetic components.
5. To make use of the multiterminal possibilities available in magnetic components and as a result reduce the number of components required.

ENCODING

This converter uses a comparison method of encoding the analog input voltage¹. Figure 3 is a simple block diagram of the converter. This block diagram will be used to illustrate, element by element, the encoding method.

The analog voltage to be encoded is applied to one input of a two-input comparison circuit. A reference amplitude pulse voltage is applied to the other input. An output signal from the comparison circuit will occur when the amplitude of the reference input pulse is greater than the amplitude of the analog voltage input.

*Published as Goddard Space Flight Center Document X-610-64-116, May 1964.

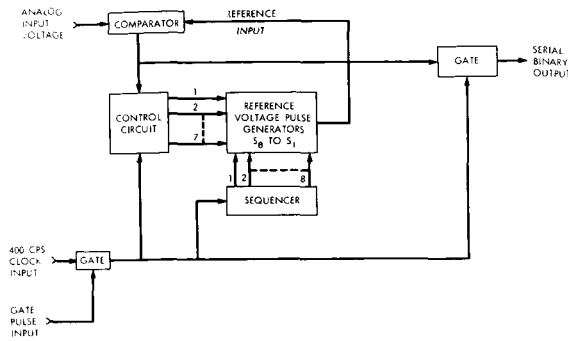


FIGURE 3.—Encoding method.

There are eight voltage sources available from the reference pulse generators. Each of these eight voltages is weighted in a binary fashion and can be controlled to sum all or any combination of these weights. If the maximum analog voltage is 5.1 volts, then the minimum analog voltage weight is .02 volts. Table I is a list of the voltage values of all eight levels.

The 400 cps clock input signal is applied to a sequencer, a control circuit and a gate. This input signal serves to time and control the formation of the binary number.

TABLE I.

Voltage Source	Binary Weight	Voltage Amplitude (volts)
S_1 -----	2^0	.02
S_2 -----	2^1	.04
S_3 -----	2^2	.08
S_4 -----	2^3	.16
S_5 -----	2^4	.32
S_6 -----	2^5	.64
S_7 -----	2^6	1.28
S_8 -----	2^7	2.56

The word gate pulse input allows eight clock pulses to be applied to the converter during the gate interval.

The sequencer provides a means of triggering the reference voltage pulse generators sequentially, from the most significant amplitude (S_8) first, to the least significant amplitude (S_1) last.

The control circuit is arranged to continue to trigger particular reference pulse generators. This

circuit can continue to hold a reference pulse generator on the remaining bits in the word after the sequencer trigger. That is; S_8 , would be triggered for 7 more clock pulses; S_7 , for 6 more; etc. through S_1 , the last bit of the word, does not require further holding.

The comparison circuit is used to determine whether or not the control circuit will continue to hold a reference voltage pulse generator. If a comparison circuit output signal *does* occur, when the sequencer triggers a reference voltage pulse generator, then the control circuit *will not* continue to trigger this reference generator on subsequent clock pulses. Conversely, if a comparison output signal *does not* occur then the control circuit *will* continue to trigger this reference generator on subsequent clock pulses. This represents the two conditions in which the reference voltage pulse amplitude has and has not, respectively, exceeded the amplitude of the analog input voltage. For each reference voltage amplitude smaller than the analog voltage input, the reference voltage generator added at that time will be held, and for each reference voltage pulse generator amplitude producing a reference sum larger than the analog voltage, the pulse generator added at that time will be dropped for the remainder of the bits in the word. Any of the 255 possible combinations can be found in just eight intervals of time.

The output gate in Figure 3 permits a "one" to appear in the output whenever there is no output signal from the comparison circuit, indicating that this binary weight is required, and a "zero" to appear in the output whenever there is an output signal from the comparison circuit, indicating that this binary weight is not required to represent the analog voltage amplitude. In this way the binary number is formed serially at the output as each comparison is made.

To illustrate by an example, assume the analog voltage input is 3.55 volts. The first of eight clock pulses causes the sequencer to trigger reference voltage pulse generator S_8 . This produces a reference voltage pulse of 2.56 volts. This voltage is less than the analog voltage therefore the control circuit is set up by the comparator to continue pulsing this reference voltage pulse generator for the remaining seven clock pulses in the word and a "one" is formed in the binary output.

The next clock pulse produces a trigger from the sequencer output number 2 and at the same time the control circuit triggers S_8 to produce a total reference voltage pulse amplitude of 3.84 volts ($S_8 + S_7$). This is a larger voltage than the analog input therefore the control circuit is not set up to trigger this reference voltage generator on the remaining clock pulses and a "zero" is formed in the binary output. The next clock pulse produces a trigger from sequencer output number 3 and at the same time the control circuit triggers S_8 to produce a total reference voltage pulse of 3.20 volts ($S_8 + S_6$). This is a smaller voltage than the analog input, therefore the control circuit is set up to hold this reference voltage generator for the remainder of the word and a "one" is formed in the binary output. If this procedure is followed for the remaining clock pulses it will be seen that the binary number 10110001 will be formed. This adds to a total of 3.54 volts or 0.01 volts less than the actual analog input voltage. The accuracy can be no better than +0.02 volts, the minimum voltage increment. This means that any binary number may represent an actual analog input voltage between the represented binary voltage number and a value just under the represented binary number +0.02 volts.

MEMORY ELEMENT CONTROL CIRCUIT

Figure 4 illustrates a familiar type of gating operation using a bistable "flip-flop" and gate combination. This d-c gate controls the flow of a signal input, allowing it to be passed or blocked, depending on the state of the flip-flop. The state

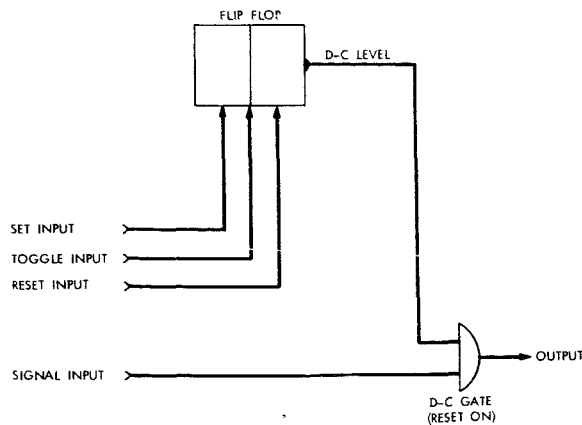


FIGURE 4.—D-C flip-flop and gate.

of the flip-flop is controlled by the set, reset and toggle inputs. The flip-flop is a memory element in this scheme which may or may not allow the signal to appear on the output and continues to "remember" one way or the other, until commanded to change.

When duty-cycle requirements are low (as was the case in our design), it is more efficient to go to an analogous pulse-type memory element for such a gating control. Figure 5 shows such an analo-

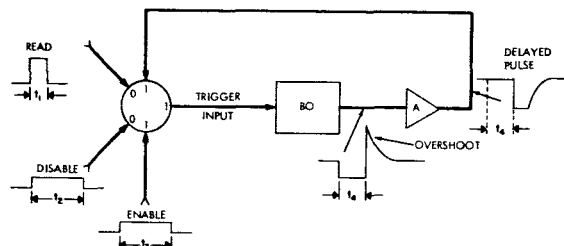


FIGURE 5.—Memory element.

gous pulse-type memory element gate arrangement.² A nonlinear-core transformer, shown by the circle, is combined with a linear-core transformer in a one-shot blocking oscillator circuit, shown by the square marked *BO*. The three inputs to the nonlinear transformer are defined as: read, enable, and disable. In Figure 4 these terms are analogous to signal in, reset and set, respectively. If the nonlinear transformer has been previously enabled, an output pulse will occur for each read pulse. If the nonlinear transformer has been previously disabled, there will be no output for each read pulse. The *BO* acts as a delay element and places the nonlinear transformer in the proper state so that, if an output does occur, the next read pulse will produce an output.

There are four pulse widths to consider when using this circuit; t_1 , t_2 , t_3 , t_4 in Figure 5.

Pulse width t_1 should be equal to or less than pulse width t_4 . This is necessary because at the end of the time t_4 the core will be switched to the "1" state by the overshoot from the *BO*. If there is a continuing read pulse during the overshoot interval the core may not be switched into the "1" state, as desired, because of the two simultaneous inputs of opposite sign.

The read pulse was arranged to switch the nonlinear core in 0.5 microseconds. This switching

time was used to obtain a large voltage per turn and at the same time not produce a trigger pulse too narrow to affect good triggering of the *BO*. More voltage could be obtained but for a smaller period of time. A Dynacor type 072S22A1 nonlinear core was used in this application. This core has a maximum switching time of the order of 20 micro-seconds.³ If the disable pulse⁴² is arranged to switch the core in 20 microseconds then the trigger voltage ratio between a read and a disable input will be 20/0.5 or 40:1. This means that the number of turns required to produce a peak voltage trigger pulse of 5 volts, 0.5 microseconds, during the read input will produce only a 5/40th volt, 20 microsecond pulse during the disable input interval. This is far below the minimum trigger voltage required to trigger the *BO*.

There is an additional trigger pulse, occurring during the read pulse interval, that must be considered. This is the noise voltage pulse produced when the nonlinear core has not been enabled and a read pulse is applied. This pulse causes the noise flux of the core to be switched. If sufficient excitation is applied to switch the full flux in the core in 0.5 microsecond, then this same excitation produces a 0.05 microsecond pulse when the noise flux is being switched. The amplitude of this pulse is 0.23v/turn in the circuit used in this converter. This voltage must be kept below the *BO* trigger voltage threshold. Tests on the 2N861 transistor in the *BO* circuit indicate that a 0.05 microsecond pulse has to have a magnitude of 4.2 volts to trigger the *BO*. A 6-turn trigger voltage winding produces a noise voltage of 1.38 volts peak, which is 1/3rd the *BO* trigger threshold.

The enable pulse width, t_3 , may have any width sufficient to switch the core. The enable pulse may be applied to the input of the amplifier A (Figure 5) instead of to a separate winding on the core. Also a single amplifier may be used to enable, read or disable a number of cores simultaneously.

Figure 6 shows one application of memory element circuits. This is a toggle-input core flip-flop made up of two memory elements with the *BO* overshoot outputs of each unit cross coupled to the other unit. This flip-flop was used to obtain two-phase pulses for driving alternate shift

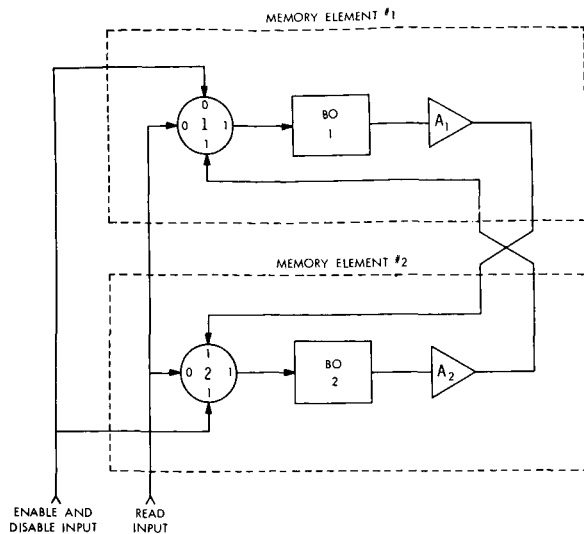


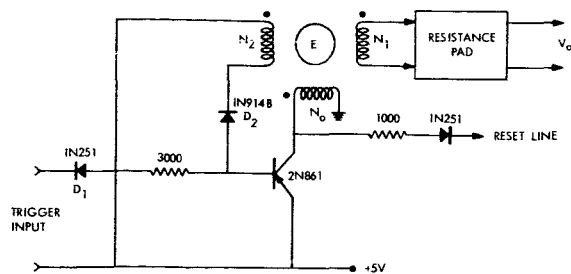
FIGURE 6.—Core flip-flop.

lines of the sequencer. In this circuit the disable and enable inputs are in series, allowing a single pulse to control both elements simultaneously. In this connection, unlike that of the single memory element, the enable-disable pulse may switch the cores at any speed. There is no requirement to switch slowly. This may be explained as follows. The enable-disable input of Figure 6 disables core number 1 and enables core number 2. If core number 1 triggers *BO* 1 the output from *A*₁ will enable core number 2, which is in the same direction as the enable input. If core number 2 switches the trigger pulse to *BO* 2 is of the wrong polarity to trigger this *BO*. This shows that there are no spurious conditions in the circuit that will result in the wrong core being enabled or disabled because of a *BO* being triggered on a disable input.

REFERENCE VOLTAGE PULSE GENERATOR

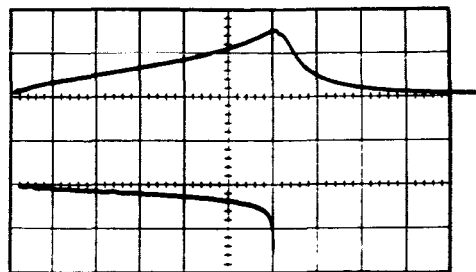
A circuit diagram of a reference voltage pulse generator is shown in Figure 7. This circuit is the same as the *BO* circuit shown in figure 8 except a nonlinear core is used and a reset is provided. There are eight of these pulse generators required for an eight-bit converter.

A resistance pad was used to adjust the amplitude of the binary-weighted voltage pulse on the output of each generator. This method of adjustment was used because there is a limit to the amount of voltage step down that can be achieved without seriously degrading the pulse waveform.



E. DYNACOR 071510A300 NONLINEAR CORE
 RESET: RETURNS TO -7VOLTS THROUGH A TRANSISTOR
 SWITCH. THIS LINE COMMON TO ALL PULSE GENERATORS.

FIGURE 7.—Reference voltage pulse generator.



LINEAR CORE WAVEFORMS. TOP TRACE IS COLLECTOR
 CURRENT, 20 ma./cm. BOTTOM TRACE IS COLLECTOR
 VOLTAGE, 0.5 VOLT/cm. HORIZONTAL RATE, 1 μ sec/cm.
 CORE: ALADDIN NO. 25-525-01

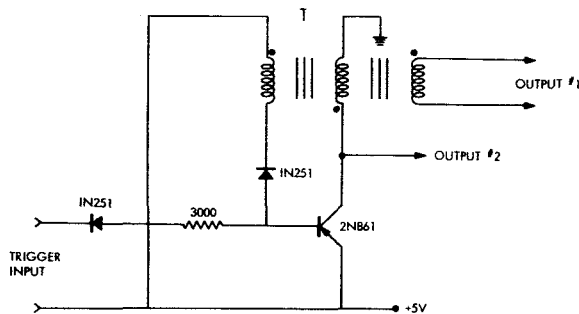
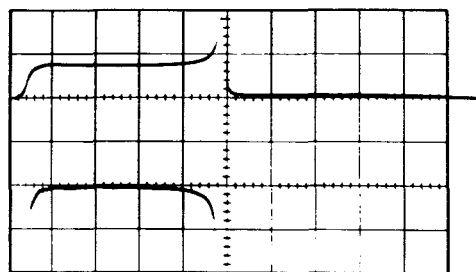


FIGURE 8.—Blocking oscillator.



NONLINEAR CORE WAVEFORMS. TOP TRACE IS COLLECTOR
 CURRENT, 50 ma./cm. BOTTOM TRACE IS COLLECTOR VOLTAGE,
 0.5 VOLT/cm. HORIZONTAL RATE, 0.5 μ sec/cm.
 CORE: DYNACOR 072522A300.

FIGURE 9.—Nonlinear and linear-core BO waveforms.

The turns ratio on all of the pulse generators from 2^5 and lower was limited to 0.33.

There are two major reasons for using a nonlinear core in this generator instead of a linear core; (a) smaller capacity coupling between the output and primary winding, and (b) better waveform. The windings on a nonlinear core may be segment wound and still maintain an acceptable waveform. This is not so in a linear core transformer in which the windings must be distributed one over the other in order to obtain an acceptable waveform. Where capacity is not a problem, the use of a linear core will usually result in a circuit containing less components. Figure 9 shows the difference between the voltage output waveform obtained with the two cores. This difference is caused by a continually changing collector current over the width of the pulse with the linear core and a substantially constant collector current with the nonlinear core. This causes a change in the transistor V_{ce} (Sat.) voltage over the width of the pulse with the linear core and results in a droop in the output pulse, while V_{ce} (Sat.) remains

essentially constant over the width of the pulse with the nonlinear core and results in a relatively constant voltage amplitude across the top of the pulse.

Figure 10 illustrates the method in which the reference voltage pulses are summed. Any stray capacitive coupling between the primary and secondary winding is not desirable because all of the voltage pulse sources are connected in series. This shunt voltage from capacitive coupling at each source will provide a spurious pulse at the output of the series-connected windings that will cause the desired inductively-coupled pulse to be modified. This spurious pulse was reduced to a negligible value by using a nonlinear core and segmenting the output windings. The only source not affected by this coupling is S_1 , because it is at ground potential and is unaffected by shunt capacity; therefore a linear core is used in this source.

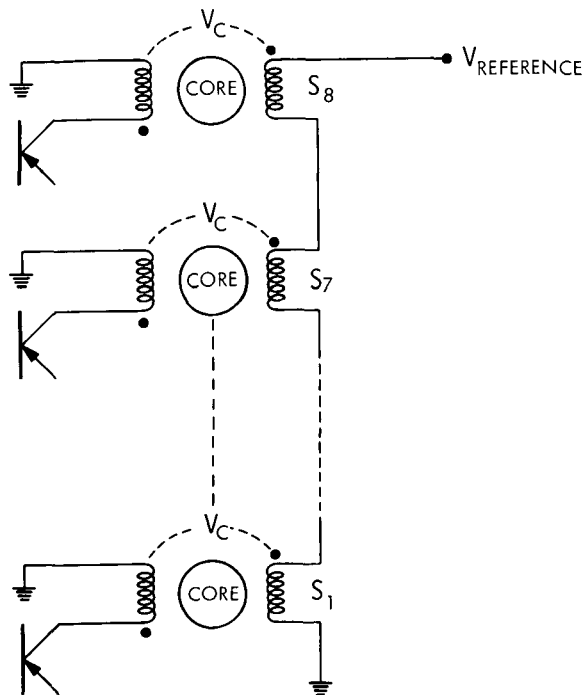


FIGURE 10.—Summing method for reference voltage generators.

The peak pulse voltage output (V_0) was measured in the circuit shown in Figure 11. It was possible to measure the amplitude of the pulse voltage to an accuracy of ± 1 millivolt out of 5 volts with this circuit.

The turns on the feedback winding (N_2) and the feedback diode (D_2) were chosen to minimize the change in V_0 with temperature. The total resistance (R_{DB}) of a series-connected diode and base-to-emitter transistor junction may have a positive, zero, or negative slope as a function of temperature, depending on the magnitude of the applied voltage. This suggests that there should be an optimum number of turns for N_2 , for a given

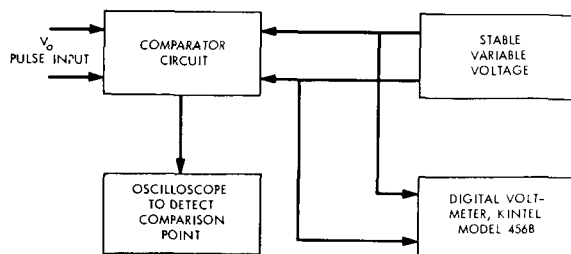


FIGURE 11.—Test set up for measuring V_0 .

diode and transistor, that will allow the change in R_{DB} to compensate for the temperature changes of the transistor collector-to-emitter saturation resistance and the effective resistance of the core. An optimum number of turns in N_2 was found by experiment and resulted in a maximum change in V_0 of less than 0.2 percent of the temperature range of -20°C to $+60^\circ\text{C}$.

COMPARISON CIRCUIT

Figure 12 is a circuit diagram of the comparison circuit.⁴ This circuit uses two dual N-P-N silicon transistors. A1 and A4 are in one package; A2 and A3, in another. The type used here is the 2N2642. This transistor has individual characteristics similar to the 2N930. The triode matching characteristics include a forward d-c current gain balance of a minimum of 0.9, a minimum base-to-emitter-voltage differential of 5 millivolts and a base-to-emitter-voltage differential temperature gradient of 10 microvolt/ $^\circ\text{C}$. These characteristics are acceptable provided the voltage differential of 5 millivolts and the gain ratio can be balanced out by the circuit.

The circuit operates in the following way. The input to A1 is biased sufficiently positive to insure that A2 is conducting and A3 is cut off. D1 is biased in the forward direction by the current supplied by the $+15$ volt supply through the 56 kilohm resistor to the collector of A2, which is at a potential of $+8.2$ volts. The resulting voltage drop across this diode produces a reverse-bias voltage across the base-to-emitter junction of transistor amplifier A5.

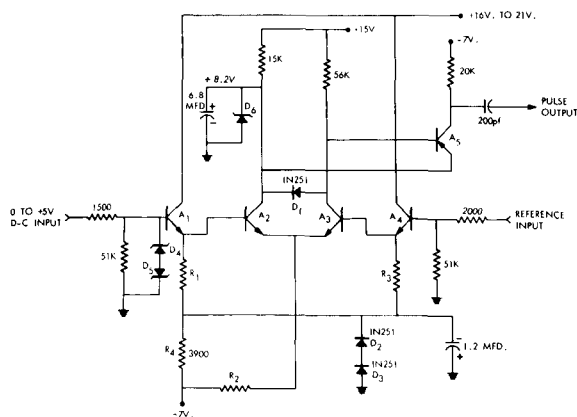


FIGURE 12.—Comparison circuit.

A3 will remain cut off as long as the base-to-emitter voltage is either below the threshold voltage of this junction or is reverse biased by A2. The d-c voltage appearing on the emitter of A3 is a function of the 0 to 5 volt input voltage. If a positive voltage pulse, with an amplitude sufficient to forward-bias A3, is applied to the reference input, A3 will begin to conduct. This will begin to bring D1 out of conduction. If the reference input pulse has sufficient amplitude and width the collector of A3 will eventually become more negative than +8.2 volts and A5 will be switched on. The resulting output pulse is further amplified and used to trigger a BO. There is sufficient gain in the system so that a change of less than 1 mv on the d-c input, at a threshold level, will cause the BO to trigger or not trigger. In other words the threshold error is less than 1 mv.

The electrometer output voltage, during range switching and voltage turn-on periods, can approach -15 volts. Also during periods of saturation the output voltage can approach +15 volts. Zener diodes D4 and D5 were placed across the input so that these large positive or negative voltages cannot appear on the base of A1.

Diodes D2 and D3 are biased in the forward direction to provide a voltage to bias "on" A1 and A4. This bias is necessary to eliminate the base-to-emitter junction threshold voltage. This diode voltage also varies with temperature in such a way as to compensate for the changes in the transistor base-to-emitter threshold.

The values of resistors R1, R2, and R3 are adjusted to eliminate the unbalance in the circuit and to set the threshold voltage to zero.

SEQUENCER CIRCUIT

The sequencer and delayed sequencer are shown symbolically in figure 13.⁵ The sequencer is similar to a ring counter in which a single one is transferred from core to core in response to the 400 cps clock pulses. The sequencer was designed so that up to two cores can be switched by the output of each core. When each core reads out it switches the next core in the sequencer, transferring the one, and at the same time a core is switched in the delayed sequencer. That particular core in the delayed sequencer then produces an output pulse after a time delay. This is a

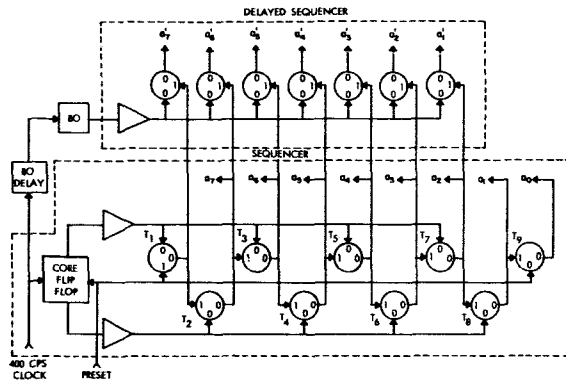


FIGURE 13.—Sequencer and delayed sequencer.

simple way of obtaining delayed pulses, a_1 through a_9 , with only a single BO delay.

The output of each stage of the sequencer is used to trigger each reference voltage pulse generator. A negative pulse is required for this trigger and is obtained from the output of core T2, in the case of the trigger for S8, at the time the one is transferred from T1 to T2. A ninth core was added to obtain a total of eight sequential pulses to trigger the eight reference pulse generators. This is required because a one is placed in T1 during preset and T1 does not produce a negative pulse on the first clock pulse of the word. The ninth core is preset at the same time as T1.

The delayed sequencer pulses are required to set up the control circuit elements. The delay allows each sequencer output to trigger a reference voltage generator and to permit the comparison circuit to produce a pulse. The delayed sequencer output pulse attempts to set up the control to hold this reference voltage generator in the active mode. If a pulse appears on the output of the comparator, the delayed sequencer pulse is prevented from enabling the control circuit and if a pulse does not appear from the comparator the control circuit is enabled to allow this reference voltage generator to be triggered on the remaining clock pulses.

The preset pulse for the sequencer is applied at the time of the leading edge of the word gate pulse so that at the time of the first clock pulse the sequencer is set up to transfer the one in core T1 to core T2 and on the eighth pulse core T9 is set up to produce an output.

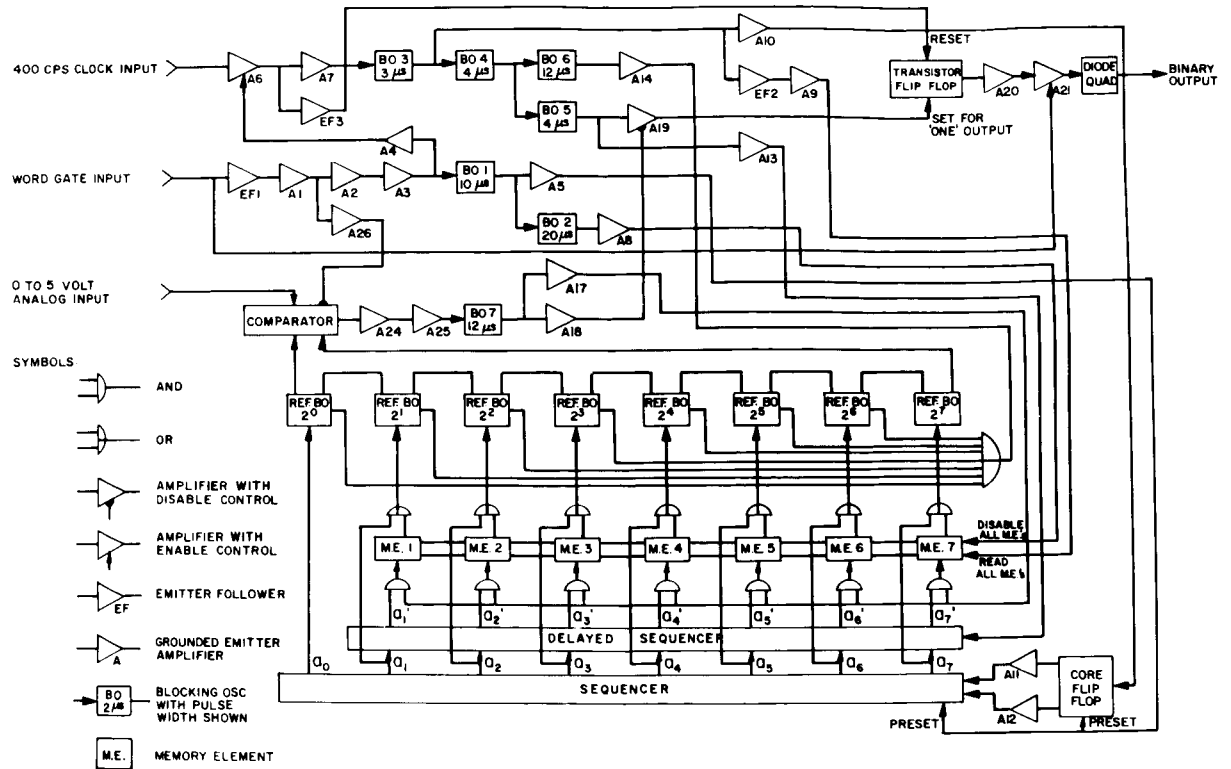


FIGURE 14.—Block diagram 8-bit A-D converter.

BLOCK DIAGRAM DESCRIPTION

Figure 14 is a complete block diagram of the converter and Figure 15 is a timing diagram showing only the pertinent waveforms required in the conversion process.

The word gate input signal is applied to *EF-1* and *A-21*. *EF-1* provides a high impedance load to the gate signal. *A-21* is provided with a source of collector load resistance voltage by the word gate signal during the word gate interval. The circuit was arranged this way so that no matter what kind of failure may occur with the converter there will be no output except during the gating interval. This was felt to be necessary because this converter shares the outputs of other experiments tied to this same point. Any malfunction within the converter could take out all of the experiments. The same reasoning applies to the need for a diode quad in series with the output signal so that a short to ground will not short out the output signals of all of the other experiments.

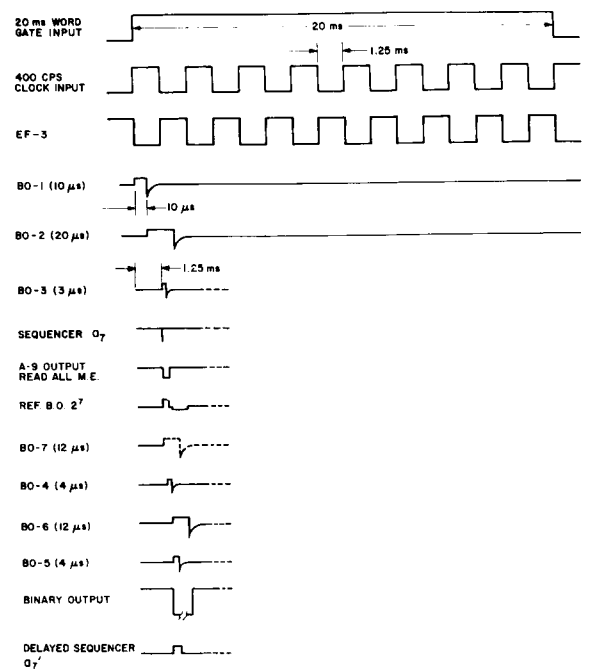


FIGURE 15.—Timing diagram.

The word gate input signal may have a maximum rise time of 50 microseconds. Amplifiers A-1, A-2, and A-3 are used to improve this rise time and invert the signal so that a negative trigger may be applied to BO-1 at the leading edge of the word gate pulse. The output of BO-1 drives A-5 which presets the core flip-flop and sequencer. The overshoot from BO-1 triggers BO-2. BO-2 drives A-8 to provide a pulse to disable all of the memory elements, clearing out the information that remained in the memory elements from the preceding conversion.

The 400 cps clock input signal is applied to a high input impedance amplifier, A-6. This amplifier is gated on only during the 20 millisecond word gate pulse interval so that there will be just 8 clock pulses appearing on the output of A-6 during each gating interval. A-6 inverts the clock signal and, on each positive-going swing of the input signal, provides a reset pulse to the transistor flip-flop through EF-3. This forces the binary output to be in the "zero" state in preparation for the next binary bit to be formed on the negative half cycle of the 400 cps clock input signal.

The rise and fall time of the clock input may have a maximum value of 5 microseconds. A-6 and A-7 improve this rise and fall time and present a negative trigger pulse to BO-3 at each negative-going half cycle of the clock input signal. This pulse is used to drive A-10, for switching the core flip-flop, and A-9, for reading all of the memory elements simultaneously. In the case of the first binary bit, 2^7 , there will be no output from any of the memory elements and sequencer output a_7 will trigger reference BO, 2^7 . Output a_7 also switches the first core in the delayed sequencer. If we assume the 0 to 5 volts analog input is larger than the 2^7 reference voltage, BO-7 will not be triggered.

The overshoot of BO-3 triggers BO-4. BO-4 provides a 4 microseconds delay to give time for the reference BO to apply a pulse to the comparator and for BO-7 to be triggered. The overshoot of BO-4 triggers BO-6 and BO-5. BO-5 applies a signal to A-19. In this case, since BO-7 was not triggered, an output will appear on A-19 and the transistor flip-flop will cause a "one" to be formed in the binary output. If BO-7 had been triggered A-18 would inhibit A-19 so

that there would be no trigger to the flip flop and a "zero" would be formed in the binary output.

BO-5 also drives A-13. A-13 provides a pulse to the delayed sequencer. This pulse causes the first core in the delayed sequencer, previously switched by sequencer output a_7 , to provide an output $a_7 \cdot a_7$ is applied to a gate. This gate is controlled by BO-7 and, since BO-7 did not trigger in this case, a_7 enables memory element number 7 so that on subsequent read pulses from A-9 reference BO 2^7 will continue to be triggered. If BO-7 had been triggered, output a_7 would have been blocked and memory element number 7 would not have been enabled.

BO-6, triggered by BO-4, drives A-14. The output of A-14 provides a pulse to reset all of the reference BO's. In this case only reference BO 2^7 will reset, however the circuit is arranged so that one or all of the reference BO's will be reset if required.

This explains the sequence of operation for the first bit. The remaining 7 bits are formed in exactly the same way, the only difference is the sequencer output transfers to the next output each time.

A-26 is an amplifier which blocks the comparator during the interval between word gate pulses. This was found to be necessary after the unit was connected to the electrometer. The output of the electrometer went negative because of a reaction between the electrometer amplifier and the comparator. This occurred only on the most sensitive range with no signal input, and caused a relaxation-type oscillation to be set up that was not corrected until the electrometer output signal was increased to +0.3 volts. This reaction occurred at some time after the binary conversion, therefore by blocking the comparator between word gate pulses it was possible to eliminate the instability and to measure down to the d-c off-set of the electrometer.

SPECIFICATIONS

The prototype converter has the following characteristics:

1. Mechanical (See figure 1).
 - (a) Size: $4\frac{1}{2} \times 3 \times 1$ inches excluding connector and mounting flanges.

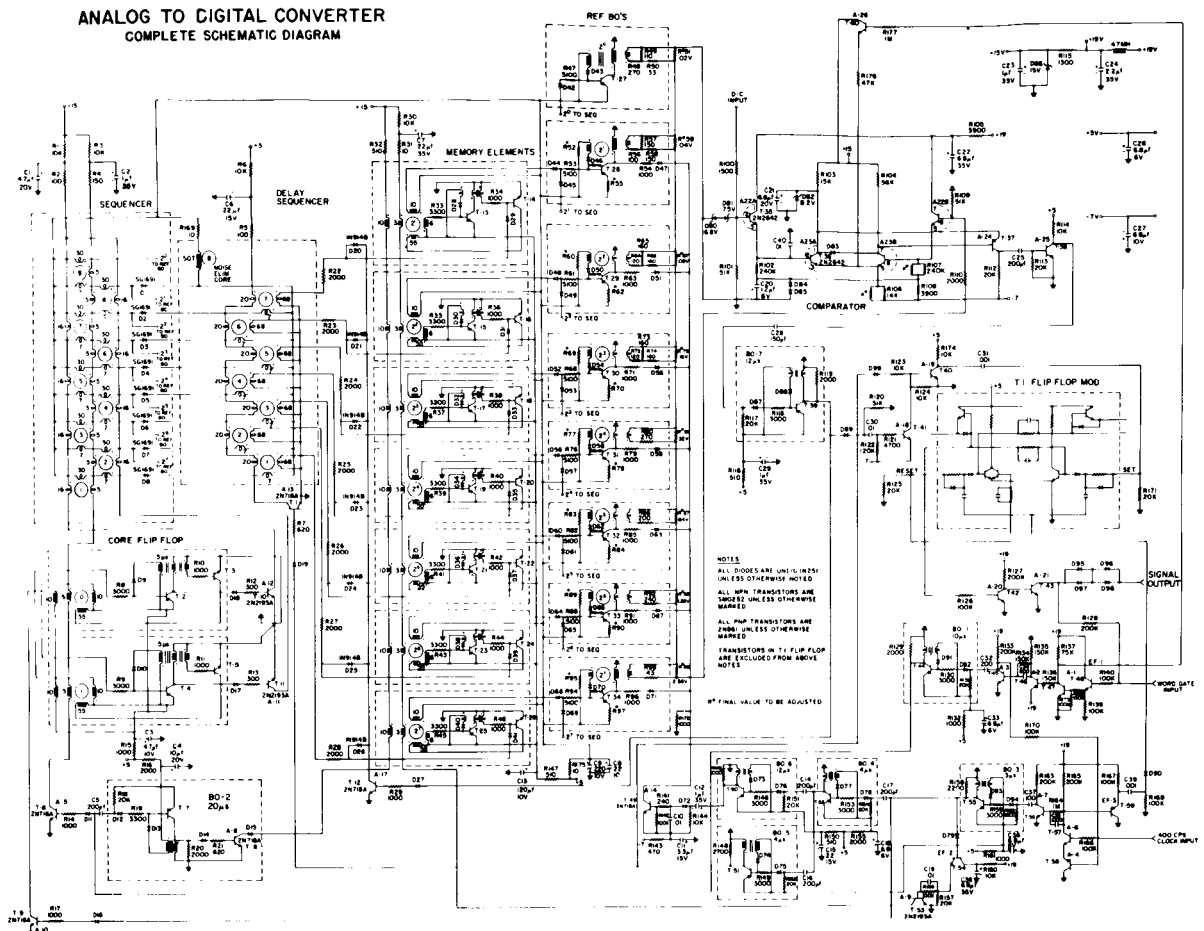


FIGURE 16.—Analog to digital converter. Complete schematic diagram.

- (b) Weight: 400 grams.
- (c) Fully encapsulated with Armstrong C-7 epoxy adhesive.
- (d) Vibration: 15G, 5 to 2000 cps.
2. Electrical
- (a) Analog input voltage: 0 to +5.1 volts d-c.
- (b) Analog input resistance:
- (1) 45.5 kilohms at -20°C . and +.04 volt analog input.
 - (2) 52.5 kilohms at $+60^{\circ}\text{C}$. and +5 volt analog input.
- (c) Temperature: -20°C to $+60^{\circ}\text{C}$.
- (d) Power supply voltages:
- (1) +16.2 to +21 volts d-c at 2.6 ma. (current measured at 18.5v).
 - (2) -7.000 ± 0.1 percent volts d-c at 1.8 ma.
 - (3) $+5.000 \pm 0.1$ percent volts d-c at 1.1 ma.
- (e) Power input: Approximately 70 milliwatts total. The converter by itself, including the flip-flop but without amplifiers, requires 20 milliwatts. The pulse logic circuit by itself requires .37 milliwatts.
- (f) Accuracy: ± 0.5 percent full scale over the temperature range of -20°C . to $+60^{\circ}\text{C}$.

- (g) Binary output: Serial with the most significant bit first. Binary "one" is zero volts and binary "zero" is +16 volts.
- (h) Binary output impedance: 200 kilohms.
- (i) Clock input requirements: 0 to +5 volts, 400 cps, square wave with maximum rise and fall time of 5 microseconds. Positive-going half cycle must coincide with the beginning of the word gate.
- (j) Word gate input requirements: 0 to +16 volts, 20 milliseconds long with a maximum rise time of 50 microseconds. Must be positive during the gating interval.
- (k) Word rate: Approximately 1.5 per second.
- (l) Clock input resistance: 100K ohm.
- (m) Word gate input resistance: 70 kilohms.

CIRCUIT DIAGRAM

The complete circuit diagram is shown in figure 16. Each of the amplifiers and blocking oscillators are labeled the same as in the block diagram. All of the circuits not already described are straightforward, therefore no attempt will be made to go into a detailed circuit description.

ACKNOWLEDGMENTS

I would like to thank Mr. Robert Lencho for his work on the breadboard of this converter and the detailed measurements made during the reference BO investigations. Also Mr. Frederick Hallberg for his work on the packaging, troubleshooting and circuit design, Mr. Ernest Nyberg, Mr. William Gallo and Mr. Alexander Bogash for their circuit board and schematic diagram layouts. This converter certainly would not have been possible without the help of these and many more individuals in the Solar Physics Branch of the Goddard Space Flight Center.

REFERENCES

1. *Notes on Analog-Digital Conversion Techniques*, edited by ALFRED K. SUSSKIND, (John Wiley and Sons and MIT Technology Press) pp. 5-54.
2. For a description of the logic symbol shown in figure 5, see *Digital Applications of Magnetic Devices*, edited by ALBERT J. MEYERHOFF, (John J. Wiley & Sons, 1960) p. 149.
3. Dynacor Engineering Bulletin No. DN-1004, figure 3.
4. For a description of comparator circuits, see *Pulse and Digital Circuits*, JACOB MILLMAN and HERBERT TAUB, (McGraw Hill Book Co., 1956) pp. 458-84.
5. See ALBERT J. MEYERHOFF, *op. cit.*, pp. 115-40 for a description of these circuits.

THE DETERMINATION OF ABSOLUTE PHOTON FLUXES AND APPLICATIONS TO CALIBRATION PROCEDURES IN THE 100Å TO 300Å RANGE*

D. H. TOMBOULIAN†

I. THE MEASUREMENT OF ABSOLUTE FLUXES

The problem of making absolute intensity measurements has recently received emphasis because of increased interest in spectral studies in the vacuum ultraviolet and in the soft x-ray region. The term "absolute" will be used to characterize an intensity measurement which is expressible in units of power per unit area or in terms of photon fluxes. The need for such measurements arises when, as in solar spectrophotometry, one wishes to compare radiant energies or photon fluxes at different wavelengths.

Ideally, absolute measurement of radiant energy can be carried out by calorimetric measurements under circumstances which guarantee that all photons in the beam are absorbed and that their energy is converted to thermal energy. This method is only practicable in the case of high intensity sources in the visible and is of very limited use in actual calibration procedures.

A second method is based on the utilization of absolute sources or radiation standards. Two such sources are of importance. One is the cavity radiator, whose spectral distribution is uniquely specified for a given temperature in accordance with Planck's law. As a second source, there is the possibility of utilizing the emission from high energy electrons accelerated in a betatron or synchrotron. As in the case of Planck's law, the spectral power radiated by a single electron is completely calculable, that is, the radiant power is expressible in absolute measure, once the maximum electron energy and the orbital radius are known.

The use of the above sources is beset with numerous experimental complexities. Electron accelerators are not generally available. So far as is known, the NBS 180 Mev synchrotron is the only accelerator in the United States which is available for use as a far ultraviolet source. There are many considerations which must be taken into account before the actual light beam may be adapted as a radiation standard. One must know the effective number of radiators in the beam, the maximum energy of electrons and the actual cross section of the electron beam. It is also necessary to see if the radial and vertical oscillations of the orbit are significant. The effects arising from the pulsed and polarized nature of the radiation must also be contemplated.

In the calibration of photographic emulsions or photoelectric devices of various kinds, one resorts to the use of thermocouples whose absolute response (microwatt per microvolt) is determined by first calibrating the thermoelectric detector against a radiation standard. This is done in the visible part of the spectrum and the calibration is assumed to hold in the far ultraviolet. Thermocouple measurements have been made down to 300Å in the case of strong radiations. Thermoelectric detectors do not possess 100 percent conversion efficiency. The diffuse scattering of the incident radiation and the kinetic energy carried away by the ejected photoelectrons is known to reduce the thermal conversion efficiency and give rise to errors which may be as high as 10 percent. Furthermore, when a thermocouple is used in the vacuum region, it becomes necessary to compare the response of the detector with its response obtained under atmospheric pressure. These conditions indicate that, even in the case of

*Published as *Goddard Space Flight Center Document X-614-64-332*, November 1964.

†Goddard Space Flight Center and Cornell University.

intense sources, the overall calibration of this type of detector is subject to large errors (~ 20 percent) in the vacuum ultraviolet.

Photoionization detectors have recently become more popular. In this category we list: gas filled ion-chambers with LiF or similar windows, ion chambers using differentially pumped entrance slits, flow-type proportion or GM counters equipped with special windows.

Although the techniques used in photoionization detectors are in need of improvement, their development has reached the state where they could be considered as reliable detectors of photon fluxes.

Consider a counter filled with a monatomic gas. Let the sensitive volume of the device be illuminated by a beam of sensibly monochromatic photons. If the energy of the incident photons exceeds the ionization energy of the rare gas (24.6 eV for He, 21.5 eV for Ne, 15.7 eV for Ar, 14.0 eV for Kr, 12.1 eV for Xe), then the interaction of such a photon with a rare gas atom may produce an electron-ion pair. For a sufficiently long path and high pressure, it may be assumed that the probability of stopping a photon in the counter is close to unity and that every photon that is absorbed will give rise to an ion pair. In other words, the total cross section of the atom is equal to its photoionization cross section. Thus in the soft x-ray region the use of ionization chambers and GM counters is based on the principle that the photon interaction results only in photoionization. This is reasonable since for low photon energies, the Thomson and Compton scattering processes are negligible, pair production is impossible and second order effects such as double excitation are highly improbable.

In particular, if a GM counter is to be used in the measurement of absolute fluxes, one must be assured that each photoelectron will give rise to one and only one pulse. As a second working principle, we may assume that the probability of detecting the photoelectron is unity. The strongest support for this statement is based on the possibility of attaining a well developed plateau in the counting rate vs anode voltage plot. Long standing experience with GM counters, where the output pulse does not depend on the charge liberated by the primary ionizing event, leads to the conviction that such counters fulfill the need

for photon flux detectors below 300\AA . Above this wavelength Samson¹ has successfully applied ionization chamber measurements of absolute radiation intensities up to about 1000\AA . The absence of convenient standard sources is no longer a serious handicap.

In the remaining portions of this report we describe the use of a GM counter as applied to the problem of calibrating (1) a Bendix photomultiplier and (2) certain photographic emulsions at wavelengths shorter than 300\AA .

II DETERMINATION OF THE EFFICIENCY OF A BENDIX PHOTOMULTIPLIER IN THE 100\AA TO 300\AA RANGE

The purpose of this investigation was to attempt a measurement of the efficiency of a Bendix magnetic photomultiplier at wavelengths distributed over the soft x-ray region. The choice of wavelengths used as dictated by the availability of characteristic valence emission bands of Be, Al and Cr, roughly 100\AA apart. These are fairly intense D.C. sources which are quite steady. They are therefore well suited for use with electronic detectors such as the GM counter, that is, for studies which involve the measurement of photon fluxes. Recent techniques in the preparation of suitable plastic windows have led to the successful use of the Geiger-Mueller (GM) counter as a device for determining absolute photon fluxes in the long wavelength region referred to above. The performance and the reliability of this device will be described elsewhere.²

It suffices to say that for not too high counting rates, the GM counter operates very satisfactorily up to and beyond 300\AA . (its use at wavelengths lying between 20\AA and 100\AA had already been demonstrated by Lukirskii et al).³

The measurements were performed in the following manner: A GM counter and Bendix photomultiplier were mounted within a grazing incidence spectrometer. The detectors were attached to a rotatable table so that the GM counter and the Bendix could be irradiated in succession by a narrow wavelength band of continuous radiation produced in the target of the soft x-ray tube. Care was exercised to see that each detector was exposed to the same beam geometrically. The counting rates of the detectors could then be

compared. These rates are given by

$$N_{GM} = N_0 \eta_{GM} \alpha_{GM} \quad (1)$$

and

$$N_{BX} = N_0 \eta_{BX} \alpha_{BX} \quad (2)$$

where N_0 is the incident flux of photons and where N , η and α denote respectively the counting rate, the efficiency and the window area sensitive to photons. In Eqs. (1) and (2), the subscripts BX and GM refer to the Bendix and the Geiger-Mueller counter. From relations (1) and (2), it follows that the efficiency of the Bendix is given by

$$\eta_{BX} = \frac{N_{BX} \alpha_{GM} \eta_{GM}}{N_{GM} \alpha_{BX}} \quad (3)$$

Relation (3) yields the desired efficiency at a specified wavelength since the quantities on the right hand side of Eq. (3) are known.

Unavoidably, the characteristic valence band appears on top of the continuous radiation from the target. At a given wavelength, the radiation incident on the counter consists of the portion of the valence band, the bremsstrahlung at this wavelength and also contributions in higher orders of the continuum at shorter wavelengths. In these tests, the contribution from the continuum was subtracted and only the intensity due to the emission band was used. The emission bands involved were those of Be K (edge at 110Å), Al $L_{2,3}$ (edge at 170Å) and Cr $M_{2,3}$ (edge at 296Å). In each case, a trace of the entire band was recorded and counting rates were compared at wavelengths near band peaks or at corresponding wavelength positions. In this manner it was possible to avoid the effect of the contamination by CK_{α} line in high orders, where such orders overlapped the emission band in question. Occasionally, it was convenient to deal with the second order spectrum of the emission band itself, because of the ease with which the band could be scanned. It is worth mentioning that (a) the observed features of the bands were in good agreement with those determined by other experiments and (b) the angular separation between various emission bands as observed by the detectors was consistent with the calculated value of the angular separation. The sources were operated at a pressure of 5×10^{-6} Torr, a target of about 50 ma and a target potential of about 1000 volts.

In the tabulation which follows we list the ratio of counting rates at the specified wave lengths.

TABLE I.—Observed Ratios of Photon Fluxes

Source	λ in Å	N_{GM}/N_{BX}	No. of indep. observations
Fe ⁵⁵	2	25 ± 2.5	2
Be K	115	80 ± 40	2
Al L	180	13.5 ± 1.2	4
Cr M	305	4.0 ± 1.0	5

The GM counter was operated on its voltage plateau and the anode voltage applied to the Bendix was 1700 volts, except in the case of the Al L band where the anode voltage was 1650 volts. In each case, the areas of the entrance slits were determined by measuring the relevant dimension by a traveling microscope.

Before we can evaluate η_{BX} we must determine the efficiency of the GM counter at the particular wavelength.

The efficiency of the GM counter is given by

$$\eta_{GM} = \frac{1}{K(\lambda)} [1 - \exp(-\alpha(\lambda)\rho)] \quad (4)$$

where $K(\lambda)$ is the attenuation due to the zapon window and the supporting mesh at the particular wavelength. The expression in the brackets is the fraction of the incident number of photons/sec stopped by the counting gas (helium). This quantity is nearly equal to unity except at 115Å where the absorption cross section is quite low. The quantity $\alpha(\lambda)$ is related to the photoionization cross section $\sigma(\lambda)$ in accordance with

$$\alpha(\lambda) = 9.65 \ell \sigma(\lambda) / T \quad (5)$$

where ℓ is the absorption path length in cm, $\sigma(\lambda)$ is the cross section in Mb and T is the absolute temperature of the gas. In formulating relation (4) it has been assumed that photoelectrons from the back wall do not contribute appreciably to the observed count and that the discharge of the counter is due to photons being absorbed by the gas and giving rise to an ion pair. The validity of the last assumption is tested by the flatness of the voltage plateau of the GM counter. One can

correct for the effect of the added counts due to photoelectrons reaching the wall. This was done in dealing with η_{BX} at 2\AA .

The quantity $K(\lambda)$ was determined in the following way: while making the counter window, an identical zapon film was mounted on a frame. The attenuation of this film was measured at 115\AA or at 180\AA using the appropriate radiation. At 300\AA , the attenuation was quite high and in addition the source was weak. In this case $K(\lambda)$ was computed by first obtaining the thickness of the film from the measured attenuation and published values⁴ of the linear absorption coefficient of zapon. Having determined the film thickness by experiment at the short wavelength, one could then compute the attenuation at 305\AA . The method was sufficiently precise for the present purposes.

The quantity $\alpha(\lambda)$ defined by Eq. (5) and appearing in Eq. (4) is determined from the experimental values^{5,6} of $\sigma(\lambda)$ at the various wavelengths. In Table II, we summarize the various quantities involved in evaluating the GM counter efficiency.

TABLE II. The Efficiency of the GM Counter

λ in \AA	$K(\lambda)$	$\sigma(\lambda)$ in Mb	$A(\lambda)$ Torr^{-1}	p Torr	η_{GM}
115	3.7	0.35	0.015	127	0.23
180	5.0	0.96	0.032	130	0.20
305	11.0	3.0	0.131	142	0.091

With the information available in Table I and Table II, the efficiency of the Bendix photomultiplier is computed by the use of Eq. (3). The results are collected in Table III and are shown in Fig. 1.

TABLE III.—The Efficiency of the Bendix Photomultiplier

in $\lambda\text{\AA}$	η_{BX}
2.....	2.3×10^{-7}
115.....	$(3.0 \pm 1.0) \times 10^{-3}$
180.....	$(1.7 \pm 0.5) \times 10^{-3}$
305.....	$(2.7 \pm 0.5) \times 10^{-3}$

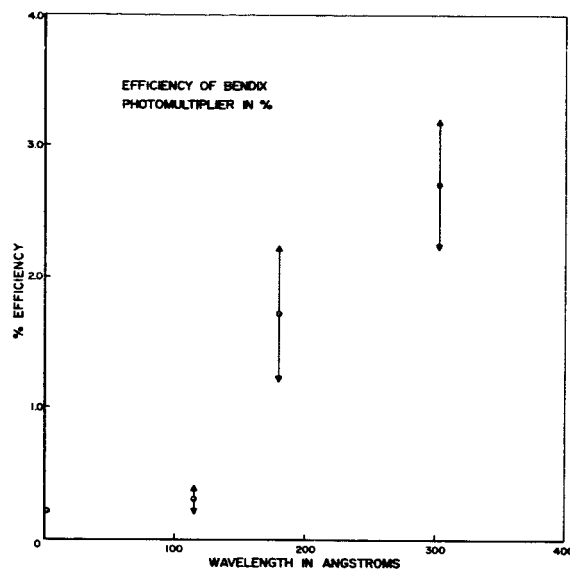


FIGURE 1.—Efficiency of Bendix photomultiplier in percent.

It appears that the Bendix is quite stable in its response and its characteristics have not changed in the course of several months. The efficiencies are plotted in the Fig. 1 along with the efficiency of a tungsten photocathode obtained by Lukirskii³ at 110\AA . A second Bendix multiplier of the same type was tested at 180\AA and its efficiency was found to be about 4 percent. This represents a variation of a factor 2. Variations of this order may originate in the response of the photocathode or in the collection efficiency of the secondary multiplication process.

III. CALIBRATION OF PHOTOGRAPHIC EMULSIONS

In the spectral region of interest two types of special emulsions, Eastman SWR and Ilford Q, are commonly utilized. The purpose of these measurements was to correlate the photographic density with the number of photons/cm² required to produce this density. The correlation was carried out at various wavelengths using the emission bands of Be and Al and also the Lyman series lines of ionized helium. The ionized helium lines were excited in a Schuler lamp. Only the first four members of the series ($\lambda \times 304\text{\AA}$, 256\AA , 243\AA and 237\AA) were made use of in the calibration.

A piece of photographic film was mounted within the grazing incidence spectrometer close to

the location of the Rowland circle. The film was inserted in the photon beam and exposed for a specified time. The photon flux was monitored before and after the exposure by the use of a GM counter whose entrance slit could be made to occupy the same position as that occupied by the film. For the purposes of photographic calibration, a series of different exposures were obtained at each discrete wavelength. In the case of the bands, the profile of the distribution was recorded photographically on film and also monitored by the GM counter. The films were developed for 3 minutes in dilute (1:10) D19 developer at 68°C. The emulsion densities were obtained from microphotometer traces in the usual way and the corresponding photon densities were deduced from the counting rates and GM scan records.

The results obtained by the use of the two emission bands are presented in Table IV. The SWR emulsion shows a response which is essentially "flat" in the number of photons needed to

produce a given density. The response of the Q emulsion is more nearly flat in energy.

TABLE IV.—Absolute Calibration of Emulsions
Emulsion SWR No. 82325

Microphotometer Deflection	Photographic Density	Photons/cm	
		at $\lambda = 115\text{\AA}$	at $\lambda = 180\text{\AA}$
10	1.00	5.60×10^9	4.80×10^9
20	0.70	3.35	3.30
30	0.52	2.50	2.40
40	0.40	2.00	1.80
50	0.30	1.50	1.40
60	0.22	1.20	1.16
70	0.16	0.92	0.92
80	0.097	0.69	0.72
90	0.045	0.46	0.50

Emulsion Q1 No. S4125

30	0.52	25×10^9	60×10^9
40	0.40	4.0	9.4
50	0.30	2.0	4.5
60	0.22	1.20	2.90
70	0.16	0.77	2.00
80	0.097	0.50	1.30
90	0.045	0.27	0.68

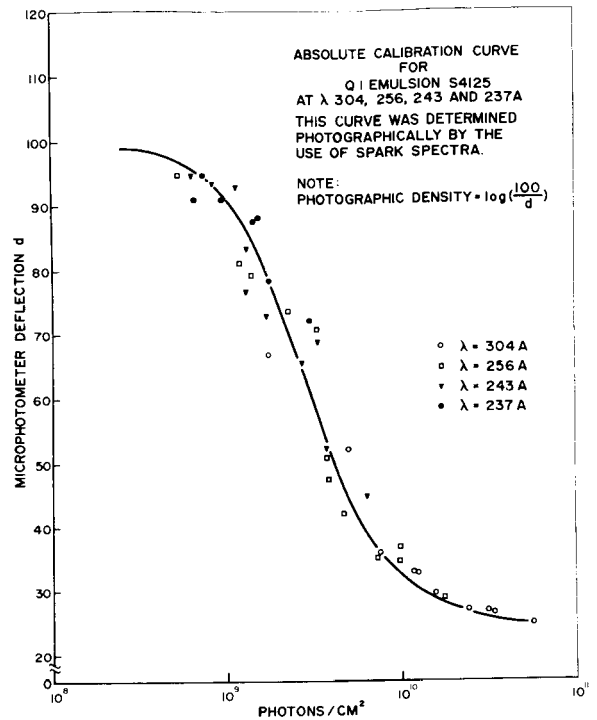


FIGURE 2.—Absolute calibration curve for Q 1 emulsion S4125 at λ 304, 256, 243 and 237 Å. This curve was determined photographically by the use of spark spectra. Note: Photographic density— $\log \left(\frac{100}{d} \right)$.

In Fig. 2 we display the various measurements of photon fluxes based on the He II lines. The solid curve was determined independently by the use of lines from a capillary discharge. The individual measurements carried out by the GM counter fall very close to the photographic calibration and convert the latter into an absolute calibration.

REFERENCES AND FOOTNOTES

1. J. A. R. SAMSON, *J. Opt. Soc. Am.* 54, 6 (1964).
2. D. L. EDERER and D. H. TOMBOULIAN, *J. of App. Optics.* In press.
3. A. P. LUKIRSKII, M. A. RAMAK, and L. A. SMIRNOV, *Opt. i Spektroskopiya* 9, 505 (1960) [English transl: *Opt. Spectry.* (USSR) 9, 262 (1960)].
4. D. H. TOMBOULIAN and D. E. BEDO, *Rev. Sci. Inst.* 26, 747 (1955).
5. D. J. BAKER, D. E. BEDO, and D. H. TOMBOULIAN, *Phys. Rev.* 124, 1471 (1961).
6. J. F. LOWRY, *Phys. Rev.* In press.
7. The efficiency at 2 Å was evaluated by taking into account the effect of photoelectrons from the back-wall of the counter.

CORRELATION OF ULTRAVIOLET EXCITED LUMINESCENCE AND INTERBAND TRANSITIONS IN HEXAGONAL ZINC SULFIDE

WILLIAM C. WALKER*

Measurements of the ultraviolet luminescent excitation spectrum, reflection spectrum, and photoemission of single crystal hexagonal zinc sulfide for the spectral region 3 to 11 eV are presented. A tentative assignment of reflection peaks to specific interband transitions is made and used to discuss the structure observed in the luminescent excitation spectrum on the basis of assumptions concerning the luminescent mechanism. Evidence is presented to support the contention that the structure in the excitation spectrum is mainly controlled by the dependence of the electron transport and trapping efficiency on the final states of the various interband transitions.

I. INTRODUCTION

Tremendous progress has been made recently in our understanding of ultraviolet absorption in solids and its interpretation through the identification of structure in the optical^{1,3} and photoemissive⁴ spectra with specific interband transitions and excitons at or near critical points in the joint density of states function^{5,6}. It is now possible to specify the initial and final states of the majority of the strong interband transitions in a broad range of solids. This advance has opened the way for detailed studies of a variety of secondary processes associated with the trapping, recombination, scattering and ejection of the excited electrons. The recent verification of the conservation of tangential electron momentum in photoemission from silicon⁷ is a fine example of the results to be reaped from such studies.

Luminescence is another example of a secondary process which should greatly profit by taking advantage of the new knowledge. Various aspects of luminescence have been studied for many years.⁸ Only recently, however, has interest revived in the use of luminescent phenomena to study intrinsic solid state properties.^{9,10} This is partly due to the fact that emphasis on the defect

nature of the luminescent mechanism has overshadowed the fact that in many cases the initial electron excitation and subsequent transport and trapping at the luminescent center are intimately connected with the host lattice. It should thus be possible to use the luminescent process as a tool to study the transport of highly excited electrons and the relative trapping efficiency for electrons in specific states as well as an aid in the identification of various interband transitions.¹¹

Previous measurements of ultraviolet induced luminescence, with the exception of some very recent papers,⁹ have been devoted mainly to the search for efficient phosphors for specific purposes.¹² As a result many of the measurements were carried out on samples which were unsuitable for detailed analysis of intrinsic effects. These studies are, however, invaluable in providing a guide for the selection of materials for detailed study.

In the present paper initial data of ultraviolet induced luminescence excitation spectra, reflection spectra and the spectral dependence of the photoemission are reported for a single crystal of hexagonal zinc sulfide in the region 3 to 11 eV. The aim of the study was to establish a correlation between the luminescent excitation process and the states of the excited electrons. Zinc sulfide was chosen for initial study because (a) the lumi-

*Goddard Space Flight Center; on leave from the Physics Department, University of California, Santa Barbara, under a NAS-NASA Senior Postdoctoral Fellowship.

nescent excitation process is known to involve the host lattice, (b) the major defects responsible for the common luminescence are reasonably well established,¹³ (c) the crystal is transparent to its luminescent radiation and (d) sufficient theoretical guides are available to allow a reasonable identification of the interband transitions.⁶

II. EXPERIMENTAL

All the measurements to be reported were carried out on a single crystal of hexagonal zinc sulfide obtained from Semi Elements, Inc. The crystal, whose size was 2 cm × 1 cm × 2 mm, was clear and colorless with no traces of the discoloration sometimes found in such crystals. Observation of the typical green phosphorescence centered at 5200 Å under 2537 Å irradiation showed that trace amounts of Cu⁺ were present to act as the luminescent center.¹³ The long lifetime of the phosphorescence, measured to be of the order of one-half hour, indicated that another impurity, probably Al⁺⁺⁺, was present to provide a source of thermally released electrons.

A McPherson, model 225, monochromator equipped with a current regulated, low pressure hydrogen discharge source was used to provide monochromatic ultraviolet radiation for the measurements. The monochromator was operated with a spectral resolution ranging from 0.02 eV (100 micron slits) for the reflection spectra to 0.2 eV (1 mm slits) for the relative photo-yield measurements. An Electro-Mechanical Research #541A photomultiplier having a dark current of 10⁻¹¹A at a gain of 5 × 10⁵ at room temperature was used to monitor both the ZnS luminescence and that from a sodium salicylate reference. Since the quantum efficiency of sodium salicylate is known to be reasonably constant over a broad spectral range,¹⁴ the ratio of the two luminescent spectra gave a direct measure of the spectral dependence of the ZnS luminescence.

Reflection spectra of the ZnS crystal were recorded with a vacuum reflectometer employing a rotatable, sodium salicylate coated light pipe and an EMI 6256N photomultiplier. The measurements were made at an angle of incidence of 20° on a cleaved surface. Because the cleaving was done in air, the reflection spectrum can not be considered that from a pure surface. Experience with ZnS and related materials, however, has

shown that while moderate surface contamination reduces the magnitude of the reflectivity it does not obscure the structure due to interband transitions.

To aid in the interpretation of the luminescent data, the spectral dependence of the photoemission was measured by collecting the electron emission from the cleaved surface of the crystal and comparing it to the luminescence response of a sodium salicylate screen. Since the detailed behavior of the photoemission is sensitive to the nature of the surface, all that could be hoped for in this measurement was an estimate of the photo-threshold and perhaps an indication of structure due to the onset of interband transitions. As will be seen, even under these relatively poor conditions, the photoelectric data can aid significantly in the interpretation of the results.

III. RESULTS AND DISCUSSION

A. Reflection and photoemissive spectra

The reflection spectrum of hexagonal ZnS taken with unpolarized light is shown in Fig. 1. Data

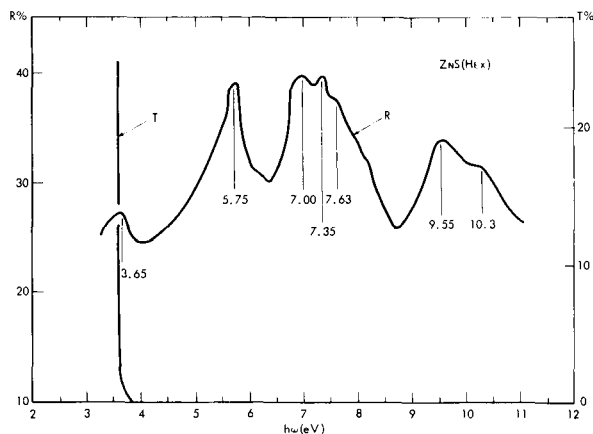


FIGURE 1.—Reflection spectrum of hexagonal ZnS. Data were taken at room temperature with unpolarized light.

were taken at about one-hundred points with an instrumental resolution of 0.02 eV. Reflection peaks were observed at 3.65, 5.75, 7.00, 7.35, 7.63 and 9.55 eV. An addition shoulder which is probably real appeared at 10.3 eV. For comparison the transmission spectrum near the fundamental energy gap is also shown. As will be discussed in more detail later, these peaks can be

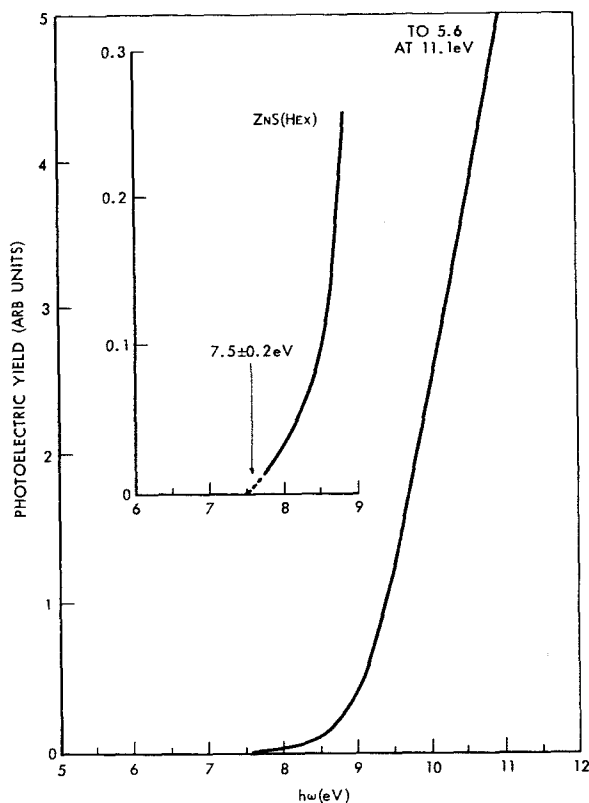


FIGURE 2.—Spectral dependence of the photoelectric yield of hexagonal ZnS. The insert shows the foot of the curve expanded by a factor of ten.

identified with interband transitions mainly at the center of the Brillouin zone and give a good picture of the order and relative magnitudes of the electronic states at this point.

The spectral dependence of the photoemission is shown in Fig. 2. Only gross features of this curve warrant attention because of the nature of the surface. As shown in the insert of Fig. 2, there is a weak threshold at 7.5 ± 0.2 eV together with a stronger increase near 9.0 eV. Beyond 9.0 eV the yield is linear with photon energy as predicted by Kane¹⁵ for direct transitions. An interpretation of these data will be given in the following section.

B. Band Structure

The assignment of structure in optical or photo-emissive spectral with specific features of the band structure is a difficult procedure and in general, requires either data on a broad range of related structures or a highly reliable set of band calcula-

tions. Since neither exist at present for wurtzite structures only a tentative assignment of the structure of Fig. 1 can be made. Fortunately, there do exist preliminary band calculations for hexagonal ZnS by Herman and Skillman¹⁶ and an analysis of experimental data by Phillips⁶ for the related material CdSo which can be used as guides in interpreting the ZnS spectrum. According to Phillips, the lowest energy direct transitions (3.65 eV) would be expected to correspond to transitions at the center of the Brillouin zone between the valence band maximum at Γ_5^v and the conduction band minimum Γ_1^c .¹⁷ Thus, the fundamental band gap is associated with a direct transition rather than an indirect one as suggested by Herman and Skillman. Indeed, it does not seem possible to get a consistent interpretation of the measurements presented here on the basis of an indirect gap. The remaining reflection peaks, with the exception of the one at 5.75 eV, are consistent with transitions at the Γ point. The 5.75 eV peak in ZnS probably corresponds to the 5.0 eV peak in CdSe which Phillips tentatively assigns to transitions near the M or K points. Using the condition of steepest slope for the transmission data of Fig. 1, a direct band edge of 3.6 ± 0.1 eV is obtained. This together with the 7.5 ± 0.2 eV photo-threshold yields an electron affinity E_A of about 3.7 eV and serves to fix the vacuum level. Assuming that Herman's numerical results at the center of the Brillouin zone are approximately correct, allowed transitions in the range 7 to 9 eV should occur between valence band states Γ_5^c , Γ_1^c , Γ_6^c . There are thus sufficient transitions at this point to account for the observed complex structure of the 7 and 9 eV peaks. Because $\Gamma_5^v \rightarrow \Gamma_5^c$ is expected to be strong,⁶ it is assigned to the main 7.00 eV peak. The remaining peaks can then be assigned to conform to the photoemission results. Since the true photo-threshold certainly exceeds 7.00 eV, the state Γ_5^c must fall below the vacuum level and hence the 9.55 eV peak is assigned to $\Gamma_5^v \rightarrow \Gamma_6^c$ rather than $\Gamma_6^v \rightarrow \Gamma_5^c$ as suggested by Phillips⁶ for the corresponding peak at 9 eV in CdSo. The 7.35 eV peak could be either $\Gamma_1^v \rightarrow \Gamma_5^c$ or $\Gamma_5^v \rightarrow \Gamma_1^c$ from energy considerations; it is assigned to the later transition in order to enable it to contribute to the photoemission near the threshold. The order of the electronic states at the center of the Brillouin zone deduced from

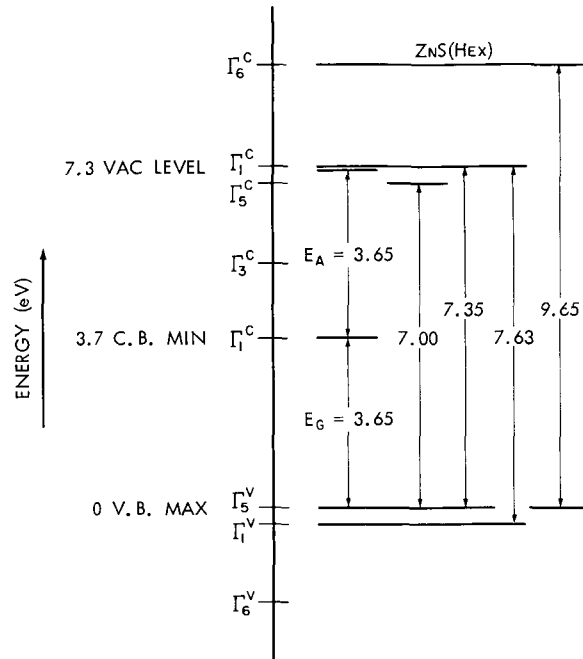


FIGURE 3.—Electronic states of hexagonal ZnS at the center of the Brillouin zone. The zero of energy is taken at the valence band maximum Γ_5 .

this analysis is shown in Fig. 3. The rapid rise in the photoyield near 9 eV is attributed in part to the onset of the $\Gamma_5^v \rightarrow \Gamma_6^c$ transition. On the basis of the photoyield data the vacuum level can be located near Γ_1^c to within ± 0.2 eV but the data are not good enough to determine if Γ_1^c is above or below the vacuum level. The author favors placing Γ_1^c just above the vacuum level and adopting the lower value of 7.3 eV for the photothreshold thus allowing $\Gamma_5^v \rightarrow \Gamma_1^c$ to contribute to the photoemission.

C. Luminescent Excitation Spectrum

The luminescent excitation spectrum measured at room temperature is shown in Fig. 4. Luminescent radiation emerging from the back surface of the crystal was monitored. The emission spectrum excited at 2537\AA consisted of the broad band centered on 5200\AA which is typical of the Cu^+ activator. The significant features of the spectrum are a rapid rise beginning at 3.9 eV, a broad peak near 5 eV with an enhancement at 5.9 eV and a decrease beginning near 6.5 eV, and a plateau starting at 7.6 eV having a magnitude about three tenths of the peak value. These

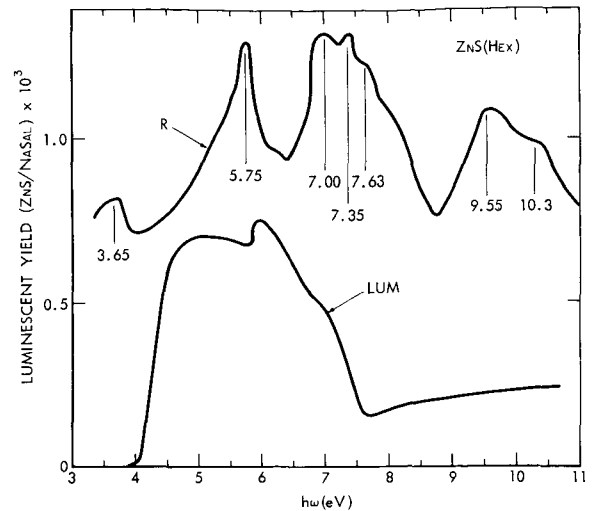


FIGURE 4.—Luminescent excitation spectrum of single crystal hexagonal ZnS . The reflection spectrum in arbitrary units, is included for comparison.

features will be interpreted on the basis of the following assumptions:

1. The luminescent mechanism is the recombination of an electron trapped at a Cu^+ site with a hole in the valence band.
2. The efficiency of the luminescence depends on the direct interband transition probability and the transport and trapping efficiency of electrons and holes.
3. The contribution of a specific interband transition to the electron-hole pair density is proportional to the product of the oscillator strength of the transition and the magnitude of the joint density of states at the singularities. It can then be shown¹⁸ that the relative contributions of two transitions are in the approximate ratio $E_i R(E_i) / E_j R(E_j)$, where $R(E_i)$ is the reflectivity at E_i .

On this basis, the luminescent onset near 4 eV can be attributed to the transitions $\Gamma_5^v \rightarrow \Gamma_1^c$. That the excitation curve does not rise abruptly at 3.6 eV is due probably to the fact that for the final state of this transition the average electron velocity $\hbar^{-1} \nabla_k E(k) = 0$ so that the transport of the electron to the trapping sites is retarded. Thus, it is not until transitions slightly removed from $k=0$ occur that significant luminescence begins. Since the 5.75 eV transition occurs for k other than zero, it is reasonable that the luminescence output

should continue to rise toward this transition. As soon as transitions to higher states set in, e.g., near 7.0 eV to the Γ_5^c state, the trapping efficiency appears to be reduced and the luminescence decreases. The fact that the break in the excitation spectrum at 7.6 eV nearly coincides with the photo-threshold and that from this point to higher energy the spectrum is nearly structureless indicates that once transitions to final states far removed from the conduction band minimum occur, the transport and trapping efficiencies are roughly averaged out to a constant value.

The energy dependence of the photon absorption depth with the associated variations in direct electron-hole surface recombination might be raised as an alternative interpretation to that given. Considerations of the optical constants for a variety of materials have shown that ϵ_2 , R , and k display approximately the same spectral dependence so that the reflectivity can be taken as a rough index of absorption. In this case, Fig. 1 shows that the transitions near 6, 7 and 9 eV correspond to roughly the same absorption coefficient so that the absorption depths for these transitions are comparable and major variations in surface recombination rates would be unlikely.

A comparison of the ratios of the luminescent response to the interband transition contribution to the electron-hole pair density for two energies gives a rough measure of the efficiency of the transport and trapping of electrons at luminescent sites and of the subsequent emission. For the transitions near 5.8, 7.0 and 9.6 eV the relation given above,¹⁸ yields the following ratios in order of increasing energy from 5.8 eV—1:0.5:0.1. Thus, electrons excited to the Γ_6^c state by 9.6 eV photons are one-tenth as efficient in producing luminescence as are those excited by 5.8 eV photons.

These measurements may also bear on the model for the luminescent mechanism. Two models, the Lambe-Klick and the Schön-Klasen models, have been proposed.¹⁹ The models differ principally by the order in which the Cu^+ sites are filled by electron trapping. In the Schön-Klasen model the Cu^+ sites are initially empty and are filled by the photo-excited electrons and subsequently radiate. In the Lambe-Klick model the Cu^+ sites are initially filled and the radiation occurs through recombination with the free hole

released by the incident light. The Cu^+ sites are subsequently reset by trapping photo-released electrons. The pertinent point of the present data is that the luminescent efficiency appears to depend on the final state of the electron through the transport and trapping process. For example, the transitions $\Gamma_5^c \rightarrow \Gamma_5^c$ and $\Gamma_5^c \rightarrow \Gamma_6^c$ having the same initial state but different final states yield widely different luminescent yields. Such dependence is more easily accounted for by the Schön-Klasen than the Lambe-Klick model.

Since the explanation proposed above for the correlation of the spectral distribution of the luminescent excitation spectrum and ultraviolet interband transitions should apply more generally, it will be of interest to extend the measurements to materials involving other luminescent mechanisms such as the alkali halides, magnesium oxide and diamond.

REFERENCES

1. H. R. PHILIPP and H. EHRENREICH, *Phys. Rev.* **129**, 1550 (1963).
2. M. CARDONA, *Phys. Rev.* **131**, 98 (1963).
3. W. C. WALKER and J. OSANTOWSKI, *Phys. Rev.* **134**, A152 (1964).
4. W. E. SPICER, *Phys. Rev. Letters* **11**, 243, (1963).
5. J. C. PHILLIPS, *Phys. Rev.* **125**, 1931 (1962).
6. J. C. PHILLIPS, *Phys. Rev.* **133**, A452 (1964).
7. G. W. GOBELI, F. G. ALLEN, and E. O. KANE, *Phys. Rev. Letters* **12**, 94 (1964).
8. P. PRINGSHEIM, *Fluorescence and Phosphorescence*, Interscience, Oxford Univ. Press (1949).
9. T. TIMUSK and W. MARTIENSEN, *Phys. Rev.* **128**, 1656 (1962).
10. J. J. HOPFIELD, *J. Phys. Chem. Solids* **10**, 110 (1959).
11. J. C. PHILLIPS, *Phys. Rev. Letters* **12**, 142 (1964).
12. F. S. JOHNSON, K. WATANABE, and R. TOUSEY, *J. Opt. Soc. Am.* **41**, 702 (1951).
13. C. C. KLICK and J. H. SCHULMAN, *Solid State Physics*, Vol. 5, Academic Press (1957), pp. 145.
14. After nearly fifteen years of use as an ultraviolet wavelength converter, there is still disagreement in the literature concerning the exact degree of spectral constancy of the quantum efficiency of sodium salicylate. See J. A. R. Samson *J. Opt. Soc. Am.* **54**, 6 (1964). In no case, however, is the structure observed strong enough to invalidate the present results.
15. E. O. KANE, *Phys. Rev.* **127**, 131 (1962).
16. F. HERMAN and S. SKILLMAN, *Proc. of the Intern. Conf. on Semicond. Phys.*, Prague (1960), page 20.

17. The symmetry notations Γ_s and Γ_a used here are reversed from those of Herman and Skillman to accord with the notation of reference 6.
18. This ratio follows if the momentum matrix element is taken to be nearly constant in the expression for the interband transition probability. The ratio of transition probabilities at E_i and E_j is then

$$P(E_i)/P(E_j) = E_i \rho(E_i) / E_j \rho(E_j)$$

where $\rho(E)$ is the joint density of states. The ratio

$\rho(E_i)/\rho(E_j)$ is given by

$$E_i^2 \epsilon_2(E_i) / E_j^2 \epsilon_2(E_j)$$

or approximately

$$E_i^2 R(E_i) / E_j R(E_j).$$

Combining these expressions we get

$$P(E_i)/P(E_j) = E_i R(E_i) / E_j R(E_j)$$

19. See reference 13, pp. 157-158.

PILE-OF-PLATES POLARIZER FOR THE VACUUM ULTRAVIOLET*

WILLIAM C. WALKER†

A pile-of-plates polarizer and analyzer for the spectral region 1200–1600Å have been constructed from cleaved LiF plates 13 mm in diameter and 0.3 to 0.8 mm thick. Both polarizer and analyzer consist of identical cylinders with $\frac{3}{16}$ inch apertures and contain eight 1 mm slots cut at an angle of 30° to the axis of the cylinder. The slots are arranged in two opposing groups of four each to give an on axis emergent beam. Measurements with a six plate polarizer gave polarizations of 82 percent and 65 percent at 1200 and 1600Å respectively. The effective transmission of the polarizer was 4.3 percent and 21.5 percent at these two wavelengths. Performance characteristics of the polarizer will be presented and discussed.

INTRODUCTION

The importance of measurements utilizing polarized radiation in giving fundamental information concerning electromagnetic interactions in a variety of fields has been well established. Due to the lack of suitable polarizers for radiation below 2000Å and the multitude of significant work which can be done in more accessible regions of the spectrum, few attempts have been made to exploit polarization measurements in the far ultraviolet. Recent theoretical¹ and experimental² advances in solid state ultraviolet spectroscopy, which have re-emphasized the significance of polarization studies should soon drastically change this situation.

The present paper reports on a pile-of-plates polarizer for the region 2000–1200Å which was constructed in order to carry out measurements to determine and identify interband transitions in non-cubic solids. In the absence of suitable dichroic materials below 2000Å, Brewster angle reflection and transmission techniques appear most easily exploited for the production of polarized UV radiation. The polarizer to be described is a conventional transmission polarizer using LiF plates. The transmission technique was chosen over the more efficient reflection method in order to give a more adaptable geometry for future

spectroscopic work. Once linearly polarized radiation has been produced it should not be too difficult to construct quarter wave plates from strained isotropic materials thus leading to the production of circularly polarized radiation. Indeed preliminary considerations using strained LiF indicates that such UV optical elements are feasible. The new areas, e.g., in optical pumping and others opened up by such an advance should be considerable.

DESCRIPTION AND RESULTS

The LiF pile-of-plates polarizer was constructed from thin plates cleaved from a cylindrical blank 13 mm in diameter. Plates varying in thickness from 0.3 to 0.8 mm were used. The polarizer and analyzer, which could be rotated while under vacuum, consisted of two stainless steel cylinders 3.00" long and 0.562" in diameter having a $\frac{3}{16}$ " axial hole. Each cylinder contained eight slots cut at 30° to the axis, giving an angle of incidence of 60°, and arranged in two opposing groups of four each to insure an on-axis emergent beam. As is well known³, optimum performance is obtained with such polarizers when the angle of incidence just exceeds the Brewster angle. Since this angle for LiF at 1200Å is 58°40' while that at 2000Å is 55°41', an incident angle of 60° represents a good compromise.

Measurements with the polarizer have been made (a) to determine its working characteristics, (b) to determine the polarization of the monochromator used and (c) to explore the polarization

*Published as *Goddard Space Flight Center Document X-613-64-87*, May 1964.

†Goddard Space Flight Center; on leave from the Physics Department, University of California, Santa Barbara, under a NAS-NASA Senior Postdoctoral Fellowship.

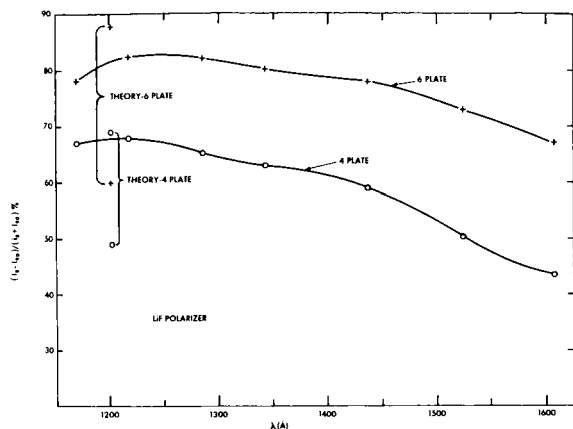


FIGURE 1.—Spectral dependence of the polarization of *LiF* pile-of-plates polarizers. I_0 and I_{90} are the transmitted intensities with the plane of incidence of the analyzer parallel and perpendicular respectively to the plane of incidence of the polarizer⁵.

dependence of the far UV reflection peaks recently observed for crystals with the wurtzite structure.⁴ The measurements concerning the operating characteristics of the polarizer will be discussed.

Figure 1 shows the polarization characteristics of the device for six and four plates. The theoretical values pertain to the two standard expressions giving the polarization for such polarizers depending on whether or not internally reflected beams are included in the transmitted beam.⁵ Since the internally reflected beams consist entirely of radiation polarized perpendicular to the plane of incidence, lower values are predicted when they are included in the analysis. The fact that the measured polarization is close to that calculated using the simple expression indicates that internally reflected beams were virtually eliminated by the natural wedging and fanning of the plates. For comparison a percent polarization of 80 percent corresponds to a rejection ratio $I_0/I_{90}=9/1$ while 50 percent polarization corresponds to $I_0/I_{90}=3/1$.

The effective transmission of the polarizing plates is shown in Fig. 2. The data are plotted so that the maximum transmission expected is 50 percent, i.e., the transmission involved is the ratio of the transmitted polarized radiation to the incident unpolarized radiation. In order to insure that the effects observed were due to polarized radiation and not transmission effects arising from

the rotation of the plates, the angular extinction dependence of the polarizer and analyzer was measured for several wavelengths. The results are shown in Fig. 3 and compared to the expected

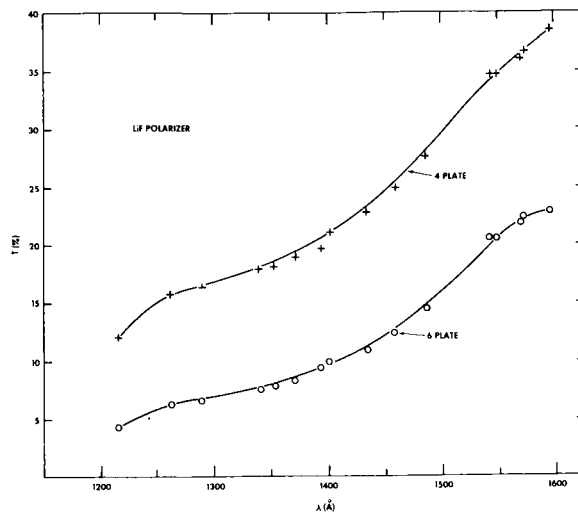


FIGURE 2.—Spectral dependence of the effective transmission of *LiF* pile-of-plates polarizers. The ratio of the intensities of the polarized transmitted beam to the unpolarized incident beam is plotted.

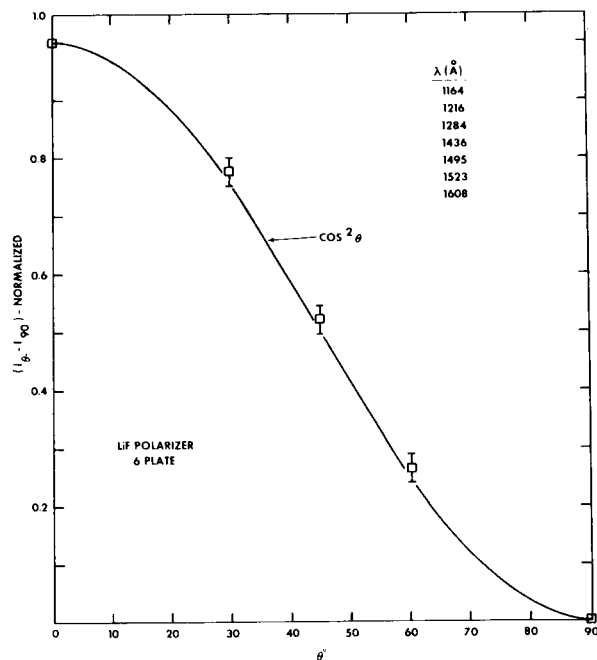


FIGURE 3.—Angular extinction characteristics of a *LiF* pile-of-plates polarizer. The measured values for the wavelengths listed all fall within the flags shown on the plotted points.

cosine squared dependence. The agreement with the cosine squared curve shows that a high degree of polarization was indeed produced.

Although the monochromator used for the measurements operated at normal incidence so that the exit beam should be unpolarized, the novelty of the UV polarizer suggested that a direct confirmation of this fact would be desirable. The results of this measurement showed that the maximum apparent polarization of the beam was in the range 1-2 percent. Since changes in transmission of the polarizer due to irregularities in the cleaved plates could easily amount to this much, it was concluded that the beam was completely unpolarized. It would be of considerable interest to repeat this measurement on Seya and grazing incidence instruments.

SUMMARY

A pile-of-plates polarizer for the 2000-1200Å region has been constructed and shown to be adequate for a wide range of polarization studies beyond the limit of available polarizers. Optimization of the quality and position of plates, e.g., by using polished rather than cleaved surfaces should markedly increase the performance of the device. Measurements of reflection spectra of crystals having the wurtzite structure have been

carried out⁶ using the polarizer and indicate that at least in one field, i.e., UV solid state spectroscopy, significant new results can be obtained. It is anticipated that the use of these polarizers will lead to an extension of other visible optical techniques such as optical pumping and the study of optical absorption of ions located in anisotropic crystalline fields into the vacuum ultraviolet.

REFERENCES

1. J. C. PHILLIPS, *Phys. Rev.* **133**, A452 (1964).
3. M. CARDONA, *Solid State Communications*, **1**, 109 (1963).
3. G. R. BIRD and W. A. SHURCLIFF, *JOSA* **49**, 235 (1959).
4. W. C. WALKER and J. OSANTOWSKI, *J. Phys. Chem. Solids*, to be published, 1964.
5. The theoretical values were calculated from the expressions

$$P = \frac{1 - \left(\frac{2n}{1+n^2}\right)^{2m}}{1 + \left(\frac{2n}{1+n^2}\right)^{2m}}$$

when internally reflected beams are neglected and

$$P = \frac{m}{m + \left(\frac{2m}{n^2-1}\right)^2}$$

when they are included, where m is the number of plates and n is the index of refraction. See reference 3.

6. W. C. WALKER and J. OSANTOWSKI, *Bull. Amer. Phys. Society* **9**, 222 (1964).

POLARIZATION DEPENDENCE OF THE FAR ULTRAVIOLET OPTICAL TRANSITIONS IN CdS*

W. C. WALKER†

The far ultraviolet reflection spectrum of CdS displays a number of distinct features in the range 5 to 18 eV which have been attributed, by analogy with the spectra of diamond and zinc blende materials, to direct valence-to-conduction band transitions.^{1,2} From a low-temperature study in polarized light out to 6 eV, Cardona³ attributed the complex 5.6-eV peak to the doublet (Δ_5 , Δ_6) - ($\Delta_1\Delta_3$) at 4.95 to 5.03 eV and a single transition analogous to $L_3'-L_3$ of the zinc blende structure at 5.58 eV. In addition, Phillips⁴ assigned the 9.5-eV peak to $\Gamma_6^v-\Gamma_5^c$. The author¹ has observed this latter peak to be double with maxima at 9.2 and 9.6 eV and suggested, it now appears erroneously, an identification of the 9.6-eV peak with the second member of the spin-orbit split $\Gamma_5^v-\Gamma_1^c$ (upper) transition. A series of three high-energy peaks at 14.4, 16.1, and 17.4 eV have been attributed to transitions originating on the cadmium 4d band.¹

Recent photoemissive studies by Kindig and Spicer⁵ have raised doubts concerning the exclusive role of direct transitions in explaining the above features of the optical spectrum. Maxima in the photoelectron energy distribution at electron energies of 0.6 and 1.8 eV which are independent of the exciting photon energy were interpreted as due to transitions to final states consisting of two maxima in the conduction-band density of states at 6.8 and 8.2 eV, respectively, above the valence-band maximum. A shoulder appearing in the energy distribution for photon energies exceeding 9.6 eV and whose position was linearly dependent on the photon energy was attributed to transitions from an initial state at

a maximum in the valence-band density of states located 1 eV below the top of the valence band. This interpretation is consistent with the optical data, provided we ignore the doublet nature of the 6-, 8-, and 9-eV peaks and postulate a third conduction-band density-of-states maximum about 5 eV above the top of the valence band. The optical transitions below 10 eV could then be accounted for by transitions from the single valence-band density-of-states maximum to the three conduction-band maxima. The three peaks at 14.4, 16.1, and 17.4 could then be attributed to transitions from another valence-band density-of-states maximum located about 9.3 eV below the top of the valence band to the same three conduction-band maxima. The neglect of the fine structure is reasonable since a fairly wide (twenty-Angstrom) bandwidth was employed in the photoemissive measurements. This interpretation requires, however, that k -nonconserving transitions occur with transition probabilities comparable to direct transitions.

In an attempt to clarify this situation the room-temperature reflection spectrum of a cleaved single crystal of CdS was measured in the hitherto inaccessible region 6-11 eV with radiation polarized perpendicular and parallel to the C axis. It was hoped that if k -conserving transitions were dominant, they would show a polarization dependence through the crystal selection rules. A recently developed pile-of-plates polarizer⁶ constructed from LiF plates was employed. It gave a polarization of 80 percent near 1216 Å (10 eV) and 70 percent near 1680 Å (8 eV). Figure 1 shows the reflection spectrum for the two orientations. Since the angle of incidence, 25°, was not zero, R_{\perp} was expected and observed to be larger than R_{\parallel} in the absence of specific crystalline anisotropy effects. Only two crystalline effects

*Published in *Physical Review Letters*, 13(2):51-52, July 13, 1964.

† Goddard Space Flight Center; on leave from the Physics Department, University of California, Santa Barbara, under a NAS-NASA Senior Postdoctoral Fellowship.

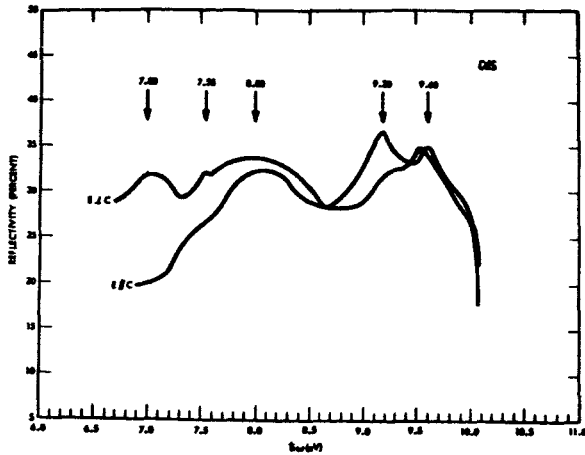


FIGURE 1.—Ultraviolet reflection spectrum of CdS in light linearly polarized perpendicular ($E \perp C$) and parallel ($E \parallel C$) to the C axis.

were observed in this region. These were a strong enhancement of the 7- and 9.2-eV peaks for $E \perp C$. These effects are interpreted as due to the crystalline selection rules governing the interband transition rather than a decrease in R for $E \parallel C$ associated with an approach to the Brewster condition for two reasons. (1) A strong reduction in R for $E \parallel C$ for an incident angle of 25° would occur only for $n \sim 0.5$. Since preliminary measurements by the author show that n is well above unity in this region, this effect is ruled out. (2) The polarization results show a definite enhancement at the two peaks in question for $E \perp C$ compared to data for unpolarized radiation. The 7-eV peak in CdS and a corresponding one at 6.3 eV in CdSe were interpreted by Phillips⁴ as due to $\Gamma_5^v \rightarrow \Gamma_5^c$ with the selection rule $E \parallel C$. This assignment was made on the basis of measurements using unpolarized light.² The present results clearly show that the selection rule for the 7-eV transition is $E \perp C$. If this transition takes

place at the center of the Brillouin zone, the only likely transition, neglecting the small (0.02-eV) spin-orbit splitting, would be $\Gamma_5^v - \Gamma_1^c$. Preliminary calculations by Herman⁷ for ZnS, however, would favor assigning $\Gamma_5^v - \Gamma_1^c$ to the 9.2-eV peak with $E \perp C$. It has not been possible to arrive at an assignment of the peaks below 10 eV in terms of transitions only at the Γ point which will agree with the polarization measurements.

The lack of polarization effects in the 8.0-eV peak and the complex nature of the 9-eV peak indicate that the situation in CdS is much more complex than in the zinc blend materials. Thus, while a strong k -conserving transition such as $\Gamma_5^v - \Gamma_1^c$ may contribute to the polarization-dependent part of the 9-eV peak, nonvertical transitions between other density-of-states maxima degenerate in energy with this transition may be responsible for the bulk of both the 8- and 9-eV peaks. Evidence for the violation of k conservation for certain transitions for which the initial electron state is strongly coupled to the proton field seems to exist for several materials.⁸ Further polarization studies at low temperature on CdS and CdSe, where spin-orbit effects may aid in the interpretation, should help clarify this complex situation.

REFERENCES

1. W. C. WALKER and J. OSANTOWSKI, to be published.
2. M. CARDONA, *Phys. Rev.*, **129**, 1068 (1963).
3. M. CARDONA, *Solid State Commun.*, **1**, 109 (1963).
4. J. C. PHILLIPS, *Phys. Rev.*, **133**, A452 (1964).
5. N. KINDIG and W. E. SPICER, *Bull. Am. Phys. Soc.*, **9**, 223 (1964).
6. W. C. WALKER, to be published.
7. F. HERMAN and S. SKILLMAN, *Proceedings of the International Conference on Semiconductor Physics, Prague, 1960* (Czechoslovakian Academy of Sciences, Prague, 1961).
8. W. E. SPICER, private communications.

AN ULTRAVIOLET ROCKET STELLAR SPECTROMETER*

DAVID U. WRIGHT, JR.

The Ultraviolet Rocket Spectrometer (UVR) was developed to measure stellar spectra in the wavelength interval 1100Å to 4000Å. The UVR consists of a 33 cm Dall-Kirkham telescope which provides input to a spectrometer utilizing a plane grating in convergent light. The $f/11.9$ system is such that the apparent coma of the plane grating is compensated for by the coma of the tilted spherical spectrometer mirror. Two exit slits are located at compromise focal positions and cover the total scan of 2900Å in two bands with a central overlap of approximately 1300Å. Spectral scan in cam operated and grating position is indicated by a marker generator related to cam rotation. The absolute response function of the flight unit is established by the simultaneous exposure of the calibrated photomultipliers (PMT) and the UVR to the same monochromatic beam of collimated light. Analyses of spectra recorded in the laboratory indicate performance consistent or better than the 3Å equivalent width of the UVR exit slits.

I. INTRODUCTION

The UVR was developed to measure stellar spectra in the wavelength interval 1100Å to 4000Å. The Kollmorgen Corporation of Northampton, Massachusetts designed and fabricated the instrument under the direction of the Goddard Space Flight Center of the National Aeronautics and Space Administration. The ray tracing techniques utilized in developing the optics system were presented to the Optical Society of America in 1963.¹

II. DESCRIPTION OF THE UVR

The UVR is intended to be mounted parallel to the longitudinal axis of the 38 cm diameter Aerobee rocket. The basic instrument is approximately 33 cm in diameter, 163 cm long and weighs 45.5 kg. As a flight unit with nose cone and miscellaneous items the weight is increased by roughly 12 kg and the length by 115 cm. It consists of a spectrometer and the largest aperture collecting optics consistent with the rocket diameter and the necessary support structure. The UVR is illustrated to scale in Figure 1.

*Published as *Goddard Space Flight Center Document X-613-64-333*, November 1964; also published in the supplement to the *Japanese Journal of Applied Physics*, May 1965.

In order to minimize the number of surfaces the plane grating is illuminated in convergent light resulting in a comatic image.² Corrective coma is introduced into the convergent beam by tilting the spherical spectrometer mirror. Both comae are linear function of wavelength. The lower "V" shaped plot in Figure 2 is the result of correcting the grating coma at midwavelength by mirror tilt. The grating coma dominates at the higher wavelengths and the coma due to the tilted mirror at the lower wavelengths. In the UVR, an $f/11.9$ system, the mirror tilt necessary to minimize the aberrations at mid-wavelength, 2550Å, is computed to be 2.3744°.

The spherical aberration contributed by the spectrometer mirror requires that the collecting optics be over-corrected. The result is a classical D-K telescope with its secondary shifted out of position longitudinally, see Figure 1. The final position of the telescope secondary is 0.584 cm from its D-K position toward the primary and causes the telescope focal point to shift 3.233 cm to the plane of the entrance slot. Under these conditions the telescope is over-corrected by the proper amount and sign necessary in order to minimize all aberrative contributions the UVR makes to line width.

The UVR will normally utilize the two exit slits

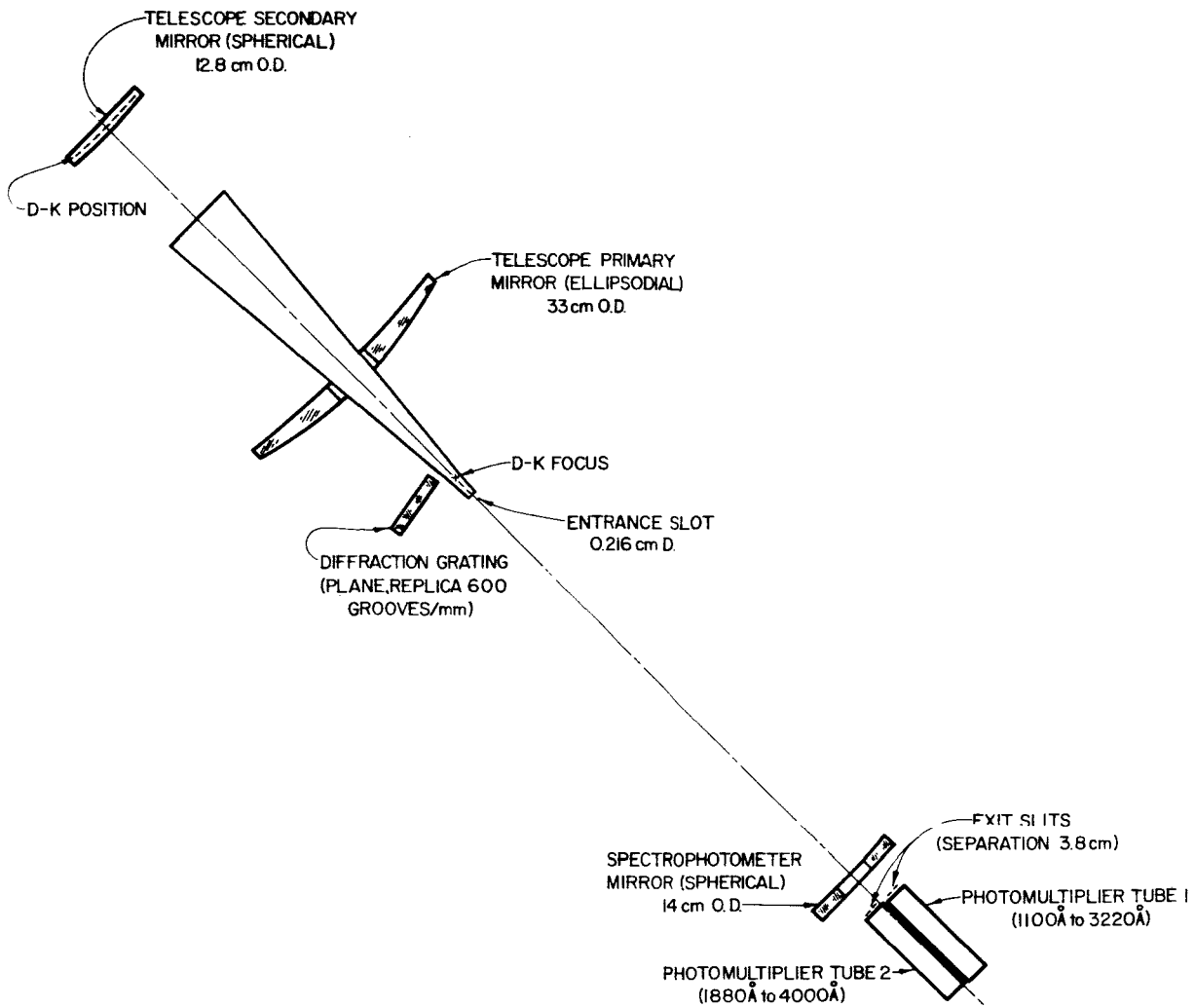


FIGURE 1.—UVR optical layout.

indicated in Figure 1. The slits are located at compromise focal positions and cover the total scan of 2900\AA in two bands with a central overlap of greater than 1300\AA . The exit slit width was chosen for the flight units to be equivalent to 3\AA or approximately 0.015 cm .

The UVR is thus essentially "exit slit limited" and the performance indicated by the respective exit slit curves in Figure 2 cannot be realized. Except as indicated by the reference to "exit slit limited," Figure 2 is therefore the computed performance of the flight units. Actually, if the grating drive could be slowed down enough, it should be possible to decipher the contribution that the UVR makes to line width.

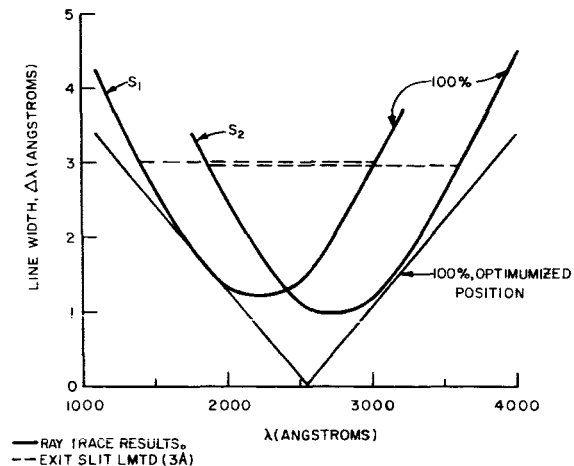


FIGURE 2.—Computed UVR performance.

Figure 3 illustrates the focal surfaces to which an appropriate emulsion must be formed in order to use the UVR as a spectrograph. The relationship of grating positions and the respective focal surfaces to the exit slits indicate that some advantage can be taken of the fact that all of the focal surfaces have a radius of 29.21 cm. In the nominal or B_0 position of the grating the optimized wavelength, 2550\AA , would fall on the optical axis and the emulsion should record the wavelengths from roughly 2000\AA to 3000\AA . Also take note in Figure 3 of the focal surface for zero order image.

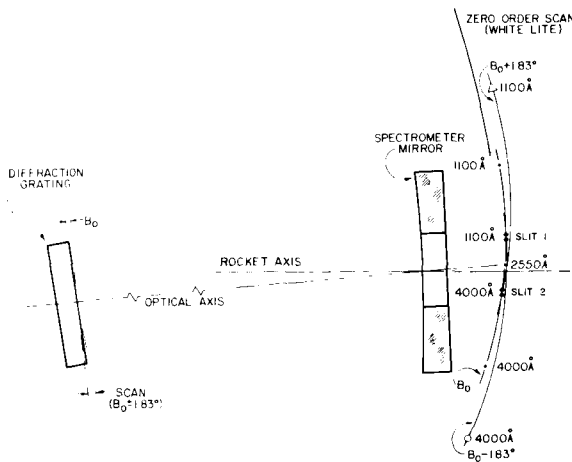


FIGURE 3.—UVR focal surfaces.

III. DEMONSTRATED PERFORMANCE OF A UVR FLIGHT UNIT

The Majority of the stars which are visible to the human eye from the ground have effective temperatures whose wavelength of maximum radiance is below 3000\AA . The earth's atmosphere is opaque to radiation of wavelength shorter than 3000\AA . Observations at the shorter wavelengths are therefore necessary in order that theoretical progress can be made.

The UVR, as a payload for the Aerobee, can spend about 300 seconds above 100 km. The entire 300 seconds are not available for observation nor is it desirable to expend the large amount of effort for one observation. It is of great importance to be able to observe several stars of choice. Therefore, the UVR flights will have programmed guidance and do require the same. The spectra of one or more, probably two, stars of 3rd magnitude

or brighter will be the targets of immediate inquiry.

The flight detectors mounted at slit #1 and #2 are EMR type 542F-08 and 542D-03 respectively. The former has a Cs-Te photocathode and a LiF window while the latter has a Bi-alkali photocathode with a sapphire window.

Figure 4 is a reproduction of Cd spectra and the related marker generator signals recorded using the flight detectors, laboratory electrometers and

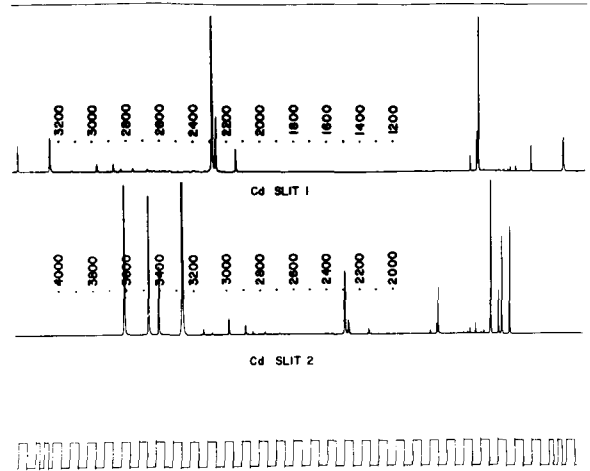


FIGURE 4.—Recorded spectra of cadmium.

recorder. Spectral scan is cam operated with the angular position of the cam indicated by a marker generator physically related to the cam. Proceeding from a point midway between the two pulses of short duration on the left of Figure 4 to the corresponding point in the pulse train on the right is a complete revolution of the cam and took 57.5 seconds. In this interval, 41.3 seconds was consumed by a slow spectral followed by 9° of cam dwell and 13.4 seconds for a fast spectral scan return and its 9° cam dwell. The prominent spectral features indicated from left to right in the record for slit #1 are 3261, 2981, 2881, 2288, 2265 and 2144\AA 's. In the same manner, the spectral features indicated in the record for slit #2 are 3611, 3466, 3404, 3261, 2981, 2881, 2288 and 2265\AA 's. The fast scan in each case is a repetition in reverse of what was recorded during the slow scan. All of the structure indicated has been identified as due to the emission from Cd.

The final calibration and alignment of the UVR is performed within a vacuum calibration tank. The tank houses a 40 cm D-K telescope which collimates the output of a vacuum monochromator. The absolute response function of the flight unit is established by the simultaneous exposure

of calibrated PMT's and the UVR to the same monochromatic collimated beam.

REFERENCES

1. H. R. MCKINLEY, *J. Opt. Soc. Am.*, *53* (1963) 518A.
2. M. V. R. K. MURTY, *J. Opt. Soc. Am.*, *52* (1963) 768.

V. GENERAL ELECTRONICS

DEMODULATION OF WIDE-BAND FREQUENCY MODULATION BY A PHASE-LOCK TECHNIQUE*

RICHARD C. BOOTON, JR.†

The wide-band phase-lock demodulator described herein was developed by the Space Technology Laboratories, Inc. for Goddard Space Flight Center. The design objective was a demodulator capable of handling frequency-modulated television signals. Base-band signals were to be standard 525 line video, with a bandwidth of approximately 3 Mc, plus an aural subcarrier at 4.5 Mc. Mathematical and physical descriptions of the demodulator, together with theoretical and experimental TV data, are presented. The principal design problem overcome during the development was the decrease of the loop time delay sufficiently to achieve a loop bandwidth of approximately 7.5 Mc without stability problems. Monitor pictures are used to compare the performance of the phase-lock demodulator with a conventional discriminator.

INTRODUCTION

With the recent launching of the first Telstar satellite, the first experimental demonstration of telephone and television communication by means of an active satellite repeater system was achieved. The first spacecraft for this purpose was of necessity limited in transmitter power capability. Now, and for some time, the communication system design must center around achieving maximum information flow per watt of satellite radiated power, even at the expense of another precious quantity—bandwidth.

The modulation technique chosen for the early satellite relays has been wide-band analog frequency, or phase, modulation. This encoding has the capability of yielding high output signal-to-noise ratio with small spacecraft transmitted power. However, full advantage cannot be taken of the power reduction capability of this modulation form with a standard-discriminator demodulator. Accordingly, considerable effort has been devoted to the development of more efficient demodulation schemes. Two examples are the Bell Telephone Laboratories FM-feedback

demodulator and the phase-lock demodulator to be described in the present paper.

The phase-lock demodulator was developed by the Space Technology Laboratories, Inc. for Goddard Space Flight Center. The design objective was a demodulator capable of handling frequency-modulated television signals. Base-band signals were to be standard 525 line video, with a bandwidth of approximately 3 Mc, plus an aural subcarrier at 4.5 Mc. A demodulator capable of handling such a TV signal can handle between 120 and 240 frequency-division-multiplex telephone channels.

Evaluation of several methods that reduce the required signal, including the FM-feedback discriminator, led to the selection of the phase-lock technique. The use of phase-lock circuits has yielded extensive experience in different applications. Much of the STL experience was obtained in connection with the design of receivers for satellites and deep space; such applications usually require extremely sensitive receivers with very narrow bandwidths. Although wide-band demodulation was different, the phase-lock technique was sufficiently promising to use for the development of a wide-band demodulator. The prime advantage of the phase-lock technique is the extreme simplicity of the circuits, which makes possible low loop time delays and, accordingly, high loop bandwidths.

*Published as *NASA Technical Note D-1680*, June, 1964. Presented at the National Telemetry Conference, May 1962, Washington, and published in the *Proceedings*, 2(6):paper 6-4. Worked performed for Goddard Space Flight Center under RELAY contract NAS 5-1302.

†Director, Communication Laboratory, Space Technology Laboratories, Inc., Redondo Beach, California.

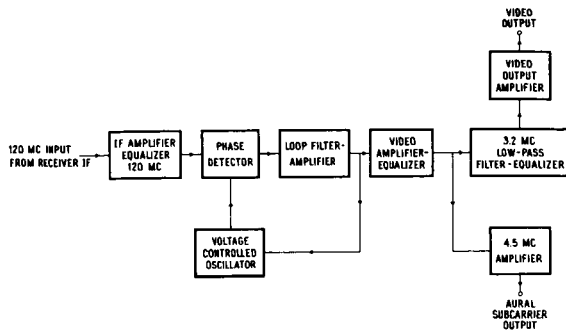


FIGURE 1.—Wide-band phase-lock demodulator for TV.

MATHEMATICAL DESCRIPTION

Figure 1 shows an abbreviated block diagram of the wide-band phase-lock demodulator developed to handle TV signals. The phase-lock loop itself is shown to the left of the figure and consists of three units: a voltage controlled oscillator (VCO) that generates a replica of the signal input, a phase detector to measure the phase relation between the signal input and the VCO output, and a third unit for amplification and filtering to control the loop characteristics. Because of the feedback action the phase of the VCO output follows the phase of the signal input, at least within the dynamic capability of the loop.

The dynamic characteristics selected for the loop were of the second order, although higher order loops have been investigated. Figure 2 gives a mathematical description of a linearized second order loop. The difference between the input phase and the VCO output phase is determined by a phase detector, whose action is indicated as a simple phase subtraction because the nonlinear effects of the actual phase detector

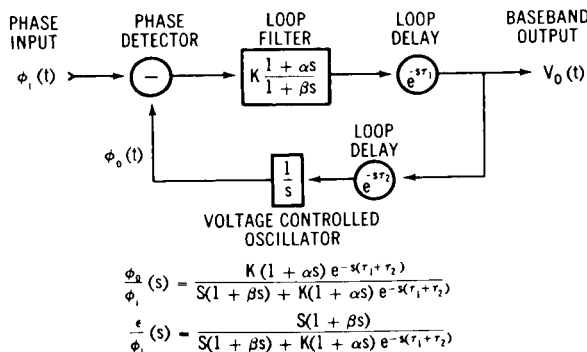


FIGURE 2.—Linearized second order phase-lock loop.

have been neglected. The loop filter is shown as a linear block that furnishes gain and a simple compensation to insure a satisfactory closed-loop characteristic. The principal effect of the VCO is indicated as an integration from the voltage input to the phase of the VCO output. In addition to the usual second order parameters, pure time delays are shown both in the forward path and in the feedback path. Although these delays are shown as pure time delays, they might in fact be a convenient approximation to higher order effects whose characteristic frequencies are so high that the dominant effect over the loop bandwidth is simply a phase shift that is linear with frequency. The two equations in Figure 2 show the mathematical form for the transfer function from the phase input to the VCO output, and the transfer function from the phase input to the phase error.

If it were possible to eliminate the time delays, achievement of any value of loop bandwidth would be a simple matter of selecting appropriate values for the constant K , α , and β . The presence of the time delays τ_1 and τ_2 , however, restricts the bandwidths that can be achieved satisfactorily. Figure 3 is a Nyquist diagram

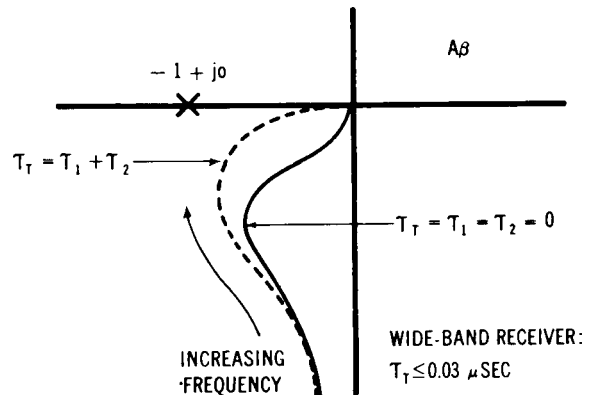


FIGURE 3.—Nyquist diagram; stability criteria.

for a circuit such as that shown in Figure 2, and indicates the tendency of the time delays to cause system instability. For a fixed total time delay, attempting to increase the gain K beyond a certain value results in an unstable system. Analysis indicated that, in order to achieve the desired loop bandwidth of approximately 7.5 Mc, the

total time delay had to be reduced to approximately 30 nanoseconds.

PHYSICAL DESCRIPTION

First, it should be noted that an appreciable amount of conventional receiving equipment is required in addition to the demodulator. Figure 4 is the block diagram of a complete receiving system. The output of the antenna-mounted preamplifier, mixer, and IF amplifier is the input to the rack-mounted portion of the receiver. The rack-mounted equipment includes the demodulator, amplifiers, filters, and equalizers. A standard FM discriminator is also provided. A decision was made to mount all this equipment on one chassis; Figure 5 shows an artist's conception of how the receiver would appear physically. Note that the phase demodulator itself consists of only two modules. Figure 6 is a photograph of the actual receiver as completed. Figures 7 and 8 respectively show the physical appearance of the phase detector and filter, and the VCO. As an illustration of the other modules, Figure 9 is a photograph of an IF amplifier.

EXPERIMENTAL MEASUREMENTS WITH SINUSOIDAL BASE-BAND SIGNAL

Measurements were made to determine the extent to which the performance of the phase-lock demodulator could be predicted by theoretical means. This comparison of experiment and theory should indicate the accuracy to which the loop parameters are known. A simple situation was chosen to simplify the analysis; Figure 10 shows the experimental setup. A sinusoidal base-band signal of 50 kc was used to frequency modulate a 120 Mc IF signal to a peak-to-peak deviation of 20 Mc. Noise was added to this FM signal, and the sum was amplified and applied as the input to the phase-lock demodulator. The noise in a 1 Mc band was measured with a low pass filter and a true rms meter. The signal component was measured with a harmonic analyzer of 6 cps bandwidth. The experimental results are shown by the circles in Figure 11.

A quasilinear analysis* has been applied† for

*Boonton, R. C., Jr., "The Analysis of Nonlinear Control Systems with Random Inputs," in *Proc. of the Symposium on Nonlinear Circuit Analysis, Polytechnic Institute of Brooklyn*, 23-24 April 1953, p. 369-391.

†Application of the quasilinear analysis to the phase-lock loop was suggested and performed by J. A. Develet, Jr.

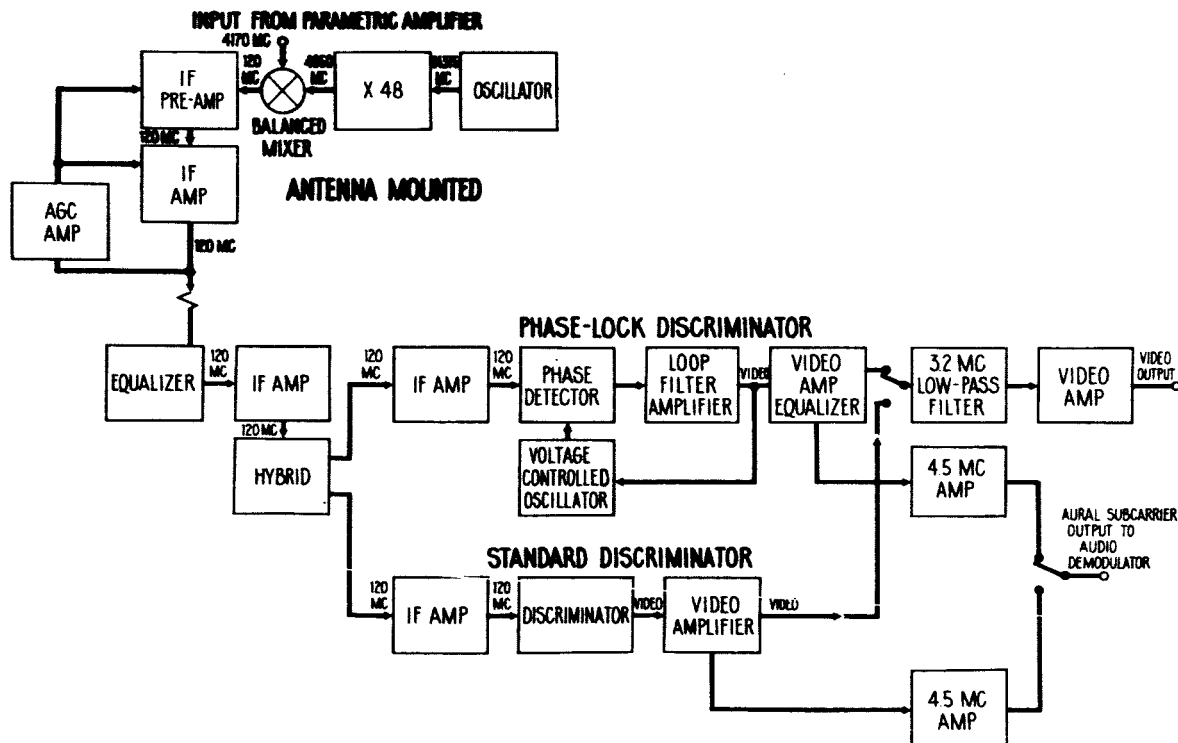


FIGURE 4.—Relay wide-band phase-lock receiver, block diagram.

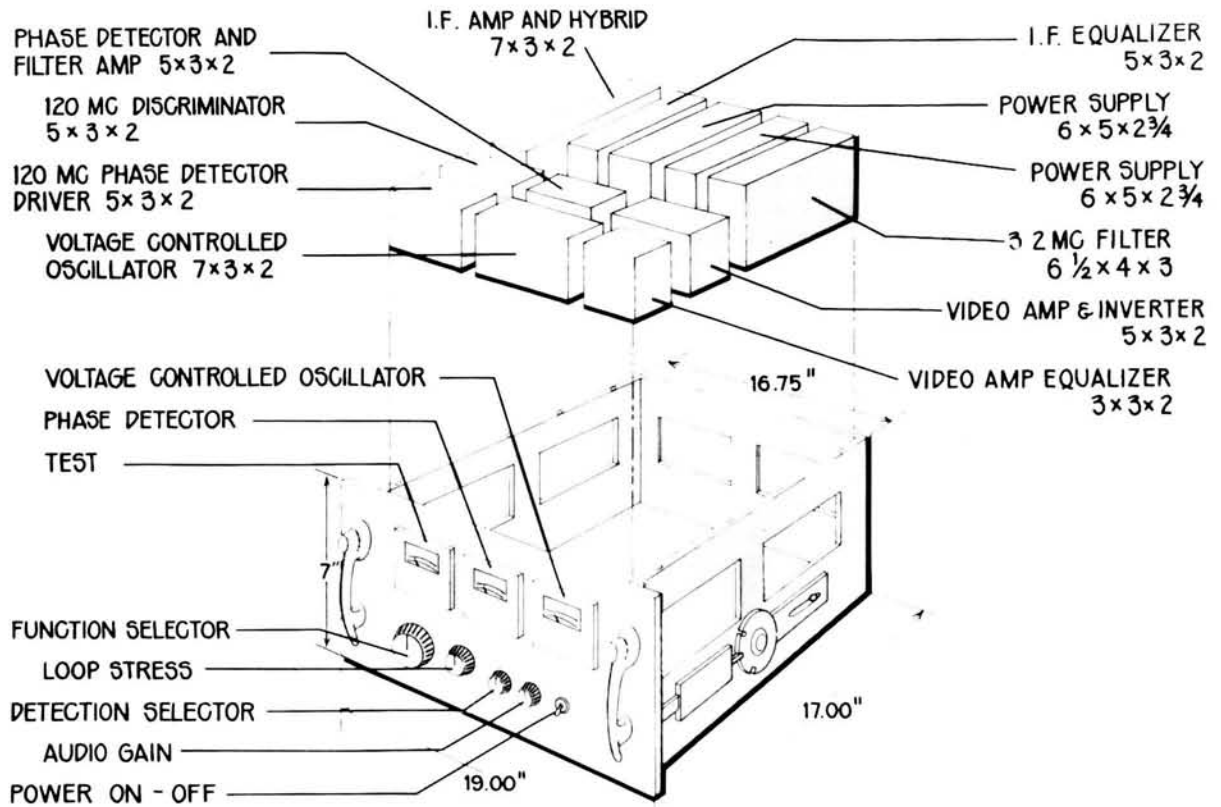


FIGURE 5.—Artist's conception of complete receiver.

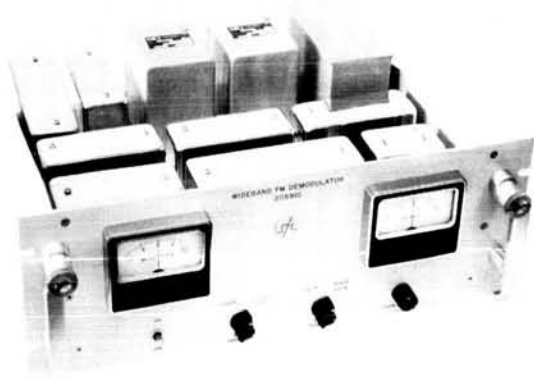


FIGURE 6.—Actual receiver.

an approximate evaluation of the expected signal-to-noise ratio in the output of the phase-lock demodulator; this analysis can be applied down to an input signal-to-noise ratio of approximately

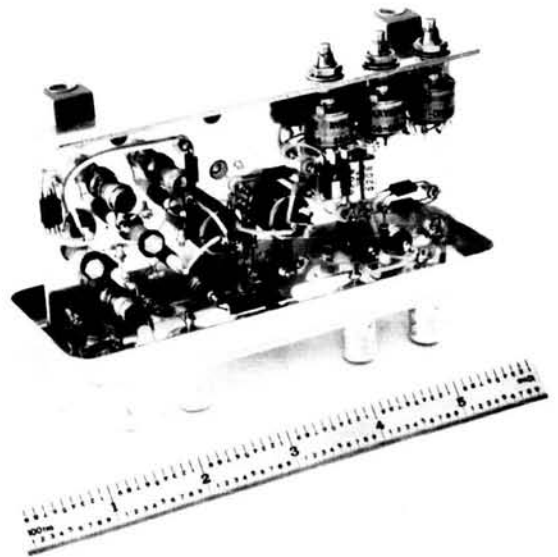


FIGURE 7.—Phase detector and filter.

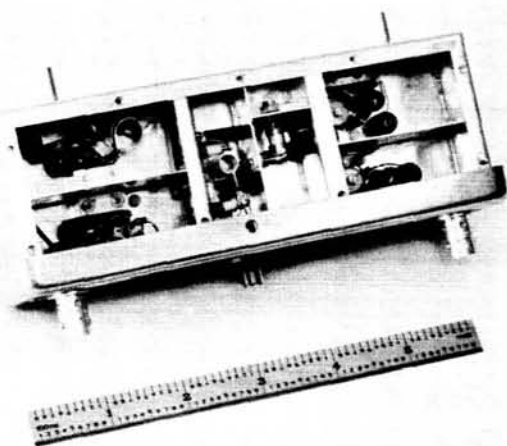


FIGURE 8.—Voltage controlled oscillator.

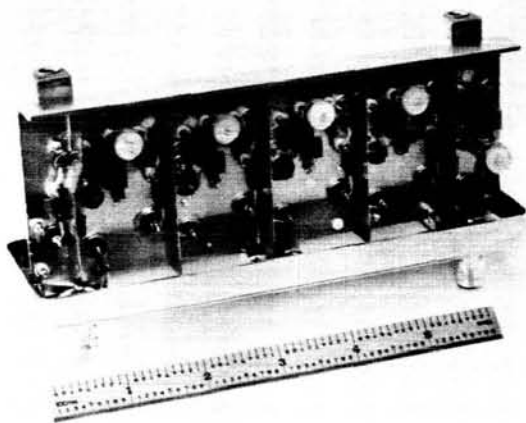


FIGURE 9.—IF amplifier.

0 db. The quasilinear analysis indicates that this value of input should correspond to a threshold in some sense. The analysis is not applicable below this point, but the output signal-to-noise ratio should decrease rapidly—as it does experimentally. For comparison, Figure 11 also shows the expected signal-to-noise ratio output from a standard discriminator with an input identical to that of the phase-lock demodulator.

EXPERIMENTAL TV DATA

The present phase-lock demodulator was designed principally for the demodulation of TV signals. Extensive measurements have been performed to determine the IF signal-to-noise ratio required to give satisfactory picture quality. As a reference point, the output of the phase-lock demodulator is compared with the output of a conventional FM discriminator. Figure 12 shows the experimental setup for such a comparison. A closed-circuit camera, together with the associated television circuitry, derives a video signal used to frequency modulate a carrier at the normal IF frequency of 120 Mc. Standard program material transmitted by a Los Angeles TV station has also been used. Carefully shaped and measured wide-band noise is added to the frequency-modulated signal, and the sum is amplified and limited at the IF frequency. The noise bandwidth is the minimum bandwidth capable of satisfactory transmitting the signal. The output of the IF amplifier limiter is an accurate simulation of the signal plus noise that

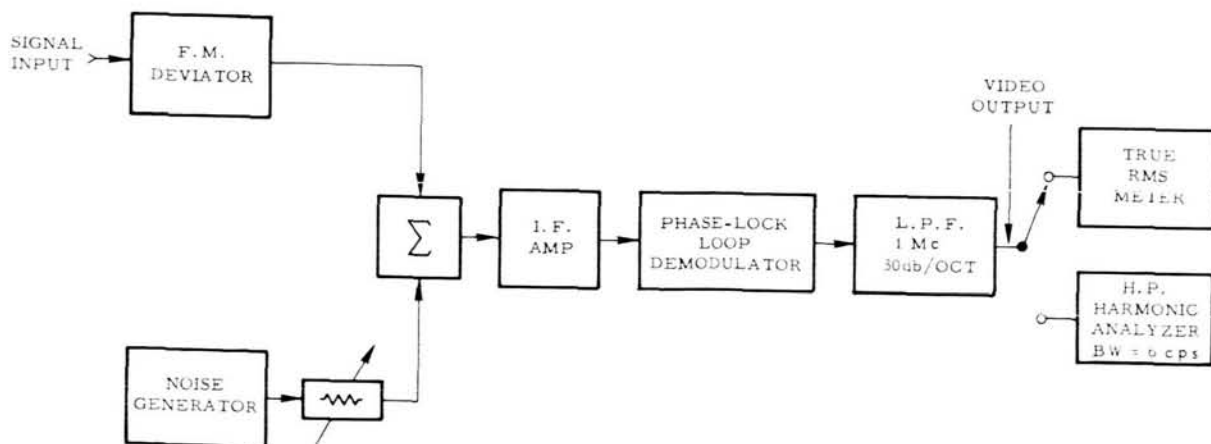


FIGURE 10.—Experimental setup for sinusoidal signal measurement.

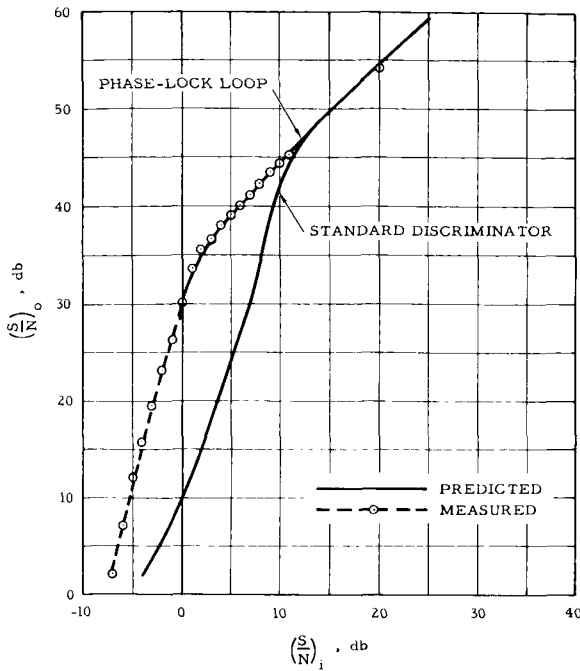


FIGURE 11.—Experimental measurement of signal-to-noise ratio.

would appear at the input to the demodulator in a communication satellite experiment. This simulated IF signal is used to drive in parallel a phase-lock demodulator and a conventional discriminator, each of which is used to control a monitor. The quantity of direct interest is the IF signal-to-noise ratio required to give satisfactory performance.

Figure 13(a) compares the output of the two monitors when no noise is added to the signal. Figure 13(b) compares the two monitor pictures with a 12-db IF signal-to-noise ratio (S/N) and illustrates that at high signal-to-noise ratios the performance of the phase-lock demodulator and

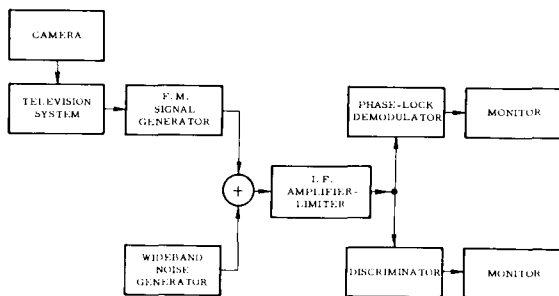


FIGURE 12.—Experimental setup for TV measurements.

the standard discriminator are approximately the same. In Figures 13(c) through (i) the two monitor pictures are shown with $S/N = 8$ db, 4 db, 2 db, 0 db, -1 db, -2 db, and -3 db respectively. As the signal-to-noise ratio is decreased, both monitors indicate a steadily worse picture quality with this important difference: The standard discriminator output degrades more rapidly than the phase-lock demodulator output. Figure 13(c) for example, illustrates the situation for $S/N = 8$ db, which is just below the threshold of 10 to 12 db commonly assumed for a standard discriminator. The discriminator output is beginning to show the effects of the noise, whereas the phase-lock picture is essentially unaffected by the noise. Comparison of the monitor pictures for the two demodulators involves a subjective comparison because of the somewhat different nature of the noise on the picture (described below). Most observers seem to agree, however, that threshold conditions for the two demodulators occur with the standard discriminator requiring approximately 5 db more signal-to-noise ratio; for example, the phase-lock monitor output for the -1 db signal-to-noise ratio has approximately the same picture quality as the standard discriminator output with a signal-to-noise ratio of +4 db.

The noise output of the standard discriminator gives a picture of uniform grain. The principal effect of noise on the phase-lock demodulator, however, is strikingly different. At random intervals the phase-lock demodulator will lose lock. After loss of lock the base-band output tends to go to full black or full white for a very short interval. After this short interval the demodulator locks up again. The losses thus appear in the monitor picture as short black and white horizontal lines.

CONCLUSIONS

The design effort discussed in this paper was directed toward the development of a discriminator that (1) could be used with a high deviation frequency-modulated signal with a base-band signal consisting of a 525 line video plus a 4.5 Mc aural subcarrier, and (2) would reduce the threshold by approximately 6 db below the threshold of a standard discriminator. The principal design problem overcome during the development was the decrease of the loop time delay suffi-

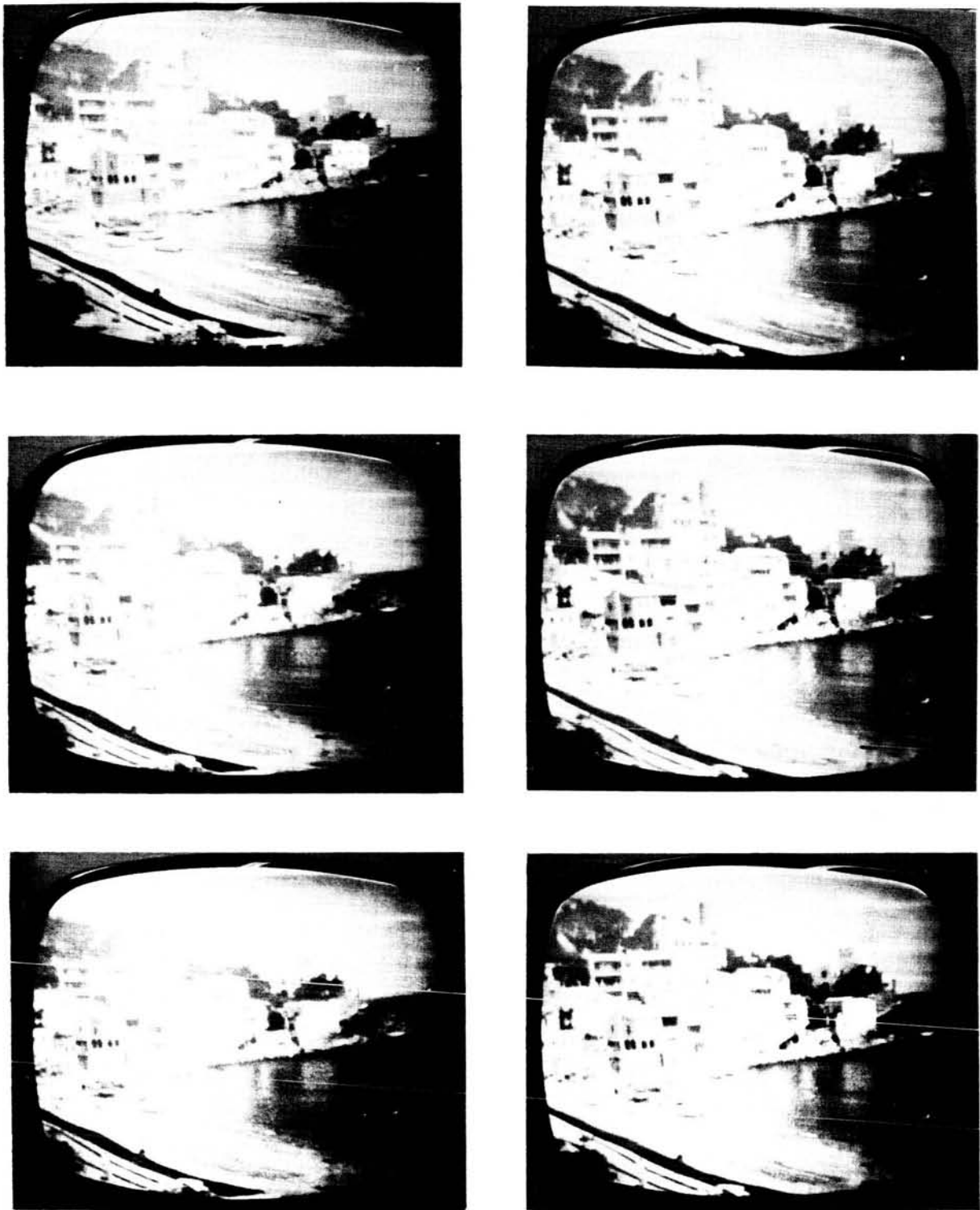


FIGURE 13.—Monitor pictures with a conventional discriminator (left-hand view) and a phase-lock demodulator (right-hand view).

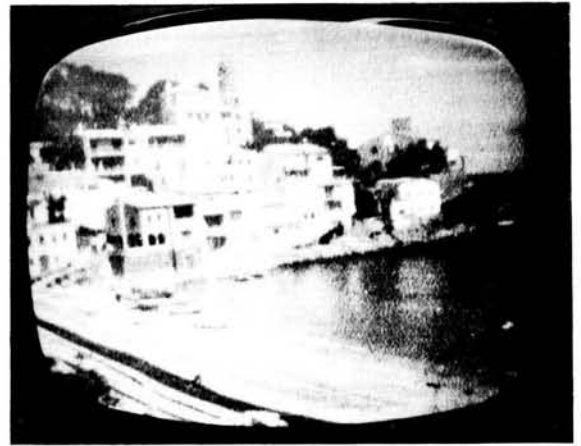
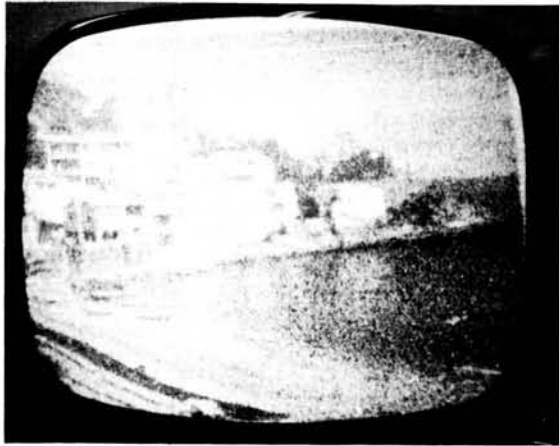
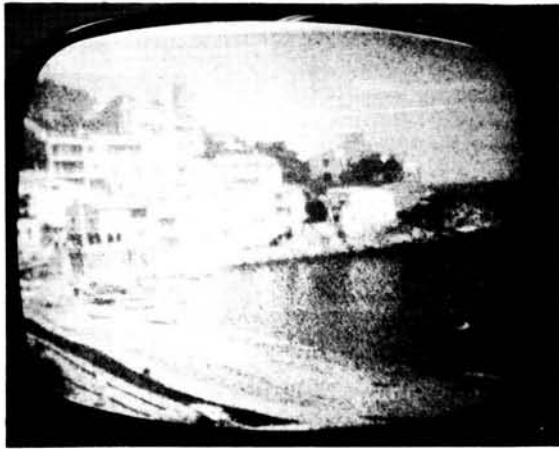


FIGURE 13.—(continued)

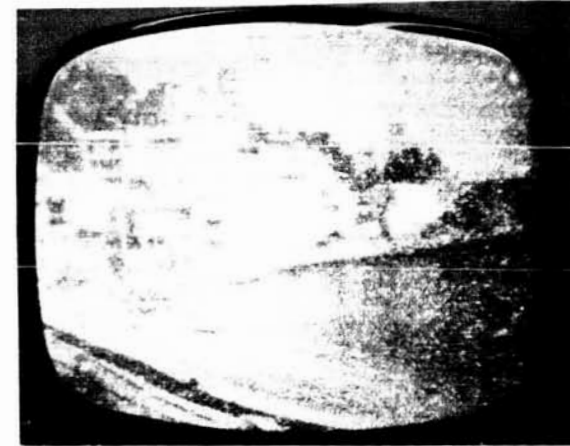
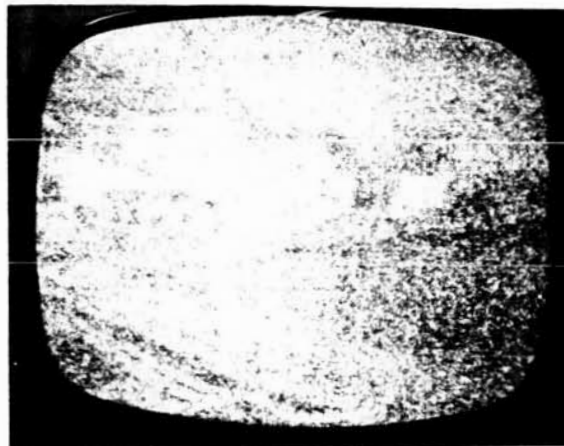
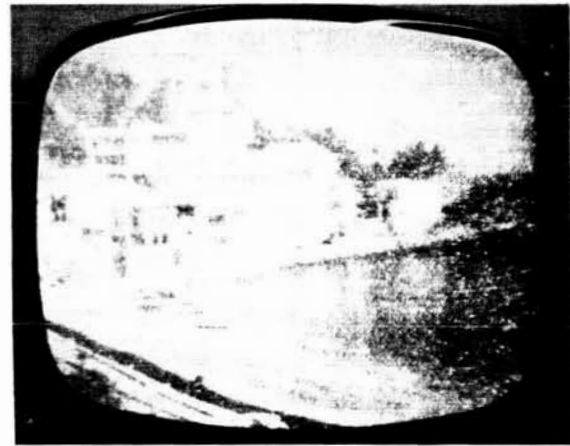
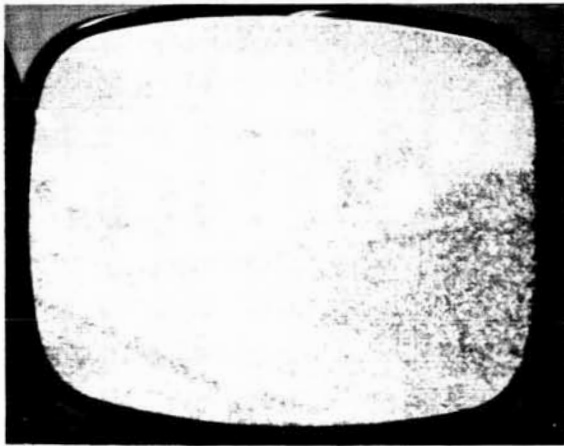
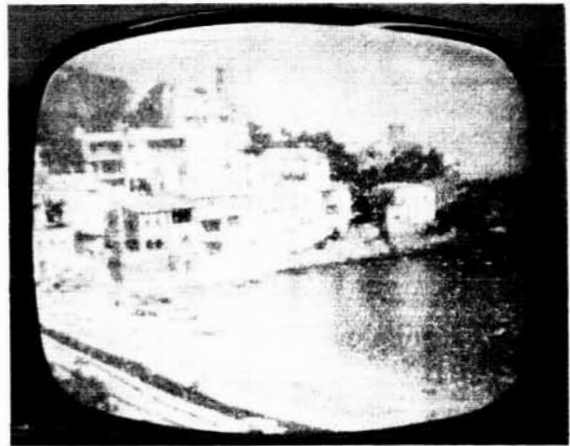


FIGURE 13.—(continued)

ciently that a loop band-width of approximately 7.5 Mc could be achieved without stability problems. This problem was solved satisfactorily by reduction of the unit time delays to the point where the total time delay is less than 30 nanoseconds. Although extensive measurements have been made to verify the dynamic properties of the loop, perhaps the best illustration that the development was satisfactory lies in the quality of the monitor pictures shown in this paper.

ACKNOWLEDGMENTS

The support and encouragement of the Goddard Space Flight Center, specifically D. G. Mazur and R. J. Mackey, are gratefully acknowledged. Special appreciation is due the members of the STL Communication Laboratory, in particular J. A. Develet, Jr., C. W. Stephens, K. H. Hurlbut, and W. A. Garber, for the development described in this paper.

HIGH EFFICIENCY COMPLIMENTARY ASTABLE PRODUCES PULSES OF EITHER POLARITY*

SERGIO CANTARANO†

Built around a complementary pair, the oscillator shown in Figure 1 has several interesting features:

1. It is able to generate pulses at very low repetition rate using a comparatively low capacitance, the limitation on the timing resistor being $R \leq \phi_1 \phi_2 R_c$,
2. Sharp positive pulses are available at the Q_1 collector, negative pulses at the Q_2 ,
3. The efficiency of the circuit is high, both transistors being simultaneously ON only during the short time of the pulse. This makes it possible to use a low R_c value, with low power consumption, obtaining a low output impedance in both the "1" and "0" states.

The circuit operates as follows: the two transistors are both biased in the active region by V_{cc} through $R+R_B$, provided that $R+R_B < \phi_{min} R_c$ to avoid saturation. This is an unstable state: regenerative action takes place and oscillations start. The pulse width τ , during which both transistors are saturated, is determined by the discharge of C through R_B+R_{in} , R_{in} being the input resistance of the transistor. Such discharge supplies the base current necessary to sustain the saturation. The OFF time T , determining the repetition rate and during which both transistors are OFF, depends

on the time constant RC . Resistor R_B limits the base current and controls the pulse width.

Representative circuit parameters are:

$R = 4.7 \text{ M}\Omega$	$C = 0.22 \text{ }\mu\text{F}$	$V_{cc} = 6 \text{ volts}$
$R_c = 3.9 \text{ K}\Omega$	$C_1 = 2\text{N}2604$	$\tau = 10 \text{ msec}$
$R_B = 12 \text{ K}\Omega$	$C_2 = 2\text{N}2623$	$T = 6.75 \text{ sec}$

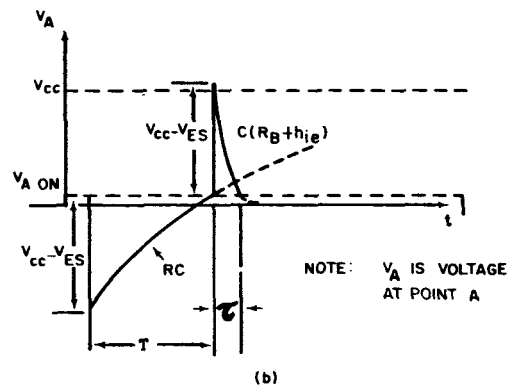
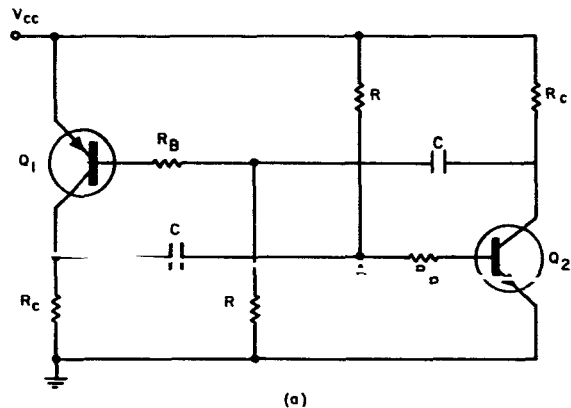


FIGURE 1.

*Published in *Electronic Design*, September 28, 1964.

†Goddard Space Flight Center NASA-National Academy of Sciences-National Research Council Senior Post-Doctoral Resident Research Associate on leave of absence from National Council of Research, Rome, Italy.

VOLTAGE AND CURRENT MONITOR FOR SCIENTIFIC SATELLITES*

WARREN R. CROCKETT

INTRODUCTION

This paper gives a detailed description of a voltage and current monitor applicable to spacecraft technology comprised of a self-balancing magnetic-amplifier circuit using d. c. bias. This circuit is capable of monitoring voltages of full-scale ranges from 1 to 3000 volts and currents ranged from 1 milli-ampere to greater than 10 amperes with an accuracy of ± 0.5 percent. The power consumed may vary from less than one milliwatt to approximately eight milliwatts depending upon the load required. In spacecraft applications the telemetry system may present a load varying, 1,000 ohms to 20,000 ohms and under these conditions the power consumed by the monitor circuit is in the order of eight milliwatts.

In spacecraft applications it is desirable to monitor several potentials such as the main-voltage converter potential, the transmitter potential and other vital spacecraft potentials that may be of interest. These potentials may vary in magnitude from a few volts to one hundred volts depending upon the size and mission of the spacecraft.

The spacecraft currents to be monitored may vary from a small trickle charge of a few milli-amperes to the solar array current or the spacecraft main bus current. These currents can be in the order of five amperes.

There are many techniques available to the design engineer to perform the above task. For example, the voltage monitor could be a simple voltage divider circuit, and the current monitor could be a Hall-Effect generator or a saturable reactor. None of these however, are adequate to cover the wide dynamic ranges encountered with

good linearity. A technique that can fulfil this requirement is the self-balancing magnetic-amplifier. This circuit is capable of accepting a d. c. or a. c. voltage or current signal over a wide dynamic range and producing a d. c. voltage at its output proportional to the input with an accuracy of ± 0.5 percent. Furthermore, when the d. c. output voltage is connected to an analog oscillator in the spacecraft telemetry, a frequency is produced which is directly proportional to this voltage. This frequency is telemetered to a satellite tracking station and stored on tapes. With the aid of a voltage vs. frequency calibration curve the magnitude of the spacecraft voltage or current being monitored can be determined.

CIRCUIT OPERATION

The self balancing magnetic amplifier circuit is shown in Figure 1. An advantage of using this type of circuit is the tendency to minimize:

1. The variation in the rectifier reverse characteristics.
2. The effects of supply voltage variations.
3. The effects of frequency variations.

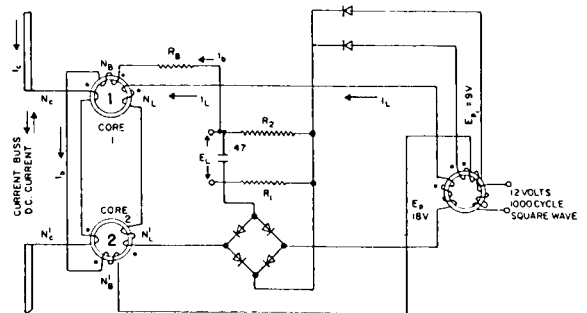


FIGURE 1.—Self balancing magnetic amplifier circuit with DC bias.

*Published as *Goddard Space Flight Center Document X-631-64-35*, February 1964.

The operation of this circuit is such that when the voltage E_p goes positive, Core 1 is driven into positive saturation through load winding N_L and at the same time Core 2 is driven into negative saturation through winding N'_L . This situation is reversed when voltage E_p goes negative. The effect of any discrepancy in the two core characteristics will tend to be reduced to a minimum. The bias current I_b supplied from the power supply E_{p1} through R_B and R_2 resistors is directly proportional to the voltage E_{p1} . The quiescent current I_Q is that current which will flow in the output load windings N_L and N'_L when the control current flowing through control windings N_c equals zero. This quiescent current I_Q is controlled by the bias current flowing in the bias windings N_B and N'_B . These windings act as additional control windings to make the quiescent current I_Q practically independent over a voltage supply E_{p1} variation from ± 5 to ± 10 percent.

CIRCUIT DESIGN

In designing magnetic amplifier circuits, three problems confront the design engineer: (1) Design of the magnetic amplifier, (2) Design of the load into which the magnetic amplifier is to operate, and (3) Design of the power supply necessary to drive the magnetic amplifier.

1. Magnetic Amplifier

Many factors enter into the selection of the proper magnetic amplifier circuits. Some of these are covered in References (1) and (2). When the magnetic amplifier is employed as a d-c instrument and the time of response is not a major problem, one may choose a magnetic amplifier of the self-balancing type with d-c bias. The reasons for this selection are as follows:

- a) An advantage of this circuit is that with properly applied d-c bias; the quiescent current can be made substantially independent of any changes in the supply voltage E_p ; and
 - b) The quiescent current can be decreased to a few micro-amperes. These become important factors when currents in the order of a few milliamperes are to be monitored.
- a. *Operating Point.* In order to reduce quiescent current, we must consider the fact that the Law of Equal Ampere Turns only holds over a

portion of its range. The transformer turns-ratio may be expressed as

$$\frac{I_L}{I_c} = \frac{N_c}{N_L} \tag{1}$$

then

$$I_L N_L = f(I_c N_c) \tag{2}$$

where

I_L = current in the load winding

N_L = the number of turns in the load winding

I_c = current in the control winding

N_c = number of turns in the control winding

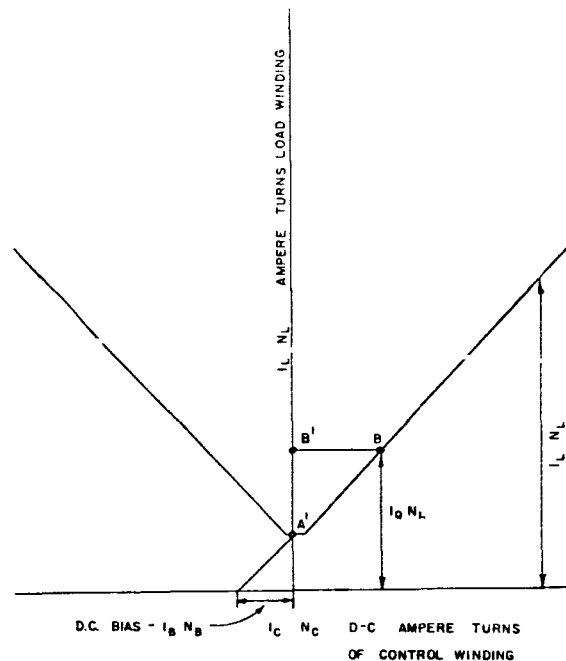


FIGURE 2.—Input vs output curve.

A plot of equation (2) is shown in Figure 2. Point B' determines the quiescent current (I_Q) in the load windings when current in the control windings is reduced to zero. It is this I_Q that causes trouble when low currents are to be monitored. W. Kramer, Ref. 5, has shown that, by adjusting the bias, point B' can be reduced to zero. To illustrate this, consider point B in figure 2 as the quiescent operating point. If a negative d-c bias is applied this point moves to point A' which results in much smaller quiescent

current I_Q . As can be seen from Figure 2, Equation (2) is not linear near A' , and therefore the circuit has very low sensitivity in this region. This nonlinearity becomes a very important problem when currents of a few tenths of a milli-ampere are to be monitored.

When the current to be monitored is 100 milliamperes or more, however the nonlinearity at point A' is so minute that for all practical applications it can be ignored. The quiescent operating point in this particular situation may be determined as follows:

1. Resistors of bias windings (N_B) equal approximately 500 ohms each.
2. The bias resistor (R_b) is 75,000 ohms.
3. Bias voltage (N_B) equals 9 volts d. c.

This gives a required bias current I_b of 0.1 ma which is well below any value of current this circuit is required to monitor. Figure 2 has been expanded in the region of A' for clarity.

b. *Magnetic Cores.* Reference 1 shows that 78 percent nickel—22 percent iron tape-wound core material is required for good operation in magnetic amplifiers with a power supply of 15 to 20 volts. This material has low magnetizing current and consumes less power than the 50 percent nickel—50 percent iron type, and was therefore selected for the current and voltage monitor circuit. Operating parameters for the transformers were selected as follows:

1. A quiescent current of of 0.1 to 0.2 milliamperes as calculated in the previous section.
2. Expected load impedance of 10 to 20 kilohms.
3. Power supply voltage $E_p = 18$ volts to insure that cores are fully saturated and constant output voltage with a possible ± 5 percent variation in supply voltage.

c. *Windings.* In selecting the number of turns for the bias and load windings, the following factors must be considered:

1. The minimum current to be measured by the monitoring circuit.
2. The power supply voltage selected.
3. Core material used.

The minimum current to be measured by the monitoring circuit is important because it determines the lowest possible I_Q . One way to achieve low I_Q is to make N_L and N_b as large as possible. With a given volume for the monitoring circuit, the best way to achieve this is to select a core with a small cross-sectional area and a large inside diameter. This will result in large windings. The dimensions of the core may be obtained from a manufacturer's specification sheet. Maximum possible number of turns is given by the expression

$$N = \frac{W}{KA_w} \quad (3)$$

where

W = core window area, in cm^2 ,

A_w = wire cross-section area including insulation, in cm^2 , and

K = a winding factor determined by the characteristics of the coil winding facilities.

When currents of 1 to 10 amperes are to be monitored, a control winding of 1 turn is sufficient, and we may set N_L and N_b equal to $N/2$. For monitoring lower currents, a good practise is to make $N_c + N_b = N_L$. The size of N_c varies inversely with the current to be monitored as shown in Table I.

TABLE I.

No. turns of N_c	Current Range Amperes
1	1→10
20	.5
100	.1
400	.025
2000	.005
10,000	.001

d. *Voltage Monitoring Operating.* The design for low current monitoring will apply to voltage monitoring applications, that is, N_c will equal N_b , and $N_c + N_b = N_L$. For our particular application N_c and N_b will equal to 1750 turns each, and N_L will equal 3500 turns. The resistance of the

winding N_c is approximately equal to 250 ohms. Now suppose the magnetic amplifier is balanced in such a way as to have a voltage gain of unity, that is $E_L/E_c=1$ where E_L =output voltage and E_c =voltage to be monitored. In this case, if 5 volts is applied to the control winding N_c , 5 volts will be produced across the load resistor. Now, suppose we want this magnetic amplifier to monitor a 25-volt source and still have an output voltage of 5 volts. The procedure is as follows:

1. With E_c set at 5 volts, observe that the current in the control winding N_c is

$$I_c = \frac{5}{250} = 20 \text{ ma} \quad (4)$$

2. With E_c equal to 25 volts, place a resistor R_c in series with the control winding such that

$$R_c = \frac{25}{20} \times 10^3 = 1250 \text{ ohms} \quad (5)$$

With this series resistor, the magnetic amplifier is capable of reading voltages up to 25 volts. This range can be extended to monitor voltages as high as 10,000 volts providing the proper size core, windings, and series resistor are chosen.

2. Load

When a magnetic amplifier is employed as a d-c instrument, it is important to have the load current I_L correspond to the input signal current of the magnetic-amplifier circuit. One such circuit could take the form of a simple voltage mixing circuit as shown in Figures 3. R_1 R_2 are connected across the load winding of Core 1 and Core 2. The two voltages E_1 and E_2 represent the rectified voltage appearing across these windings when Cores 1 and 2 are conducting respectively. The average voltage across R_L is equal to the difference between E_1 and E_2 . When these two outputs are equal, the total voltage applied to the load resistor is zero.

The bias applied to the magnetic amplifier is such that the total m.m.f. controlling Core 1 is equal:

$$I_b N_b = I_c N_c \quad (6)$$

and that controlling Core 2 is

$$I_b N_b = -I_c N_c \quad (7)$$

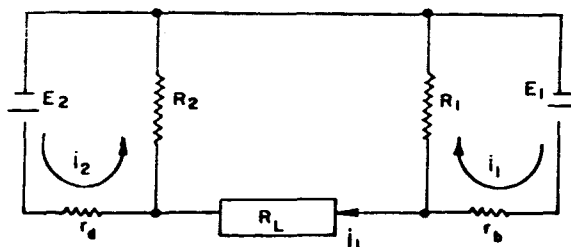


FIGURE 3.—Magnetic amplifier load circuit.

Thus, if the control current is increased from zero, the control ampere-turns on Core 1 increase while those on Core 2 decrease. If the control current is decreased, the effect is reversed.

The analysis of Figure 3 is governed by the amount of bias applied to the magnetic amplifier circuit. A full analytical treatment of this subject may be found in Reference 6.

3. Power Supply

To produce the driving voltage and frequency necessary to supply the magnetic amplifier, a dc-to-ac converter using transistor switching and a square-loop magnetic core was used. This type of circuitry gives a constant volt-second waveform. A great deal of technical literature describing the design and operation of such circuits has been published, and many elaborate circuits have been built verifying this literature. Basically, all of these circuits can be resolved into the circuit as shown in Figure 4. See Appendix for design equations.

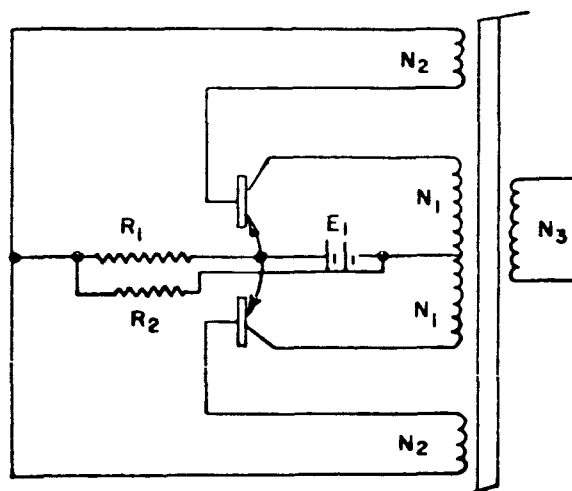


FIGURE 4.—DC to AC converter.

Factors to consider in selecting the magnetic core and transistors are given in the following sections:

a. Magnetic Core:

1. Core size is an inverse function of operating frequency.
2. For low power levels and light loads, 78 percent nickel, 22 percent iron tape-wound core material should be selected.

b. Transistors

1. The reverse collector-to-emitter voltage breakdown should be more than twice E_1 (see Figure 4).
2. Cut-off frequency should be at least 10 times the power supply operating frequency.

DESIGN EXAMPLE

1. Design Procedure

Suppose it is desired to design a current sensor to meet the following requirements:

1. Current to be monitored is (0→5) amperes.
2. A linearity of ± 1 percent.
3. Output impedance 10 K ohms or less.
4. Output is (0→5) v.d.c.
5. A 12 v.d.c. source is available.
6. The current sensor must consume less than 10 milliwatts.
7. Environmental temperature will vary from -20°C to $+60^\circ\text{C}$.

A self balancing magnetic amplifier will be used to monitor this current. Schematic for the entire voltage and current monitor including power supply is shown in Figure 5.

Step 1. Selecting the Core Material

There are two different types of material to choose from.

- (a) 50 percent nickel, 50 percent iron
- (b) 78 percent nickel, 22 percent iron

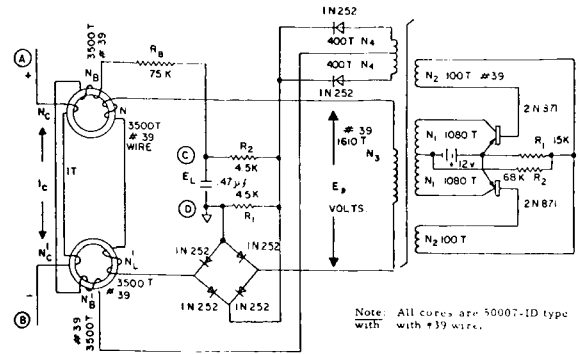


FIGURE 5.—Magnetic amplifier and DC to AC converter circuit.

For low power, small magnetizing current, light loads, and high efficiency the 78 percent nickel, 22 percent iron material is a better choice. Tape-wound cores using 1 mil thick tape were chosen. (See section under magnetic core in this report).

Step 2. Core Size to be Used

There are several factors that influence the core size selection; they are as follows:

- (a) The total weight allotted for the magnetic amplifier.
- (b) Frequency of operation.
- (c) Size wire to be used for each winding.
- (d) The smallest current the magnetic amplifier is required to monitor.

For the interplanetary monitoring platform (Explorer 18) power and space were the prime considerations. The smallest current to be monitored was 100 milliamperes, and the largest was 5 amperes. The frequency of operation was chosen to be approximately 1000 cycles per second. From these considerations the core selected had the following dimensions: $ID = .625''$, $OD = .750''$, height = $.125''$, core cross section area = $.040 \text{ cm}^2$.

Step 3. Selecting the Number of Turns for N_L and N_B Windings

Using equation (3):

$$N = \frac{W}{KA_w}$$

By using tape-wound core catalog for this size core the window area W , and the wire cross section

area Aw are given. These dimensions give the designer some idea of the upper limit for N . By using A.W.G. Size #39 wire, it was possible to put 7000 turns on this size core. Therefore,

$$N = 7000 \text{ turns,}$$

$$N_B + N_L = 7000 \text{ turns,}$$

and

$$N_B = N_L = 3500 \text{ turns}$$

From Table 1, Page 466

$$N_c = 1 \text{ turn}$$

Step 4. D.C. Bias Circuit

From the graph in Reference 1, Page 67, and using 78 percent nickel, 22 percent iron type cores the supply voltage E_p was chosen to be 18 volts. The bias voltage was selected to be approximately one-half of E_p , or 9 volts.

The actual magnetization current for this particular design, where the supply voltage E_p is 18 volts and the load resistor is approximately 10 K ohms, is .12 milliamperes. From this a value of R_B may be computed.

$$R_B = \frac{9}{.12} \times 10^3 \approx 75 \text{ K ohms}$$

Step 5. Load Circuit

Resistors R_1 and R_2 form a voltage mixing circuit for the magnetic amplifier. The values of R_1 and R_2 are selected at approximately 5 K ohms. However more will be said about this in a later section on magnetic amplifier balancing. The .47 μf capacitor is to reduce the ripple voltage that will appear across the load resistor and this value may be increased or decreased depending upon the amount of ripple voltage allowed in the load.

Step 6. Power Supply Design (See Appendix)

The power supply circuit is a dc-to-ca converter with the following characteristics:

1. Input voltage is 12 volts dc.
2. Frequency of operation 1000 cps.

3. Core material: 1 mil thick tape, 78 percent nickel, 22 percent iron.
4. Core size: ID = 0.625", OD = 0.750", height = 0.125", core cross section area, 0.040 cm^2 .
5. Load current approximately 4 milliamperes.

Now using equation (8)

$$N_1 = \frac{12}{4(7000)(.04) \times 1000} \times 10^8 =$$

$$N_1 \cong 1080$$

N^2 is chosen such that it will supply one volt to the base of the transistor. Then from equation (10)

$$N_2 = N_1 \frac{1}{E_1}$$

$$N_2 = 1080 \frac{1}{12}$$

$$N^2 \cong 100 \text{ turns}$$

Then N_3 is found by the desired output voltage. In our case we have chosen 18 volts; and from equation (11)

$$N_3 = N_1 \frac{E_0}{E_1}$$

$$N_3 = 1080 \frac{18}{12}$$

$$N_3 = 1610 \text{ turns}$$

Then

$$V(N_2) = \frac{N_2}{N_1} E = 1.2 \text{ volts}$$

So R_1 and R_2 may be found by using equations (13). Thus,

$$R_1 = \frac{1.2}{\frac{1610}{1080} \times \frac{4 \times 10^{-3}}{100}}$$

$$R_1 \cong 15 \text{ K}$$

and R_2 becomes

$$R_2 = \frac{(12)^2 \times 100}{.25}$$

$$R_2 \cong 66 \text{ K ohms}$$

2. Balancing Procedure

A magnetic amplifier of the type described above is balanced when equation (2) $N_L I_L = F(N_c I_c)$ produces a linear curve.

Referring to Figure 5, the balancing procedure is as follows:

- (a) Place a 5 ampere d-c current supply across terminals (A) and (B).
- (b) Place a 4 place digital voltmeter across terminals (C) and (D).
- (c) Replace R_B , R_2 , and R_1 with decade resistors.
- (d) Set the 5 ampere current supply to zero.
- (e) Set R_1 and R_2 to 5000 ohms and set R_B to 75000 ohms. Increase R_B until the voltage across (C) and (D) is a minimum.
- (f) Increase the current across terminals (A) and (B) to 5 amperes. Adjust R_1 and R_2 until the voltage across (C) and (D) reads 5 volts.
- (g) Reduce the current supply to zero and note the zero reading. If the previous zero reading has changed, increase R_B until a minimum voltage appears across (C) and (D).
- (h) Again increase the 5 ampere supply to 5 amperes and adjust R_1 and R_2 until 5 volts is read across (C) and (D).
- (i) Reduce the current supply to 4 amperes and note the voltage across (C) and (D). If the digital voltmeter does not read 4 volts adjust R_1 and R_2 until 4 volts is read on the voltmeter.
- (j) Repeat steps (h) and (i) until the error is reduced to a minimum.
- (k) When (j) is satisfied the magnetic amplifier is balanced and equation (2) satisfied.

With a little practice this balancing procedure will become routine.

3. Environmental Test and Results

The voltage and current monitoring circuit shown in Figure 5 was assembled and exposed to the following tests:

1. Thermal-vacuum test
 - (a) 24 hours soak at -30°C and 1×10^{-5} mm Hg
 - (b) 24 hours soak at $+70^\circ\text{C}$ and $1 = 10^{-5}$ mm Hg
2. Acceleration test
 - (a) 49.4G thrust load combined with 13.6G radial load due to spin of 206 RPM with radius = $11\frac{1}{4}$ "
3. Vibration
 - (a) Random all axis 20–3000 cps 54G
4. Temperature variation -50°C to $+70^\circ\text{C}$
5. Magnetic flux radiation with 5 amperes flowing in the control winding N_c . The Magnetic flux radiated from this circuit is less than 2 gamma at 18 inches.

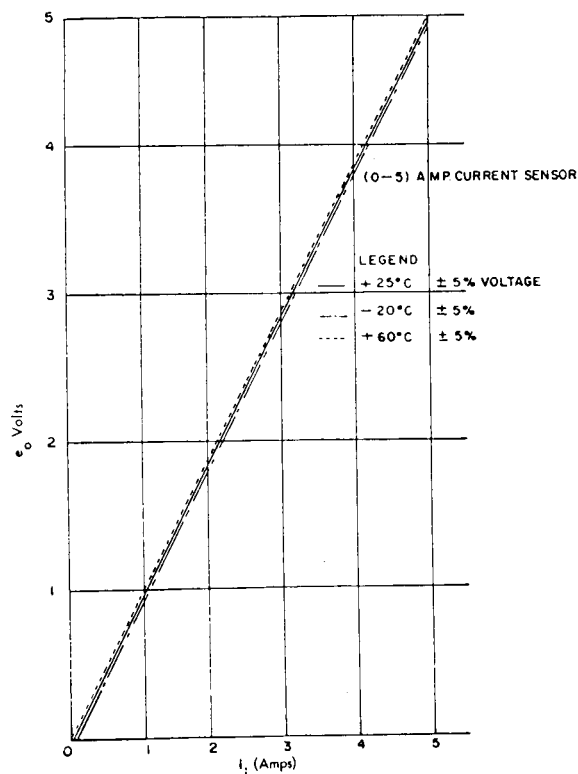


FIGURE 6.—Effects of temperature on output voltage.

Figure 6 shows the results of a temperature variation from -20°C to $+60^{\circ}\text{C}$, and a supply voltage variation of ± 5 percent. It can be seen that temperature and voltage variations have little effect on the current monitoring circuit.

The total power required for the circuit is 18 milliwatts. 10 milliwatts are required for the power supply, and 8 milliwatts for the magnetic amplifier.

Following satisfactory operation during tests, the circuit of Figure 5 was flown on the interplanetary monitoring platform Explorer 18 satellite. This satellite was launched from Cape

Kennedy 26 Nov. 1963. A modified version of this circuit will be flown on the United Kingdom Satellite.

CONCLUSION

The circuit described in this report offers a very good solution for measuring currents and voltages of any magnitude with a high degree of linearity and accuracy. Currents as low as one micro-ampere and voltages as high as ten thousand volts have been measured by this method. (see References 1 and 5). From the quality of data received, to date, from Explorer 18 Satellite, tends to verify this conclusion.

APPENDIX

POWER SUPPLY DESIGN

The design equations used to develop the circuit of Figure 4 are indicated below.

$$N_1 = \frac{E_1}{4B_m A F} \times 10^8 \quad (8)$$

Where E_1 is the applied d-c voltage B_m and A are flux density and effective cross-section of the core. F is frequency of operation. N_2 is determined from the transistor characteristics

$$I_b = \frac{I_L}{\beta} \quad (9)$$

N_2 should supply at least one volt to the base of the transistor under all conditions.

$$N_2 = N_1 \frac{1}{E_1} \quad (10)$$

The number of turns for N_3 is determined by the desired output voltage to the load thus:

$$N_3 = N_1 \frac{E_0}{E_1} \quad (11)$$

The values for R_1 and R_2 are derived in the following manner:

$$E_{(N_2)} = \frac{N_2}{N_1} E_1 \quad (12)$$

Then

$$R_1 = \frac{E_{n2}}{N_3 \frac{I_L}{N_1 \beta}} \quad (13)$$

and

$$R_2 = \frac{E^2 \beta}{W(\max)}$$

where β is transistor current gain W max. power rating of transistor operated class A. Further information on this subject can be obtained in Reference 3.

REFERENCES

1. DR. W. A. GEYGER, *Magnetic Amplifier Circuits*, McGraw-Hill 1957.
2. H. F. STORM, *Magnetic Amplifiers*, John Wiley & Son 1955.
3. R. L. VAN ALLEN, *Designing DC-DC Converters*, Tech Note, Magnetics Inc.
4. DR. W. A. GEYGER, *Miniaturized Current Transformers for Multirange Operation of Milliammeter*, *Communication & Electronics*, AIEE, Jan. 1961.
5. W. KRAMER, *Proceedings of the Intermag Conference*.
6. FROST-SMITH, *Magnetic Amplifiers*, John Wiley & Son 1958.

A HYBRID RANGING SYSTEM FOR SPACECRAFT*

R. T. FITZGERALD, P. ENGELS, W. SHAFFER, E. HABIB, AND M. MITCHKO

This report describes, in engineering terms, the principles and practice of satellite ranging by direct measurement of the elapsed time for a hybrid combination of sidetone-modulated and pseudo random noise-modulated, CW, RF energy to travel from the ground, to a spacecraft and back, after being transponded in a phase coherent manner at the spacecraft. A ranging subsystem utilizing sidetone-PN techniques was successfully built at the Goddard Space Flight Center and is now ready for final loop tests and integration into the GSFC Range and Range Rate System, to give the latter universal cisplanet ranging capabilities.

The Hybrid Ranging System functions as follows: a demand pulse synchronized to WWV stores a binary pattern composed of transmitted bits and starts a time interval counter which is readied for stoppage when the identical signal is received. Simultaneously, sidetones, harmonically related to one another, are transmitted in alternately reversed phases, received in reversed phases and are used for vernier stop-control of the time interval counter. Thus, the non-ambiguous advantages of PN encoding are exploited without the cost of the wide bandwidth required for precise PN-only range measurements; and the precision and narrow bandwidth advantages of sidetone ranging are exploited without the cost of the engineering problems associated with ambiguity resolution by fractional cycle per second sidetones.

1. INTRODUCTION

A spacecraft ranging system which utilizes both sidetone and encoded signals has been built in the Goddard Space Flight Center laboratories and has worked as a subsystem—it is now ready for final loop tests. This universal hybrid system capitalizes upon both the advantages of sidetone ranging which permits fine distance definition under restricted bandwidth conditions, and upon the advantages of pseudo random encoding which permits unambiguous ranging without the need for many low frequency sidetones and their associated problems. The new system eliminates the basic disadvantages of previous ranging systems: the unnecessarily wide bandwidths and complex coding techniques of the encoded-signal-only system; and the fractional cycle per second problems of the sidetone-only system at cisplanetary distances.

The hybrid system will be exploited to enhance the highly successful GSFC sidetone ranging system which has functioned with precision on

Syncom in a special tracking network and which has made accurate range measurements on IMP in the general tracking network. With this modification, the GSFC Range and Range Rate System will have a universal ranging capability applicable to cislunar tracking problems such as Apollo and cisplanet tracking problems such as non-ambiguous, apogee-ranging on projected IMP missions.

The purpose of this report is to present an overall, but brief description of the Hybrid ranging system. The Hybrid modification is compatible with the highly successful Goddard Space Flight Center's Range and Range Rate System, and augments its capabilities. In this modification of the sidetone system, known encoded ranging techniques are combined with direct RF transit time measurement and sidetone techniques to create a more optimum design than previously known designs.† To distinguish the new system

†A number of the coding techniques described in this report were created by members of the Jet Propulsion Research Laboratory and were derived here from information contained in JPL Research Summary Reports 36-1, 36-2, 36-3 . . . etc. issued quarterly starting in 1960. Additional coding concepts were derived from other reports listed in the Bibliography.

*Published as *Goddard Space Flight Center Document X-531-64-71*, April 1964.

it is called the Hybrid System; it utilizes a "TIMER" coding technique (from: *Time Interval Measurements with an Encoded Ranging System*); it relies primarily upon direct measurement of the time required for both sidetones and encoded, CW, RF signal to travel from the ground to the spacecraft and back, after being transponded in a phase coherent manner at the spacecraft. In this report the TIMER techniques are emphasized because they are new; but the significant advantages and the successful field performances of the precise sidetone ranging system are the base upon which the Hybrid ranging system is built.

This report is divided into five sections as follows:

Basic TIMER System, Some Properties of Maximal Linear Codes, Some Properties of Acquirable Codes, Overall Description, and Hybrid System. An Appendix describes the practical digital circuitry used to demonstrate the feasibility of the Hybrid version of the TIMER System. More expanded and detailed descriptions of individual circuits, codes, and, specifically, the Goddard Space Flight Center's Range and Range Rate System are contained in the reports listed in the Bibliography.

2. BASIC TIMER SYSTEM

Basically, the TIMER System determines the time delay between the transmission and reception of an unambiguous PN-encoded pattern. This technique which is closely related to simple pulse radar methods, requires no complex computer operations. The principles of range measurements made with this system are illustrated in the simplified block diagram shown in Figure 1.

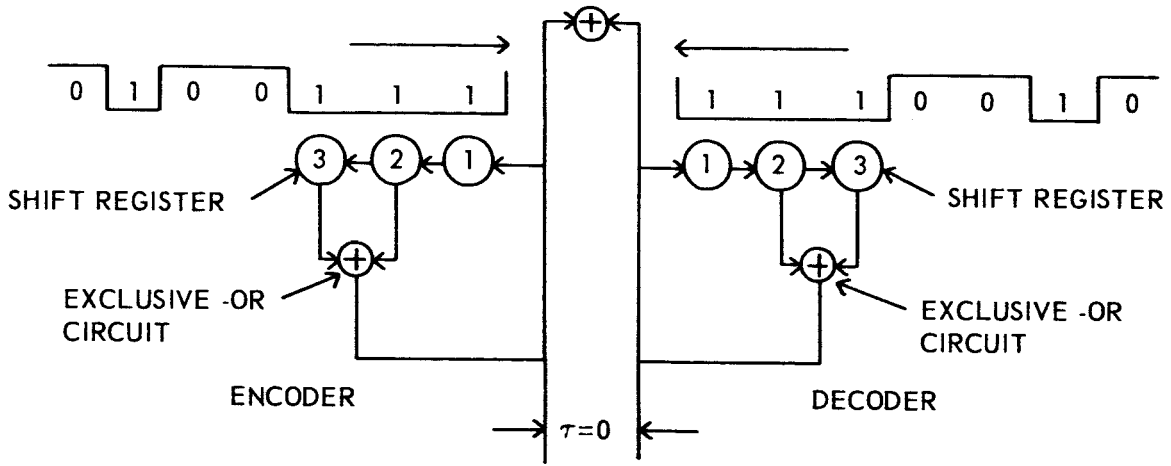
A simple maximal linear code, 1 1 1 0 0 1 0, is used as an example only. If this code is continuously transmitted as 1 1 1 0 0 1 0 1 1 1 0 0 1 0 1 1 . . . etc., it can be broken up into seven unambiguous, sequential three-bit groups of 1 1 1, 1 1 0, 1 0 0, 0 0 1, 0 1 0, 1 0 1, and 0 1 1. Each sequential group of three bits is unambiguous and definitive in the sense that within the total code each three-bit word or pattern occurs only once; each word is unique. Thus, if a transmitted pattern composed of three bits is stored and compared to the delayed, received pattern, received within the time interval of the whole code, only

one pattern out of those received will identically match the transmitted pattern.

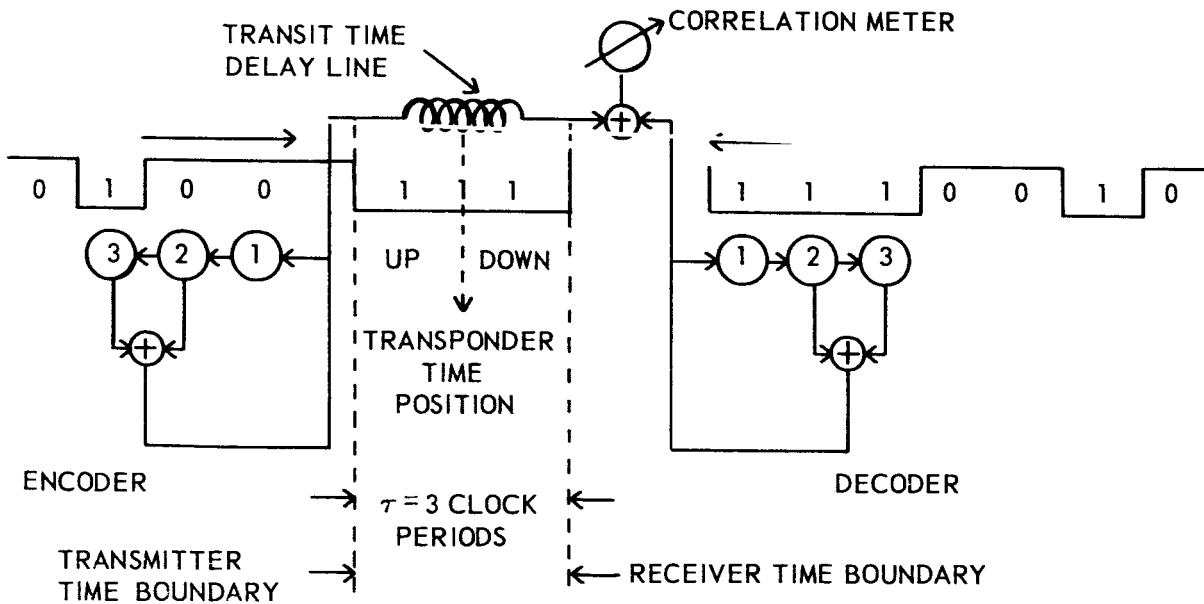
The encoder which modulates the transmitter, and the decoder which is "pattern-matched" to the received code, are each composed of a three-stage shift register with appropriate feed-back. The contents of the shift register flip-flops correspond during each bit-time interval to one of the three-bit patterns described above, and determine unambiguously the particular bit in the long code that is being transmitted or received as the case may be. Thus, if transit time, τ , is zero (the encode and decode generators are synchronized and their outputs are compared with no time delay) then the instantaneous patterns in both generators are identical (Figure 1A). As the two generators are separated by a distance corresponding to a transit time of $\tau = K$ (a constant), and the decode generator is kept locked to each received bit, the contents of the decode generator at time τ correspond to the contents of the encode generator at time zero (Figure 1B).

If, on the other hand, the receiver space boundary moves towards (or away from) the transmitter boundary (equivalent to a continuous change in transit time) to simulate transponder velocity, and at the same time the decode generator creates bits at a faster (slower) rate to keep up with the faster (slower) rate at which bits are received, the previous arguments still hold. Thus, when the leading edge of a particular transmitted bit reaches the receiver time boundary, it is met, so to speak, by the corresponding bit from the decode generator—and the contents of the decode generator at time τ will match the previous contents of the encode generator at time zero. In brief, the measurement of τ by determining the elapsed time between the formation of an encode pattern and the formation of an equivalent decode pattern, is closely related to simple pulse radar techniques, and is in essence independent of Doppler effects. In addition, the present techniques are compatible with long time constant, narrow bandwidth, phase-locked loop acquisition techniques with their improved signal-to-noise properties.

In the following sections, the simplified example of the encoder given above is expanded into a more practical encoder suitable for lunar and cisplanet range measurements.



A. ZERO DISTANCE CASE



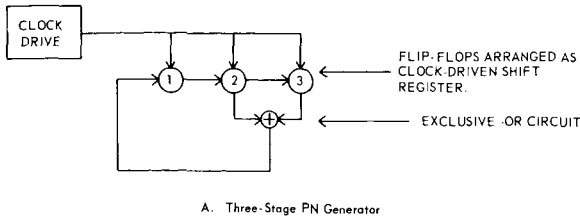
B. TRANSIT TIME DELAYED CASE

FIGURE 1.

3. SOME PROPERTIES OF MAXIMAL LINEAR CODES

Some of the pseudo random noise (PN) properties that are exploited in the TIMER System

are illustrated below. As shown in Figure 2A, if three flip-flops are connected as a shift register, and the outputs of the last two stages are combined in a circuit, designated \oplus or exclusive-or



A. Three-Stage PN Generator

Time	State of Flip-Flops			Input to First Stage on Next Shift
	(1)	(2)	(3)	
t_0	1	1	1	0
t_1	0	1	1	0
t_2	0	0	1	1
t_3	1	0	0	0
t_4	0	1	0	1
t_5	1	0	1	1
t_6	1	1	0	1
t_7	1	1	1	0

B. Tabulation of Sequential States of Three Stage PN Generator

FIGURE 2

circuit,* which feeds a one back to the input whenever the last two stages are in different states, a code is generated which has almost ideal correlation properties. The states of the several flip-flops are tabulated in the table given in Figure 2B, and the generated code corresponds to the contents of the third flip-flop as indicated in the table. This code is called a maximal linear binary code or m-sequence (see Bibliography).

The output code is 1110010 and is repeated indefinitely as 111001011100 The length of the code, L , is $2^n - 1$ where n is the number of flip-flops (i.e. the number of stages in the shift register). If this code is shifted and added to the unshifted code, according to the rule that a one is generated wherever an element of the original and an element of the shifted identical code are different (exclusive-or or mod-2 addition), the result will *always* be the *same code shifted* again as shown below (except when the code is shifted zero, or a multiple of L bits):

original code 11100101110010

*The exclusive-or function, $F = A \cdot \bar{B} + B \cdot \bar{A}$, is variously indicated in the technical literature as, exclusive-or, \oplus , mod-2 addition, modulo-2 addition, addition without carry and as a function which is true when two binary outputs disagree, and false when two binary outputs agree.

code shifted
by one bit \oplus 1110010111001

resultant code 0010111001011

original code 11100101110010

code shifted
by two bits \oplus 111001011100

resultant code 011100101110

“ “ “

“ “ “

“ “ “

“ “ “

original code 111001011100101

code shifted
by five bits \oplus 1110010111

resultant code 0101110010

Thus, these PN codes, have the highly important property that when shifted and compared, mod-2 with the original code, the same code is generated. And the resultant code has the same distribution of 1's and 0's as the original code which *always* has one more 1 than 0. It is noted here that the resultant code has four 1's and three 0's, and that a 1 indicates a disagreement, and a 0 indicates an agreement between the original and the shifted codes.

When the above-noted shift and add properties of the PN code are related to the simple binary correlation function described below, the nearly perfect correlation properties of PN codes become apparent. Binary correlation can be defined as follows:

$$\text{correlation} = \frac{\text{no. of agreements} - \text{no. of disagreements}}{\text{no. of agreements} + \text{no. of disagreements}}$$

or

$$C = \frac{A - D}{A + D}$$

Where C is the correlation factor, A the number of agreements, and D the number of disagreements. Inserting the numbers illustrated:

$$C = \frac{3 - 4}{3 + 4} = -\frac{1}{7} = -\frac{1}{L}$$

The correlation factor for 1 to $L-1$ shifts is always:

$$C = -\frac{1}{L} = \frac{-1}{2^n - 1}$$

because, for these codes, in general,

$$C = \frac{\frac{L-1}{2} - \frac{L+1}{2}}{L} = -\frac{1}{L}$$

In words, this is true because the exclusive-or or mod-2 acquired code *always* contains one more disagreement than agreement; the acquired code is the original, but shifted, maximal linear pseudo random code.

Whenever the code is unshifted (or shifted a multiple of L bits) and added mod-2 to the original code (disagreements are 1's and agreements are 0's), all corresponding elements of the two codes agree, $A=L$, $D=0$, and

$$C = \frac{A-D}{A+D} = \frac{L-0}{L+0} = 1$$

In short mathematical form, the correlation factor derived by the processes illustrated above and the conclusions reached, can be described, for PN codes, as:

$$\begin{aligned} C(k) &= \frac{1}{L} \sum_{n=0}^{n=L-1-k} a_n \oplus a_{n-k} \\ &= -\frac{1}{L} \text{ for } k=1 \text{ to } L-1 \\ &= 1 \text{ for } k=0 \text{ or } mL \end{aligned}$$

Where a_n is an element of the original code

a_{n-k} is an element of the shifted code

m is an integer

k is an integer corresponding to the number of shifts of the shifted code.

n is an integer.

To summarize, the correlation property of the PN code is essentially (for longer codes) zero when not matched exactly and unity when matched exactly. The PN codes possess nearly perfect correlation properties.

In the TIMER System, the local decoder code is "locked on" to the received code by automatic

circuitry which measures the correlation factor of the decoder code with the received code; the decoder code is driven by the received clock plus Doppler and shifted periodically by extra shifts until correlation is obtained. Long time constant integration and corresponding narrow bandwidth circuits enhance the signal to noise properties of this acquisition process.

Although a three flip-flop generator was selected as an example, codes of this type can be simply made for lengths up to $2^{34}-1$ (about 10^{10}) and longer. It is not always true that the mod-2 or exclusive-or circuit is connected to the last two stages; other connections, as described below and in the reports listed in the Bibliography, are necessary for certain lengths.

In Table 1, some of the simpler connections for m -sequences are listed.

TABLE 1.—Linear Maximal Generators with Simple Connections*

n	Stages Tapped	n	Stages Tapped
2	2, 1	18	18, 11
3	3, 2	20	20, 17
4	4, 3	21	21, 19
5	5, 3	22	22, 21
6	6, 5	23	23, 18
7	7, 5	25	25, 22
9	9, 5	28	28, 25
10	10, 7	29	29, 27
11	11, 9	31	31, 28
15	15, 14	33	33, 20
17	17, 14		

*See Bibliography, Peterson, Appendix C.

4. SOME PROPERTIES OF ACQUIRABLE CODES

To handle longer codes practically, they are broken up into shorter "acquirable" codes. These longer codes, composed of the shorter codes, can be relatively easily generated as illustrated in Figure 3. The generated code, W_n , is equal to $x \oplus yz$ —it is the mod-2 sum of the x code with the y "logical and" z codes.

For the example illustrated in Figure 3, the overall code length L , is the product of the shorter codes $L_x L_y L_z$, or $(2^3-1)(2^4-1)(2^5-1) = 3255$ bits. To acquire such a code, the shorter codes are "locked on" to the total received code, W_R ,

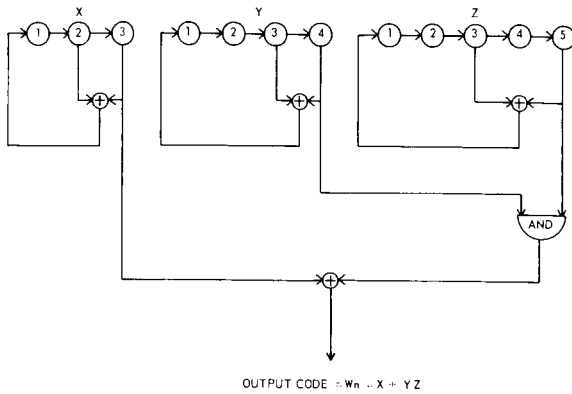


FIGURE 3.—Acquirable code generator.

usually sequentially, and the maximum number of shifts necessary to acquire the long code is $Lx + Ly + L_z$ or $7 + 15 + 31 = 53$ shifts—the x_m code is correlated with W_R first, then the $x_m \oplus y_m$ combined codes, and finally the $x_m \oplus y_m \oplus z_m$ codes; after this process W_m the complete decoder code, $(x_m \oplus y_m \oplus z_m)$ is completely “matched” or correlated with the received code, W_R .

The almost ideal correlation levels of $-\frac{1}{L}$ and 1 for single-code to single-code correlation, are reduced to approximately $-\frac{1}{L}$ and $\frac{1}{2}$ for the “acquirable-code” correlation processes described above. As shown below, the “acquirable-code” acquisition process results in a change from a sequence of approximately random 1’s and 0’s (which exhibit nearly zero correlation) to a sequence of 1’s and 0’s of approximately $\frac{1}{4}$ ones and $\frac{3}{4}$ zeros (which exhibit nearly $\frac{1}{2}$ correlation), when a given short code is correlated with the longer code. For example,

$$W \oplus x = x \oplus yz \oplus x = yz$$

and the probability code y and code z have the value of 1 at any time is approximately $\frac{1}{2}$ times $\frac{1}{2}$ or $\frac{1}{4}$. Thus

$$C = \frac{A - D}{A + D} = \frac{\frac{3}{4} - \frac{1}{4}}{1} = \frac{1}{2}$$

Also, if the x and y codes are correlated,

$$\begin{aligned} W \oplus x \oplus y &= x \oplus yz \oplus x \oplus y \\ &= y \oplus yz \\ &= y \cdot \bar{y}z + \bar{y} \cdot yz \text{ (by expansion)} \\ &= y \cdot (\bar{y} + z) = y \cdot \bar{z} \text{ (by DeMorgan's theorem)} \end{aligned}$$

and again the probability that y “logical and” z have the value of 1 at any time is approximately $\frac{1}{4}$ as above.

In the TIMER System, negative logic and -6 volts levels for true statements are used. As shown later, -3 volts indicate zero correlation, and -1.5 volts indicates 50 percent correlation. Automatic shifting and recognition circuitry acquire and lock the decoder codes with the received code, by measuring the change in voltage level from $-3v$ to -1.5 volts. When all codes are acquired, the correlation indicator goes to 0 volts, i.e. 100 percent correlation.

5. OVERALL DESCRIPTION

By a simple extension of the principles described in the section called the Basic TIMER System, the advantages of *acquirable* codes can be exploited in the TIMER System as illustrated in Figure 4. Just as each one of the seven possible three bit words in the three flip-flop PN generator determines the bit being transmitted, so each of the four bit words and five bit words respectively for the four and five bit flip-flop PN generators are definitive. And the summation 12 bit word, composed of the 3, 4, and 5 bit words, defines exactly the particular bit of the 3255 bit code that is being transmitted. A timer can be initiated at the instant in time that a particular 12 bit pattern in the encoder generators is formed, and this pattern can be stored in a register; the timer can be stopped at the instant in time that the decoder generators, previously locked to the received code, match the stored encoder pattern. The timer, then, will indicate the transit time.

The timing actions that take place are the following: a demand pulse synchronized for timing

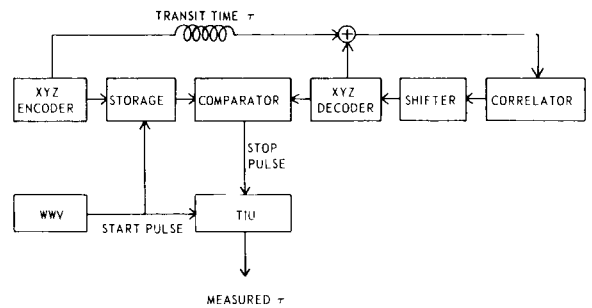


FIGURE 4

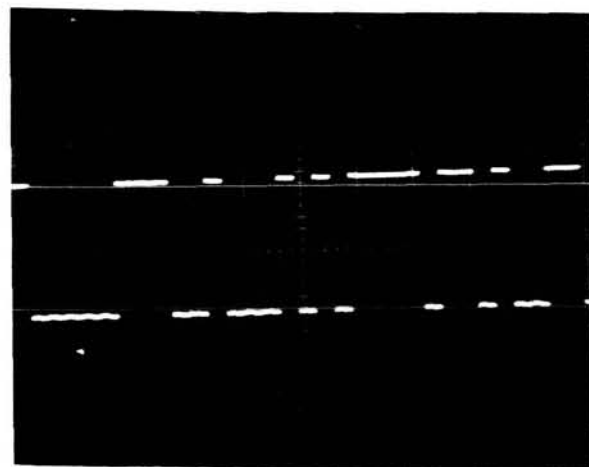
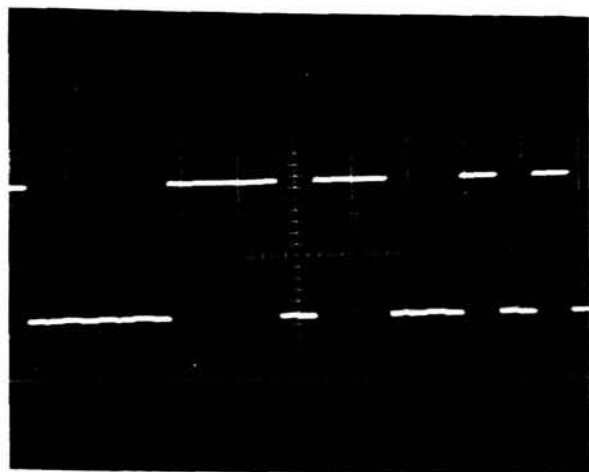
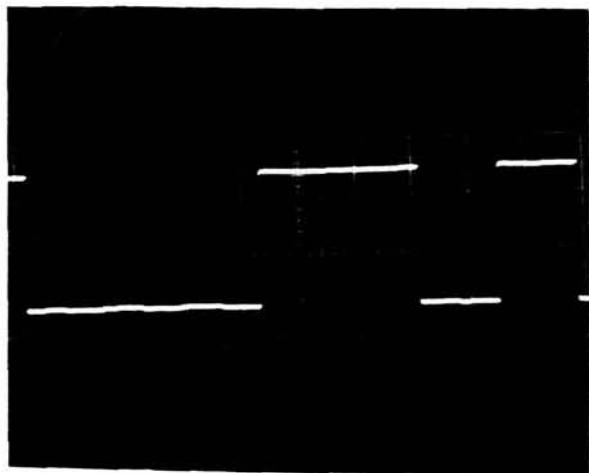


FIGURE 5a

purposes to WWV, stores the existing 3 bits of the X_n code, 1 1 1 0 0 1 0, in a storage register; simultaneously the existing 4 bits of the y_n code 1 1 1 1 0 0 0 1 0 0 1 1 0 1 0, and the existing 5 bits of the Z_n code 1 1 1 1 1 0 0 0 1 1 0 1 1 1 0 1 0 1 0 0 0 0 1 0 0 1 0 1 1 0 0, are also stored in the storage register. After the elapse of transit time τ , the 12-bit contents of the X_m , Y_m , and Z_m decode generators, will match the stored patterns from the encode generators. Figure 5 shows the form factors of these codes in various combinations, as displayed on an oscilloscope. Since the



FIGURE 5b



FIGURE 5c

WWV demand pulse started at time interval unit (TIU) at the instant the encoder patterns were stored, and since the comparator unit stopped the TIU at the instant the decoder patterns matched the stored patterns, the TIU, a simple fixed frequency counter, determines the transit time in terms of the invariant speed-of-light constant. In the Hybrid application of the TIMER System, the fine stop control of the TIU is controlled by the zero crossing of a relatively high frequency sidetone; the TIMER is used to unambiguously "ready" the TIU for the stopping action of the sidetone.

5.1 Correlation Circuit

An important unit in the TIMER System is the correlator which integrates the mod-2 sum of the received code and the decoder code. When the mod-2 sum is a series of approximately equal numbers of 1's and 0's, the codes are uncorrelated; a selectable and relatively long time constant RC circuit averages the $-6v$ and 0 volt levels corresponding respectively to 1's and

0's and indicates the average value of -3 volts. As noted before, the same RC circuit indicates a level of approximately -1.5 volts when one of the short codes is correlated. As each short code is correlated, an automatic shifting process is terminated automatically. After each short code is acquired by this process, a manual switch (which will be automated for the systems used in the tracking network) is thrown to start the automatic acquisition of another short code. Several selectable time constants for integration are available depending upon signal energy levels. In the actual Hybrid System, the lengths of the x , y , and z codes are selectable according to mission requirements for unambiguous ranging; the clock rates, 8, 32, 160, 800, or 4000 cps are also selectable according to range requirements.

The correlation process is, of course, completed before any transit time measurements are made. An automatic lock prevents range data recording unless the correlator unit indicates 100 percent lock-on.

6. HYBRID SYSTEM

The Hybrid System is essentially an extension of existing range and range rate units. Two of the major spacecraft ranging systems are the Sidetone type designed and successfully used by the Goddard Space Flight Center and the PN type under development by the Jet Propulsion Laboratories. Since both of these types are described in much detail in a number of the reports given in the Bibliography, the discussions in this report are directed primarily towards describing the technical units employed in the TIMER System. Nevertheless, a few statements about current ranging systems will help to define the Hybrid System in a comparative sense.

The Sidetone ranging system as designed by GSFC employs a number of sidetones whose frequencies are integrally related by a factor of 4 or 5. These sidetones, suitably modulated on a carrier, are transmitted, coherently transponded, and received. After phase-lock loop acquisition of each sidetone in the ground receiver, range measurements are made in essence by measuring the elapsed time for all sidetones transmitted in alternately reversed phases to be received in alternately reversed phases after being transponded in a phase coherent manner in the space-

craft. The lowest sidetone employed is selected to be unambiguous for the distances to be measured (transit time, τ , is less than the period, T , of the lowest frequency sidetone).

The PN ranging system under development by JPL utilizes a PN code imposed upon a clock frequency, equivalent to a single sidetone, and, after phase locked loop acquisition of the clock frequency and correlation of the PN code, range measurements are made, in essence, by determining the phase difference between the transmitted clock and received clock, and by computing the equivalent transit time values of each shift of the short codes which compose the longer acquirable code; after this process, range information is updated by adding in distance changes determined by counting Doppler cycles.

The Hybrid System, in its optimum form, utilizes several sidetones in the manner of the GSFC Range and Range Rate System, and, in addition, imposes a PN code on the lowest sidetone employed to increase non-ambiguous ranging distance. Range measurements are made by direct measurements of the elapsed transit time.

With this orientation, it is noted that the stop pulse action illustrated in the block diagram of Figure 4 is improved by the phase definition of the 20 and 100 Kc sidetones (see Appendix). The Hybrid System thus utilizes sidetone technology and PN technology to determine range by direct transit time measurements. The Hybrid System capitalizes upon both the advantages of sidetone ranging which permits fine distance definition under restricted bandwidth conditions, and upon the advantages of pseudo random encoding which permits unambiguous ranging at long distances without the need for extremely low frequency sidetones and their associated phase-locked-loop acquisition problems.

The Hybrid System, optimally, utilizes pseudo random encoding of a sidetone, for example 4 Kc, and also utilizes higher sidetones of 20 Kc and 100 Kc for fine distance measurements. It could be used equally well with a single clock frequency, (PN only ranging), but the higher clock frequency necessary for fine ranging, and the higher frequency code would occupy more spectral bandwidth than necessary; longer codes and longer or more complex acquisition processes would also be necessary.

As presently designed, the GSFC Range and Range Rate System solves the ambiguity problem by use of a novel technique of alternately phase reversing each higher frequency transmitted sidetone; this technique permits each half period of a lower frequency sidetone to gate "on" the next higher frequency half period of sidetone at the receiver. By properly arranging a series of flip-flops, the fine determination of range measurement is held to the phase definition of the 100 Kc sidetone which is $\pm 0.36^\circ$ in the GSFC Range and Range Rate System; this definition is equivalent to a time error of 10 nanoseconds or about five feet of distance. This ultimate distance definition was practically achieved in tracking the Syncom at a range of about 20,000 miles. The Hybrid System retains this capability but utilizes the PN code to fix the time position of the first 4 Kc gate—the 20 Kc and 100 Kc gates then function as in the present system. In Appendix A, the fundamental units for ranging with the TIMER System are illustrated with their logic equations.

6.1 Cisplanet Ranging

A question arises as to how far the long code is shifted when one of the shorter codes is shifted a given amount. In the TIMER System no computation based on this analysis is employed at cislunar distances. At cisplanet ranging distances some of the TIMER System technology can be profitably employed by utilizing this computation. In the latter case, it is desirable to make range measurements as soon as the ground-satellite-ground loop is filled with RF energy and the decode generator is matched to the received encoded signal (i.e. the received signal is acquired), without waiting the additional time required for a given generator pattern to travel the loop distance. Very briefly, this result can be accomplished by employing the previously described encoder register plus an additional decoder storage register. Range measurements can then be made by measuring the phase difference between encode-decode clocks, and by determining the phase shift distance of each encode-decode short code one to another. The advantages of this method at cisplanet distances are: one, no additional time beyond acquisition time is required for ranging; two, all ranging measurements are independent; three,

code reacquisition is not necessary to obtain independent measurements; and four, sampling rates can be increased as far as practically required.

In order to utilize the cisplanetary techniques outlined above, the contents of the decode generator, as well as the contents of the encode generator, must be stored in a storage register at demand time. Each decoder short code is shifted until the encode-decode short codes match; the number of shifts of each code plus the encode-decode clock phase difference at demand time determine spacecraft range at demand time minus D/C where C is the velocity of light and D is the one-way distance to the spacecraft. Either of the following two equations, derived by means of the Chinese Remainder Theorem of Number Theory, determines the equivalent shift of the long code in terms of the individual shifts of the short codes.

$$\begin{aligned} & -465 \times 2 \text{ (} x \text{ shift)} - 217 \times 2 \text{ (} y \text{ shift)} \\ & + 105 \times 13 \text{ (} z \text{ shift)} = \text{(total code shift)} \\ & \text{mod } 3255 \end{aligned}$$

$$\text{or: } 465 \times 5 \text{ (} x \text{ shift)} + 217 \times 13 \text{ (} y \text{ shift)} + 105 \times 13 \text{ (} z \text{ shift)} = \text{(total code shift)} \text{ mod } 3255$$

It is noted that, the actual computation of distance can be performed at a central computer, instead of at each network station, with little loss in communications efficiency because the number of bits required to transmit the distance information is practically equal to the number of bits required to transmit the code shift and phase information.

For completeness the following fact is noted: if it is desired to shift the long code a given amount, say N shifts, each short code is shifted $N_{(\text{mod short code})}$. For example, if the long code is to be shifted an amount 33, the x code is shifted $33_{\text{mod } 7}$ or 5, the y code $33_{\text{mod } 15}$ or 3, and the z code $33_{\text{mod } 31}$ or 2.

6.2 Conclusion

In conclusion, this report has attempted to describe the principles and practical design features of an improved system for spacecraft ranging with a hybrid combination of sidetones and encoded signals. Only by integrating the

information presented here, with other information available from the reports listed in the Bibliography, can a complete understanding of all the units necessary for a practical ranging system be obtained. In this brief presentation of principles and practice no deliberate attempt has been made to minimize the essential complexity

of spacecraft ranging. By its very nature, spacecraft ranging draws on the ultimate of current technology. This report is intended to enhance this technology by one more contribution, and will have succeeded only when the principles presented here are reduced to practice and are utilized in the range and range rate network.

APPENDIX A

This Appendix describes the basic, practical, logic circuits for a Hybrid type ranging system employing PN TIMER and Sidetone techniques. Design criteria and operational feasibility are demonstrated.

Figure A-1 is a simple block diagram of the TIMER System. Sidetones of 100, 20, and 4 Kc are produced in the Sidetone Generator. The 4 Kc signal is used to drive the Encoder which produces the PN 3255-bit code and is also used as the phase modulated subcarrier for the PN code. The 100 Kc, 20 Kc and PN Code plus 4 Kc are then phase modulated on a carrier and transmitted. The signal is received and coherently transmitted back to the ground by the Transponder. This signal is received by the ground

equipment and from it is derived the Decoder Clock frequency (Encoder Clock frequency plus Doppler) and the 100 Kc and 20 Kc Sidetones plus Doppler which are fed to the appropriate phase-locked loops. The Decoder produces a shifted Code identical to that of the Encoder but produces code bits at the clock-plus-Doppler frequency. The Correlator locks the decoder generator to the received code and then by means of the Transit Time circuit, determines the range. Note that the Sidetone Generator, Transmitter, and Receiver equipment shown (but not the PN coders and code modulators and demodulators) already exist as part of the Goddard Range and Range Rate System.

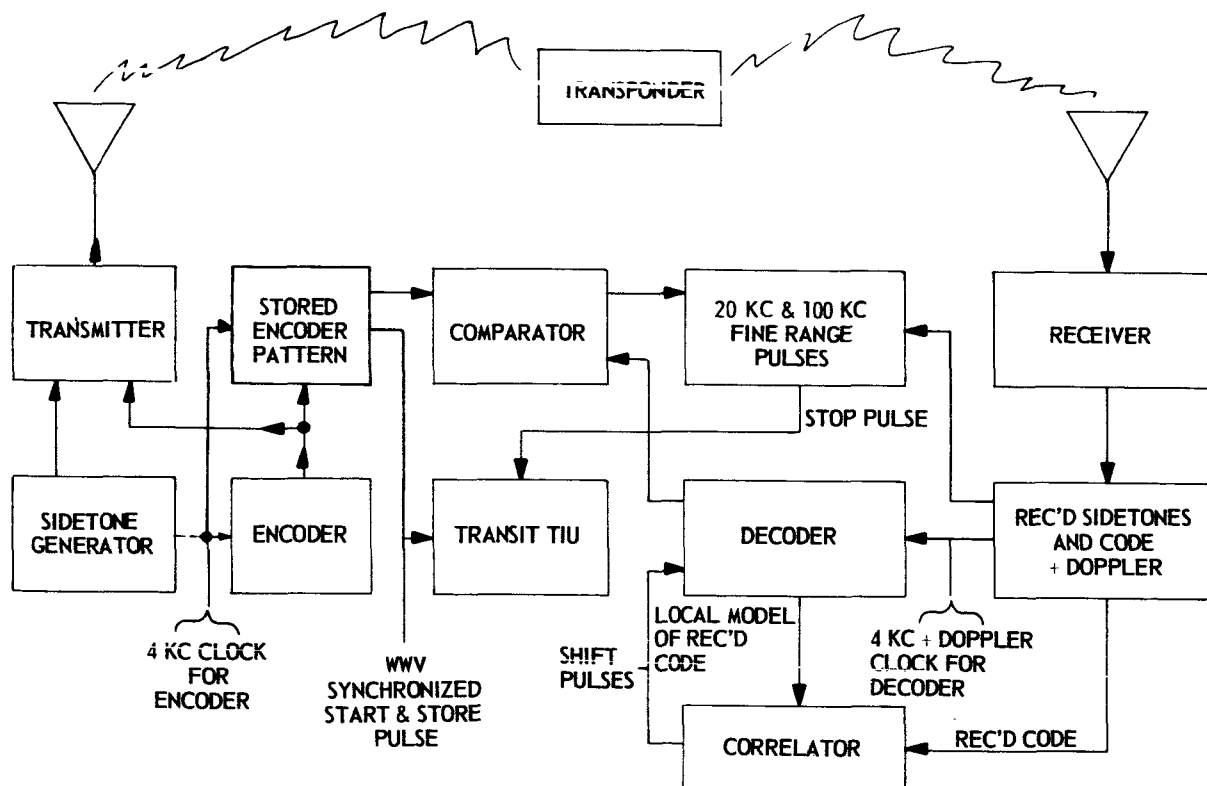


FIGURE A1.—Simplified system block diagram.

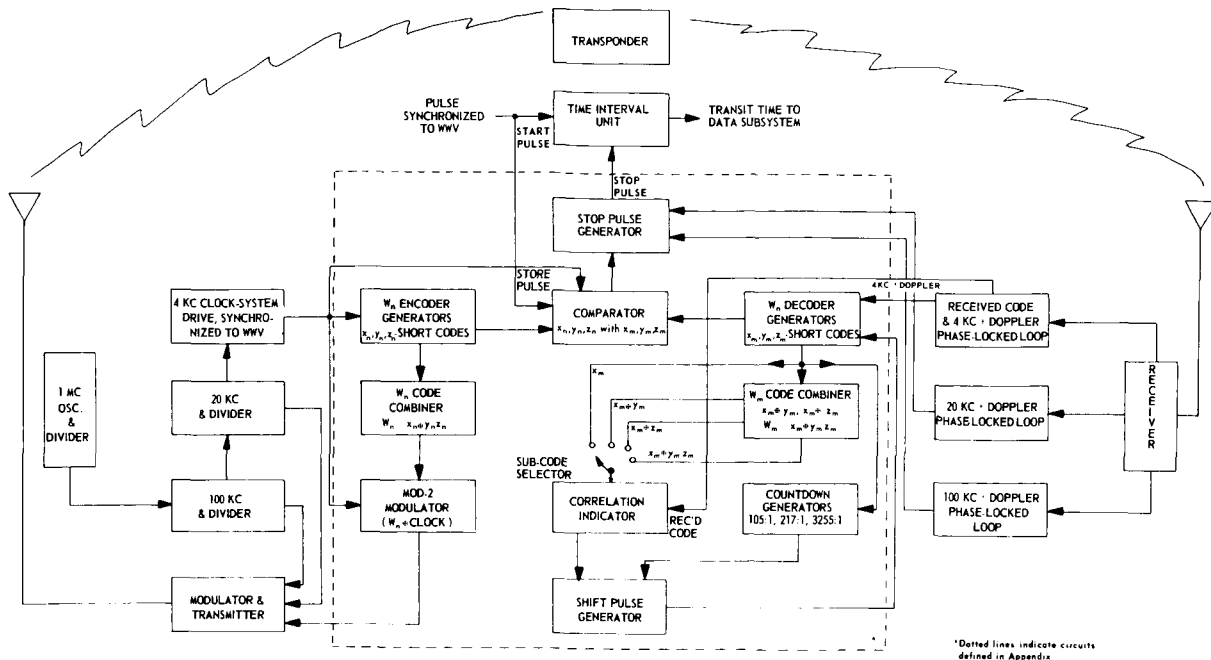


FIGURE A2.—System block diagram.

Figure A-2 shows the system components in more detail. The TIMER System can be broken down further into the following subsystem designations: the transmitter subsystem which includes the transmitter, the W_n^* Encoder Generators, the W_n Code Combiner, and the Mod-2 Modulator; the receiver subsystem which includes the receiver, the W_m Decoder Generators, the W_m Code Combiners, and the Countdown Generators; the correlator which includes the Correlation Indicator and Shift Pulse Generator; and the ranging subsystem which includes the Comparator, Stop Pulse Generator, and the Time Interval Unit.

Some of the functional operations of the system components are the following: A local 1 MC oscillator and frequency synthesizer are used to provide frequencies of 100, 20, and 4 Kc. The 4 Kc square wave which is synchronized with the 20 Kc, 100 Kc and WWV frequencies is used to drive the W_n Encoder Generator, the Mod-2 Modulator, and the Comparator.

W_n Encoder Short-Code Generators function to produce the X_n , Y_n , and Z_n short codes and provide outputs for operating the storage registers in

*The subscripts n and m are used throughout the Appendix to denote encoder and decoder related items, respectively.

the Comparator, and for combining codes in the W_n Code Combiner. The Code Combiner provides the output $X_n \oplus Y_n Z_n$ for operating the Mod-2 Modulator. Additional clock energy, not illustrated, is added to the 4 Kc subcarrier to permit 4 Kc clock phase-lock-loop acquisition in the receiver, independent of code lock-on.

The W_m Decoder Short-Code Generators function to produce the X_m , Y_m , and Z_m short codes and outputs for operating the Comparator gates, the W_m Code Combiners, and the Countdown Generators. The W_m Code Combiner outputs, $X_m \oplus Y_m Z_m$, $X_m \oplus Y_m$, and $X_m \oplus Z_m$ are used for correlation indication. The Countdown Generators produce countdown pulses of 3255:1, 217:1, and 105:1 to provide different shift rates for the Shift Pulse Generator.

The Correlation Indicator serves to detect and display correlation of the received code with W_m .

Correlation is obtained by shifting W_m with respect to the received code by means of extra shift pulses (added to the clock plus Doppler shift pulses) from the Shift Pulse Generator.

The Comparator is used to compare the transmitted code with the received code thereby providing the time basis for range determination; the Comparator also serves to activate the Stop

Pulse Generator. Actual range is determined by the Time Interval Unit which is activated by a WWV demand pulse and deactivated by the Stop Pulse Generator.

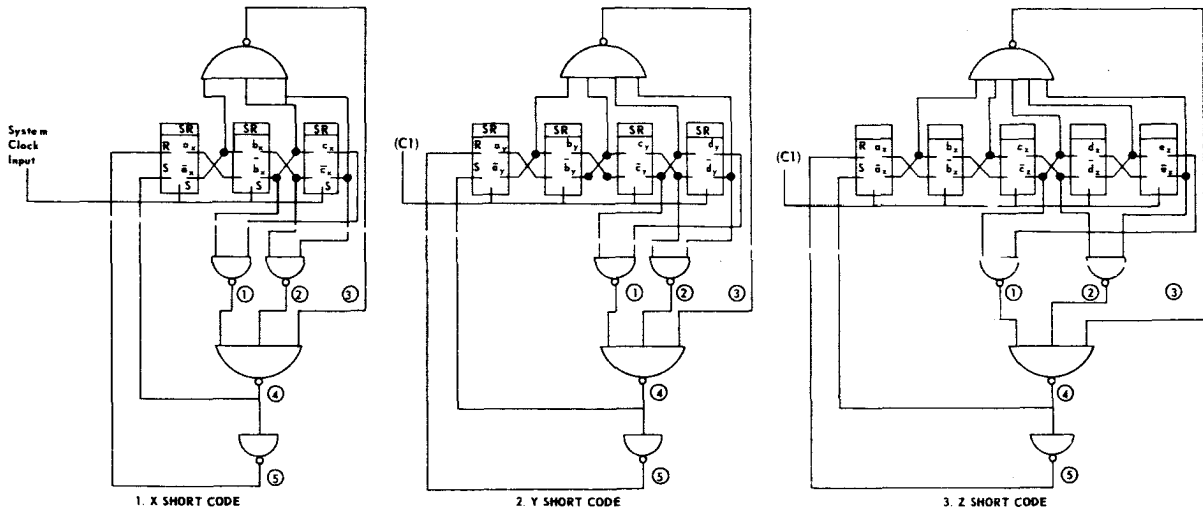
The sequence of operations is as follows: The output of the Mod-2 Modulator, $W_n \oplus C1$, where $C1$ is the 4 Kc System Clock, together with the 100 and 20 Kc frequencies, is phase-modulated on a carrier and transmitted.

The complex signal is coherently transponded and received by existing equipment in the Goddard RARR system and, after proper conditioning, is detected by the 100, 20, and 4 Kc phase locked loops. Each loop contains the particular sidetone plus its Doppler frequency. The received 4 Kc Clock plus Doppler constitutes the Decoder clock; this signal also contains the received PN code information. Initially, the

received code and the local code generated by the Decoder generators, W_m , are uncorrelated, i.e., are out of phase by an undetermined number of bits.

The correlation condition is detected and displayed by the correlation indicator, which determines sequentially the correlation levels of the decoder X_m , Y_m , and Z_m short codes, with respect to the total received combined code. The short codes are actually correlated as X_m , $X_m \oplus Y_m$, $X_m \oplus Z_m$, and $X_m \oplus Y_m \oplus Z_m$, in sequence.

The shift pulse generator shown is used to shift, separately, the X_m , Y_m , and Z_m short codes. When one of these short codes is correlated with its component in the received code, the Shift Pulse Generator is automatically turned off by a pulse generated by the Correlation Indicator. This allows the shift pulse to be directed by



1. X-SHORT CODE

1.1. $F_{(1)} = \overline{b_x \cdot c_x}$
 1.2. $F_{(2)} = \overline{b_x \cdot c_x}$
 1.3. $F_{(3)} = \overline{a_x \cdot b_x \cdot c_x}$
 1.4. $F_{(4)} = F_{(4)}$
 $= \overline{(b_x \cdot c_x) \cdot (b_x \cdot c_x) \cdot a_x \cdot b_x \cdot c_x}$
 $= \overline{b_x \cdot c_x + a_x \cdot b_x \cdot c_x}$
 1.5. $F_{(5)} = \overline{b_x \cdot c_x + a_x \cdot b_x \cdot c_x}$

2. Y-SHORT CODE

2.1. $F_{(1)} = \overline{c_y \cdot d_y}$
 2.2. $F_{(2)} = \overline{c_y \cdot d_y}$
 2.3. $F_{(3)} = \overline{a_y \cdot b_y \cdot c_y \cdot d_y}$
 2.4. $F_{(4)} = F_{(4)}$
 $= \overline{(c_y \cdot d_y) \cdot (c_y \cdot d_y) \cdot a_y \cdot b_y \cdot c_y \cdot d_y}$
 $= \overline{c_y \cdot d_y + a_y \cdot b_y \cdot c_y \cdot d_y}$
 2.5. $F_{(5)} = \overline{c_y \cdot d_y + a_y \cdot b_y \cdot c_y \cdot d_y}$

3. Z-SHORT CODE

3.1. $F_{(1)} = \overline{c_z \cdot e_z}$
 3.2. $F_{(2)} = \overline{c_z \cdot e_z}$
 3.3. $F_{(3)} = \overline{a_z \cdot b_z \cdot c_z \cdot d_z \cdot e_z}$
 3.4. $F_{(4)} = F_{(4)}$
 $= \overline{(c_z \cdot e_z) \cdot (c_z \cdot e_z) \cdot a_z \cdot b_z \cdot c_z \cdot d_z \cdot e_z}$
 $= \overline{c_z \cdot e_z + a_z \cdot b_z \cdot c_z \cdot d_z \cdot e_z}$
 3.5. $F_{(5)} = \overline{c_z \cdot e_z + a_z \cdot b_z \cdot c_z \cdot d_z \cdot e_z}$

Legend: $F_{(1)}$, $F_{(2)}$, etc., refer to the existing functions at the related points designated by a circled number on the above logic diagram.
 $F_{(x)}$, $F_{(y)}$, etc., refer to the functions which generate the encoding-decoding PN codes x, y, etc., respectively.
 Letters a, b, etc., refer to particular shift register stages.
 The subscripts in terms c_x , c_y , etc., refer to shift registers which are associated with codes x, y, etc., respectively.

FIGURE A3.—Encoder-decoder generators.

manual switching to the other W_m short code generators. Because of this self-turn-off feature, no shift pulses are generated after complete correlation is obtained.

When complete correlation exists and on demand from WWV, the Time Interval Unit is activated; at the same time the 12-bit Encoder word is stored in the registers of the Comparator. When the outputs of the registers of the Decoder generators (running synchronous with the received code) match the stored code, a stop pulse is generated which readies the Time Interval Unit for the stop command from the highest sidetone used. The Time interval reading obtained is the range in transit time units.

Figure A-3 shows the Encoder-Decoder short code generators. The design of each is such that a maximal linear short code is generated, i.e., a code of $2^n - 1$ bits is obtained, where n is the number of stages in the shift register. Thus the X , Y , and Z short codes contain, respectively, 7, 15, and 31 bits. These generators are implemented with modulo-2 (exclusive "or") logic by means of the gated feedback from appropriate stages of the register. At (4), a logical "1" will be generated when the outputs of each stage differ and a logical "0" when the outputs are alike (either two "1's" or two "0's"). Without input (3) (a self-starting mechanism) the outputs at (4) are represented by the functions:

$F(x) = b_x \oplus c_x$, $F(y) = c_y \oplus d_y$, $F(z) = c_z \oplus e_z$, where $F(x)$, $F(y)$, $F(z)$ represent the logic functions, respectively of the x , y , and z short code generators, and b_x , c_x , c_y , d_y , c_z , and e_z represent outputs of particular stages of those generators. Note that for the maximal linear Z-code, outputs c and e (stages 3 and 5) are tapped rather than d and e , (the last two stages).

For system reliability, a self-starter is incorporated into the design of each Short Code Generator to assure code generation each time the system is turned on. A logical "1" will be generated by the self-starter at (3) for all conditions except when the reset outputs (\bar{a} , \bar{b} , \bar{c}) of the generators are all in the "1" state (or the set outputs are all "0's"). Then a logical "0" is generated at (3) which provides proper gating at gate (4). Without this feature the codes would not be generated if all stages were in the zero state at turn-on. With the self-starter, the

functions of X , Y , and Z at (4) are:

$$F(x) = b_x \oplus c_x + \bar{a}_x \cdot \bar{b}_x \cdot \bar{c}_x,$$

$$F(y) = c_y \oplus d_y + \bar{a}_y \cdot \bar{b}_y \cdot \bar{c}_y \cdot \bar{d}_y,$$

$$F(z) = c_z \oplus e_z + \bar{a}_z \cdot \bar{b}_z \cdot \bar{c}_z \cdot \bar{d}_z \cdot \bar{e}_z.$$

Figure A-4.1 and A-4.2 show the W_m (Decoder) Sub-Code Combiners. Each Sub-Code combiner adds, modulo-2, the X_m short code with one of the

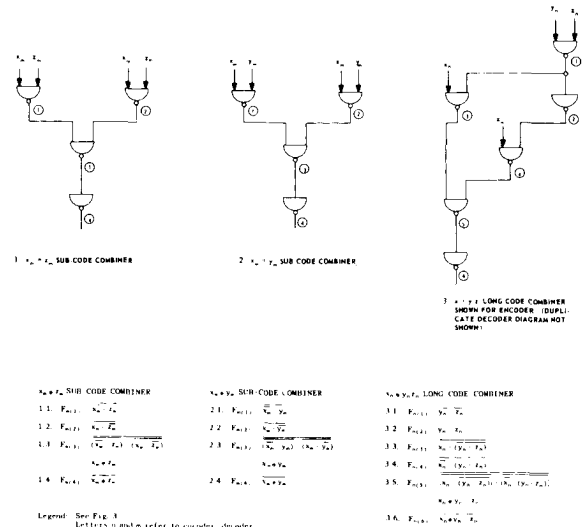


FIGURE A4.— W_m sub-code combiners and $W_n - W_m$ code combiner.

other short codes. The output at 1.3 is the $X_m \oplus Z_m$ sub-code containing 217 bits, i.e., the X_m short code length (7 bits) times the Z_m code length (31 bits). The output at 2.3 is the $X_m \oplus Y_m$ sub-code containing 105 bits, i.e., the X_m short code length times the Y_m short code length (7×15).

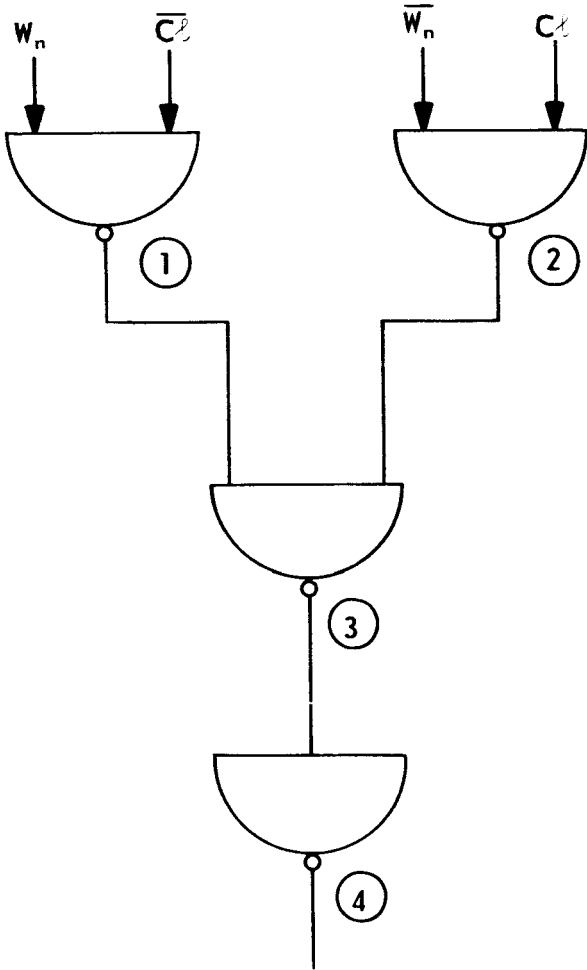
The long-code combiner shown in Figure A-4.3 illustrates the circuitry used to combine the X , Y , and Z codes into the long code $X \oplus Y \oplus Z$. This diagram is applicable to both the encoder combiner which generates W_m , and the decoder combiner which generates W_m .

The long-code combiner generates by simple logical addition and modulo-2 logic the combination of the three short code generators. The output at 3.5 is the long code $X_n \oplus Y_n \oplus Z_n$ (for the encoder) and $X_m \oplus Y_m \oplus Z_m$ (for the decoder) containing 3255 bits, i.e., the product of the X , Y , and Z short code lengths, $7 \times 15 \times 31$.

The Mod-2 Modulator in Figure A-5 combines, modulo-2, the clock, C1, with the long code, $W_n = X_n \oplus Y_n Z_n$, giving at (3) $W_n \oplus C1$. Addi-

tional Clock energy is added (not shown) to permit received clock phase-lock-on independent of code correlation.

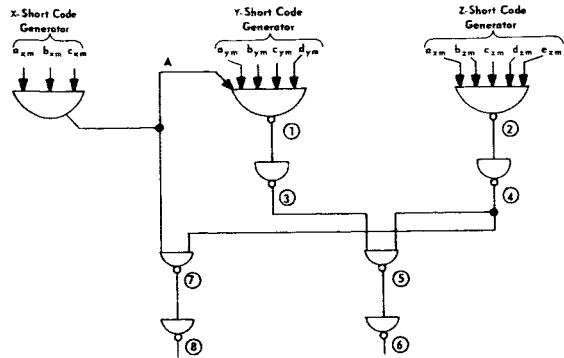
In Figure A-6, the Countdown Generators generate a countdown pulse of 3255:1 at (6) each



1. $F_{n(1)} = \overline{W_n \cdot C1}$
2. $F_{n(2)} = \overline{\overline{W_n} \cdot C1}$
3. $F_{n(3)} = \overline{(W_n \cdot C1) \cdot (\overline{W_n} \cdot C1)}$
 $= W_n \oplus C1$
4. $F_{n(4)} = \overline{W_n \oplus C1}$

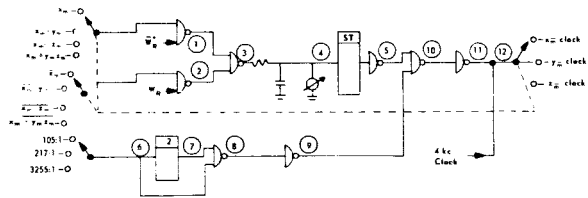
Legend: See Figures 3 and 4
 C1 ≡ System clock input
 W_n ≡ decoder long code

FIGURE A5.—MOD-2 modulator.



- Let $A = a_{2m} \cdot b_{2m} \cdot c_{2m}$
 $B = a_{2m} \cdot b_{2m} \cdot c_{2m} \cdot d_{2m}$
 $C = a_{2m} \cdot b_{2m} \cdot c_{2m} \cdot d_{2m} \cdot e_{2m}$
1. $F_{n(1)} = \overline{A \cdot B}$
 2. $F_{n(2)} = \overline{C}$
 3. $F_{n(3)} = \overline{A \cdot B} = \text{Countdown of } 105:1$
 4. $F_{n(4)} = \overline{C}$
 5. $F_{n(5)} = \overline{A \cdot B \cdot C}$
 6. $F_{n(6)} = \overline{A \cdot B \cdot C} = \text{Countdown of } 3255:1$
 7. $F_{n(7)} = \overline{A \cdot C}$
 8. $F_{n(8)} = \overline{A \cdot C} = \text{Countdown of } 217:1$

FIGURE A6.—Countdown generators.



1. $F_{n(1)} = \overline{W_n \cdot X_n}$
 2. $F_{n(2)} = \overline{W_n \cdot Y_n}$
 3. $F_{n(3)} = \overline{(W_n \cdot X_n) \cdot (W_n \cdot Y_n)}$
 4. $F_{n(4)} = 1$
 5. $F_{n(5)} = 1$
 6. $F_{n(6)} = CD$
 7. $F_{n(7)} = \overline{CD}$
 8. $F_{n(8)} = \overline{CD} \cdot \overline{CD}$
 9. $F_{n(9)} = CD_2 \cdot \overline{CD} \cdot \overline{CD_2}$
 10. $F_{n(10)} = \overline{CD_2} \cdot 1$
 11. $F_{n(11)} = \overline{CD_2} \cdot 1$
 12. $F_{n(12)} = \overline{CD_2} \cdot 1 \cdot C1$
- Legend: Subscript R refers to received code
 1 Continuous correlation
 CD Countdown pulse
 \overline{CD} Countdown pulse divided by two (used as extra shift pulses added to clock shift pulses)

FIGURE A7.—Correlation indicator and shift pulse generator.

time the code is repeated and all generators contain a logical "1" in the set position; a countdown pulse of 105:1 at (3) when all the set outputs of the X_m and Y_m generators are in the logical "1" position; and a countdown pulse of 217:1 at (8) when all the set outputs of the X_m and Z_m generators are in the logical "1" position.

Figure A-7 shows both the Correlation Indicator and the Shift Pulse Generator. By combining, separately, X_m , $X_m \oplus Y_m$, $X_m \oplus Z_m$ with W_R (the received code), mod-2, and then integrating, correlation of each W_m -short code with the W_R -long code can be detected and displayed on the meter at (4). When the codes are not correlated, the voltage reading at (4) is approximately -3 volts. When one of the W_m -short codes is correlated with its corresponding component in the received W_R -long-code, the voltage rises to approximately -1.5 volts. Note

that on complete correlation of W_m with received W_R , the voltage reading at (4) is 0 V, (because with mod-2 logic, all ones or all zeros gives an output of 0 volts.)

The Schmitt Trigger, sensitive to the change in Voltage (from -3v to -1.5 volts), triggers gate (5) which then turns off gate (10), stopping the shift pulse (9). The output at (12) then changes from $F(12) = CD \cdot \bar{I} + C1$ to $F(12) = C1$ (where CD = Countdown, I = integration, and $C1$ = clock). At complete correlation, the Decoder runs synchronously with the received code at the received clock-plus-Doppler rate.

The Comparator, Stop Pulse Generator, and Time Interval Unit shown in Figure A-8 function to provide a range reading.

On demand from WWV, the TIU is activated and simultaneously gate (12) is triggered, stopping the system clock input. The transmitted code

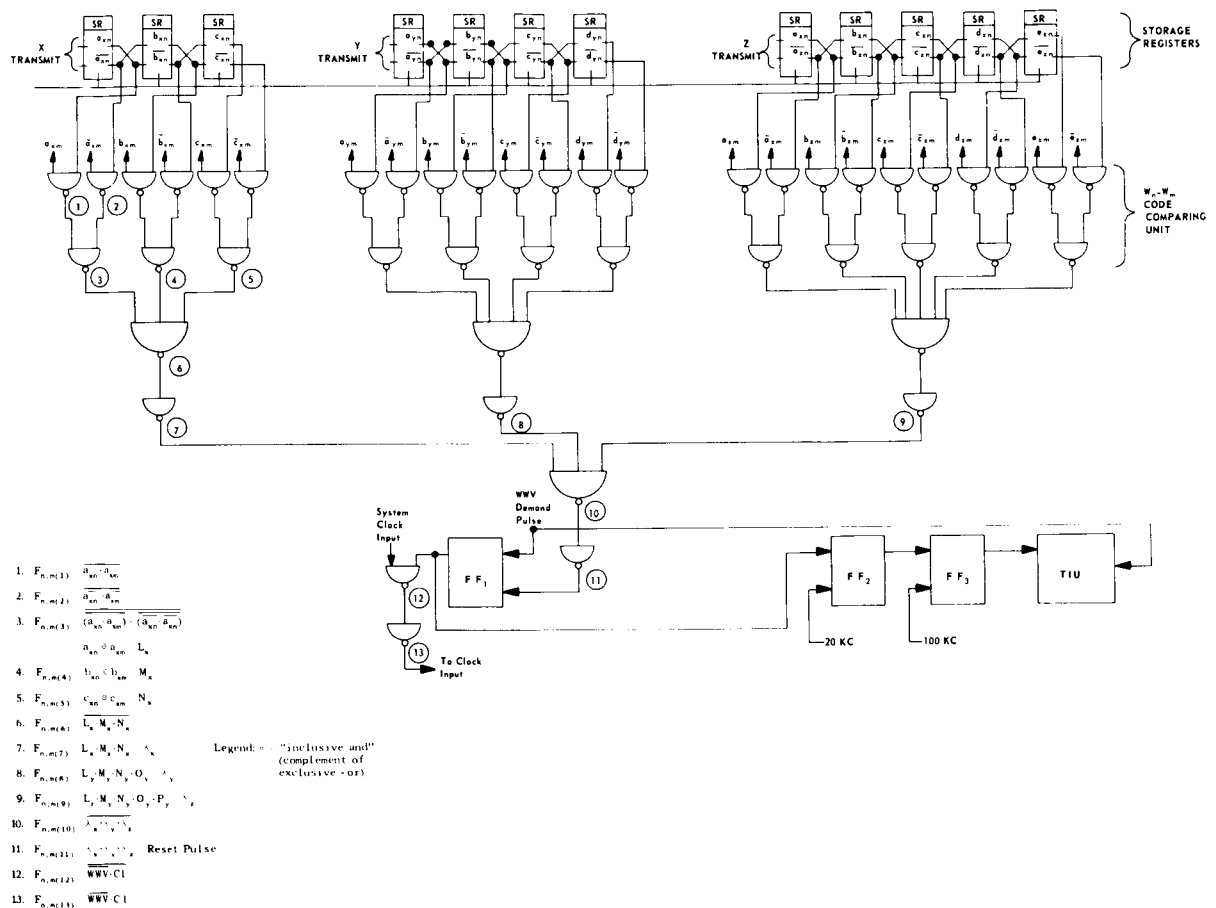


FIGURE A8.—Comparator and stop pulse generator.

(the 12-bit Encoder word) is stored in the registers of the Comparator. The demand pulse is timed to occur on the negative going portion of one of the 4 Kc square wave clock cycles or 125 microseconds ($\frac{1}{2}$ period) after the previous positive going portion has shifted both the short-code generators and Storage Registers. This technique insures that the Storage Registers contain the proper transmitted code, i.e., the three, four, and five-bit Encoder words. The next shift changes the contents of the generators but does not change the contents of the registers. This shift occurs 125 microseconds after demand time, on the next positive going portion of the 4 Kc square wave. Note that during the 125 microsecond period previous to demand time, the Encoder three, four, and five-bit words have been "on the air." The total transit time τ , is, therefore, the measured transit time, τ' , plus 125 microseconds.

By means of the Comparator gates, the locally generated model of the received code is compared with the stored code. When matched, a reset pulse is generated which turns off FF_1 . Simultaneously, FF_2 is turned on. When the positive going portion of the 20 Kc frequency tone occurs, FF_2 is turned off. At the same time, FF_3 is turned on. In the manner described above, the 100 Kc frequency tone will turn off FF_3 and simultaneously the TIU will be deactivated. A constant time interval consisting of the sum of the half periods of the 100, and 20 Kc sidetones is

included in the above transit time reading. This constant time is subtracted to obtain the true distance in transit time units.

BIBLIOGRAPHY

1. "Development of a Range and Range Rate Spacecraft Tracking System," HABIB, KRONMILLER, ENGLIS, FRANKS, GSFC Technical Note D-2093, August 1963.
2. "Syncom Range and Range Rate System Design Evaluation Report," 6 October 1961, Contract No. NAS 5-1605, Space Technology Laboratories, Inc.
3. "Goddard Range and Range Rate System Design Evaluation Report," 23 November 1962, Contract No. NAS 5-1926, Motorola, Inc.
4. "High Accuracy Satellite Tracking Systems," 31 March 1961, Contract No. NAS 5-1187, Advanced Communications Research and Development (ADCOM, Inc.).
5. "High Accuracy Satellite Tracking Systems," 30 September 1961, Contract No. NAS 5-1187, Advanced Communications Research and Development (ADCOM, Inc.).
6. *Error Correcting Codes*, W. W. PETERSON, M.I.T. Press and John Wiley and Sons, Inc., 1961.
7. "On Coded Phase Coherent Communications," A. J. VITERBI, Jet Propulsion Laboratory Technical Report 32-35.
8. "Coding Theory and Its Application to Communications Systems," BAUMERT, EASTERLING, GOLOMB and VITERBI, Jet Propulsion Laboratory Technical Report 32-67.
9. Research Summaries 36-1 to 36-14, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California.
10. Space Program Summaries 36-11 to 36-20, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California.

PHOTOEFFECTS AND RELATED PROPERTIES OF SEMI-CONDUCTING DIAMONDS*

C. JOHNSON,† H. STEIN,‡ T. YOUNG,§ J. WAYLAND|| AND W. LEIVO**

A number of properties of natural semiconducting diamonds have been investigated. The studies were concerned with the optical transmission, photoconductivity, mobility and activation energies of carriers, also the lifetimes and trapping of carriers. Optical transmission as a function of the temperature was measured in the region $200\text{ m}\mu\text{--}13\mu$. Optical absorption characteristic of type IIb diamond was observed at $2\cdot43$, $3\cdot40$, $3\cdot56$ and $4\cdot07\ \mu$ with slight peaks elsewhere; in other respects the absorption was similar to that of a type IIa diamond. Photoconductivity was determined as a function of excitation wavelength, intensity of illumination, temperature and electric field strength. A number of maxima in photoconductivity occur in the visible and near infrared ranges which are not resolved in the optical transmission measurements. The photoconductivity increases with increasing temperature until approximately 400°K after which it decreases. The maxima become better resolved and shift to shorter wavelengths with decreasing temperature. The long wavelength side of the photoconductivity in the range $0\cdot4\ \mu$ to $2\cdot4\ \mu$ may correspond to an acceptor level at $0\cdot52\ \text{eV}$. Hall measurements show *p*-type conductivity with a room temperature hole mobility of $1300\ \text{cm}^2/\text{V sec}$. Assuming a specific model and using the Hall data, an activation energy of $0\cdot35\ \text{eV}$ can be obtained. This corresponds to the optical absorption at $0\cdot35\ \text{eV}$, and it is in agreement with temperature resistivity measurements. Lifetimes and trapping times of carriers were determined by various methods. A number of lifetimes and/or recombination times ranging from $9\ \mu\text{ sec}$ to $84\ \text{min}$ were obtained.

1. INTRODUCTION

The optical and electrical properties of natural diamonds vary considerably from specimen to specimen. In fact, the continuation variation of the characteristics to several extremes was the origin of the general classification into the three general groups—type I, type IIa and type IIb. The early classification of diamonds into types I and II was mainly determined by their optical absorption. The type II diamonds transmit light further into the ultraviolet range than do the type I, also the type I diamonds have optical absorption near 9μ which is not present in type II diamonds. In 1952 CUSTERS¹ discovered a rare type diamond which could conduct an electrical current

but which otherwise had the characteristics of type II diamonds. Thereupon, Custers separated the type II class into types IIa and IIb. Types I and IIa diamonds are good insulators and type IIb are conductors of electricity.

The optical absorption of types I and IIa diamonds had been studied extensively for many years previous to the discovery of the type IIb diamonds. The origin of much of the optical absorption in types I and IIa diamonds is gradually becoming understood. SMOLUCHOWSKI and LEIVO examined a type IIb diamond (diamond DS-1 in this report) and found that it behaved as an extrinsic semiconductor.² The plot of $k \ln \rho$ against $1/T$ gave a slope of $0\cdot35\ \text{eV}$ where k is the Boltzmann constant, ρ the resistivity and T the absolute temperature. Hall measurements showed a *p*-type conductivity. There have been many recent investigations of type IIb diamonds.³ The unavailability of diamonds with carefully controlled impurities and the very rare occurrence of natural type IIb diamonds complicates the investigation of the electrical and optical properties.

*Published in *Journal of the Physics and Chemistry of Solids*, 25: 827-836, 1964. This research was supported in part by the U.S. Air Force Office of Scientific Research and the U.S. Air Force Cambridge Research Laboratories.

**All authors affiliated with the Research Foundation and Physics Department, Oklahoma State University, Stillwater, Oklahoma, in addition to the following addresses.

†Goddard Space Flight Center.

‡Sandia Corporation, Albuquerque, New Mexico.

§International Business Machines Corporation, Endicott, New York.

||Eglin Air Force Base, Florida.

In order to correlate various effects, it was considered desirable to investigate a number of inter-related properties in the same semiconducting specimens. Reported herein are studies of the optical absorption, temperature dependence of photoconductivity, lifetimes and trapping of carriers and the mobility of carriers in natural semiconducting diamond. The diamonds used in the investigations are of type IIb and originated in the Premier Mine, South Africa. One diamond (diamond DS-2) is cut to rectangular dimensions of $2.3 \times 3.5 \times 6.5$ mm. Diamond DS-2 is blue on one end and nearly free from coloration on the other end. All the other diamonds are blue diamonds.

2. EXPERIMENTAL PROCEDURES AND RESULTS

2.1 Optical absorption

The infrared absorption measurements were obtained with Perkin-Elmer Spectrophotometers Models 12C and 21, and a Beckman Model IR-4. The near-infrared, visible and ultraviolet absorption spectra were taken on Beckman Spectrophotometers Models DK-1 and DK-2. The IR-4 has a resolution of 0.01μ at 10μ , and the DK-2 has a resolution of 2 \AA at 220μ . The temperature of the diamond was lowered by passing dry nitrogen through a liquid nitrogen bath and then around the diamond which was mounted in a metal cryostat. The cryostat was equipped with quartz windows for the visible and ultraviolet regions and with sapphire windows for the infrared region. The temperature was increased in the same manner, the liquid nitrogen bath being replaced by a furnace. The temperature of the diamond was determined to within $\pm 2^\circ\text{C}$ with a thermocouple.

The absorption characteristics of diamond DS-2 are shown in Figs. 1-6. Changes in absorption with temperature are shown in Figs. 3-6. The blue end of diamond DS-2 exhibits optical absorption bands in the $1-4 \mu$ region which are not normally associated with a type IIa diamond. Maxima of the special absorption occur at 2.43 , 3.40 , 3.56 and 4.07μ , with slight peaks elsewhere. The clear end exhibits absorption which is nearly that of type IIa diamonds. Room temperature resistivities using a four probe method are $3.6 \times 10^5 \Omega\text{-cm}$ for the clear end and $65 \Omega\text{-cm}$ for the

blue end. The characteristic absorption in the blue end increases with decreasing temperature. Decreasing the temperature to 150°K changes the wavelength at which the ultraviolet absorption starts to become significant from 236 to $226 \mu\mu$; however, the wavelength at which extinction occurs remains nearly unchanged (Fig. 6). Increasing the temperature to 360°K does not change the wavelength at which the ultraviolet absorption starts to become significant, but it shifts the extinction wavelength from 225 to $231 \mu\mu$. Diamond DS-1 shows the same absorption peaks as the blue end of diamond DS-2.

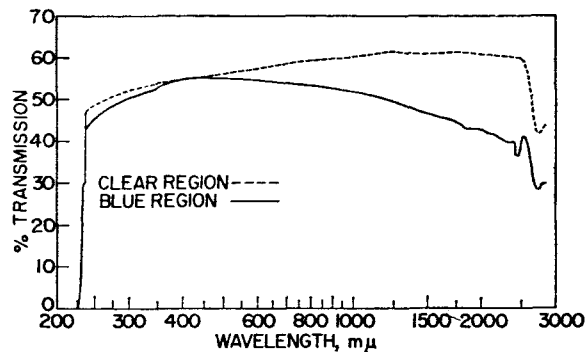


FIGURE 1.—Optical transmission of diamond DS-2 for two separate parts of the specimen.

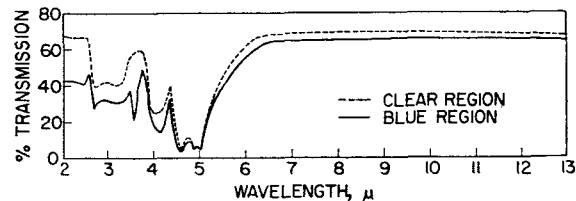


FIGURE 2.—Optical transmission of diamond DS-2 for two separate parts of the specimen.

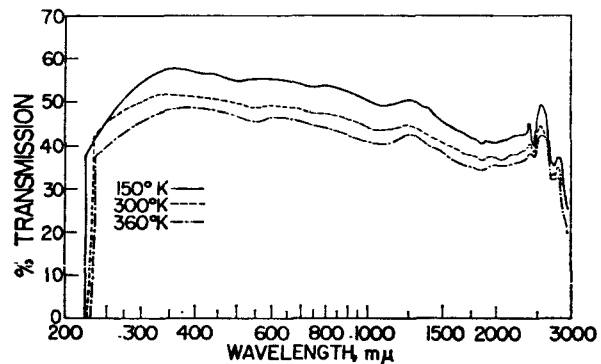


FIGURE 3.—Temperature dependence of the optical transmission of diamond DS-2.

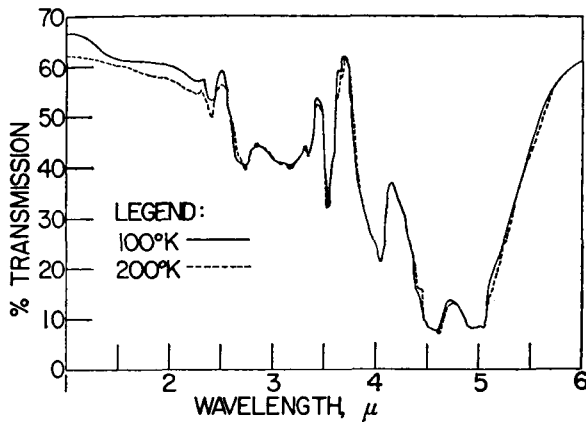


FIGURE 4.—Temperature dependence of the optical transmission of diamond DS-2.

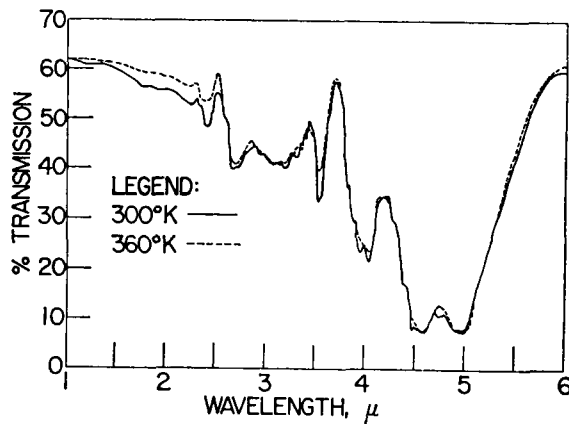


FIGURE 5.—Temperature dependence of the optical transmission of diamond DS-2.

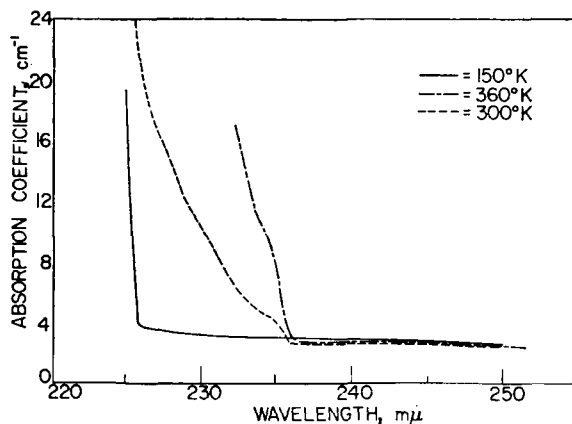


FIGURE 6.—Temperature dependence of the absorption coefficient of diamond DS-2 in the region of the fundamental absorption.

The curves have been adjusted for continuity when changing detectors or light sources. As such, the value of the percent transmission cannot be taken as absolute. The discontinuity in the transmission curve can arise from the following causes: First, light beam is not uniform in intensity and adjusting the focus shifts the region of maximum intensity to different parts of the diamond which is not homogeneous with regard to optical absorption. Second, the detectors may not give uniform response over their surfaces. If the diamond does not have exactly parallel faces the energy may be incident on a different area of the detector after passing through the diamond than the area used in establishing the 100 percent transmission.

The birefringence of the diamond was also examined. The birefringence pattern in the blue part of diamond DS-2 is different from the pattern in the clear region. The clear end exhibits parallel streaks similar to those observed by RAMAN and JAYARAMAN in what were probably type II diamonds.⁴ The blue end of diamond DS-2 exhibits a cross-hatched pattern.

2.2 Photoconductivity

Photoconductivity in semiconducting diamond was studied as a function of intensity of illumination, spectral distribution, electric field strength and temperature. It was found that the semiconducting diamonds, besides being photoconducting in the ultraviolet, have a number of photoconducting maxima in the visible and infrared regions.

Photocurrents were measured with a Beckman DK-1 spectrophotometer which was altered to record photocurrents directly. A tungsten lamp was used for the visible and near infrared source and a hydrogen lamp for the ultraviolet source. The light was chopped at 480 c/s. The spectral distribution of the tungsten light source was determined using a vacuum thermopile with a quartz transmission window. The curve of the spectral distribution at the crystal surface is shown in Fig. 7. The ultraviolet spectral response curve was obtained indirectly with a General Electric FJ-76 phototube. The diamond was placed in the regular sample compartment of the DK-1 spectrophotometer and was shielded from stray

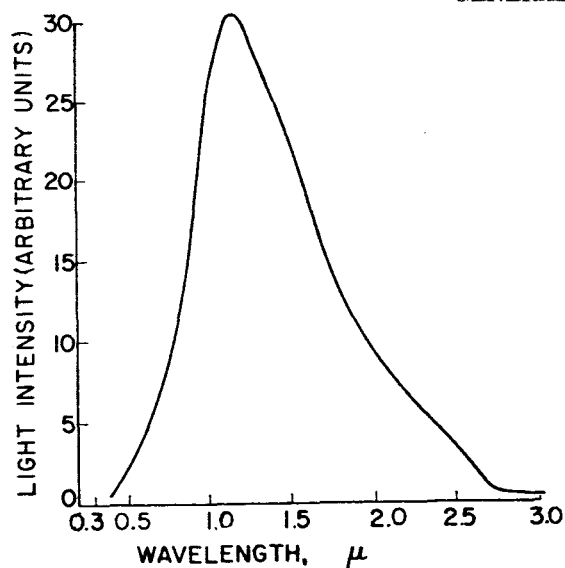


FIGURE 7.—Intensity distribution of the tungsten light source at the exit slit of the monochromator.

radiation. The recorder was calibrated by applying known differences in potential to the amplifiers thus making it possible to record photocurrents directly. The spectral response curves of the photoconductivity in the diamond are reduced to photocurrent per incident photon per second.

The crystal holder for low temperature measurements consisted of a double-walled evacuated container with quartz windows. Sodium chloride windows were used for infrared measurements. The cooling for low temperature work was obtained by passing dry nitrogen through a liquid nitrogen bath and then into the inner container. The currents measured were of the order of 10^{-10} A. Although photoconductivity measurements were made on other semiconducting diamonds with similar results, the results given here were obtained on specimen (DS-2).

2.2(a) *Dependence of photocurrent upon intensity of illumination.* The dependence of the photocurrent in diamond DS-2 was determined as a function of the intensity of the radiation at constant wavelength. It was found that the photocurrent depends linearly upon the intensity of the radiation. A typical curve is shown in Fig. 8. In this case the excitation wavelength was $600 \text{ m}\mu$ and the maximum photocurrent was approximately 10^{-7} A.

Ultimately the sensitivity of a photoconductor is not measured by the current per photon but by

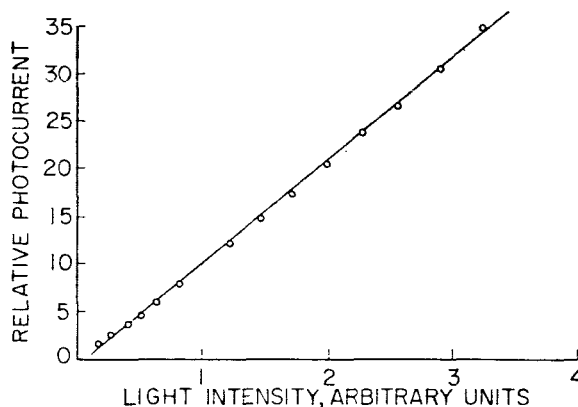


FIGURE 8.—Dependence of the photocurrent in diamond DS-2 upon the intensity of irradiation.

the signal to noise ratio for a given amount of incident radiation.⁵ A good signal to noise ratio can be obtained in the clear end of the diamond at room temperature because the resistivity of this end is such that an appreciable field can be obtained without a very large dark current. A large dark current introduces an appreciable amount of noise. The major source of noise originates with the contacts and is dependent upon the electric field direction. Reverse bias, which occurs when the blue end is negative, introduces a large amount of noise. With forward bias, the noise level is considerably lower, and direct illumination of either or both of the contacts does not add to the noise level of the photocurrent.

2.2(b) *Dependence of photocurrent upon electric field strength.* The dependence of the photocurrent with electric field strength was studied in order to determine the presence of potential barriers in the crystal. The photocurrent is a linear function of the field in the measured range. The field was measured using potential probes thus eliminating the potential drop across the contacts.

2.2(c) *Spectral response of photocurrent as a function of temperature.* The spectral response of the photoconductivity was determined in the wavelength range $0.200\text{--}20 \mu$ at various temperatures between 127 and 450°K . Figure 9 shows the spectral distribution of the photocurrent in the visible range at room temperature. The curve gives the photocurrent per incident photon/sec vs. the wavelength in microns. Upon increasing the sensitivity of the circuit to a maximum and opening the slit to 2 mm , additional

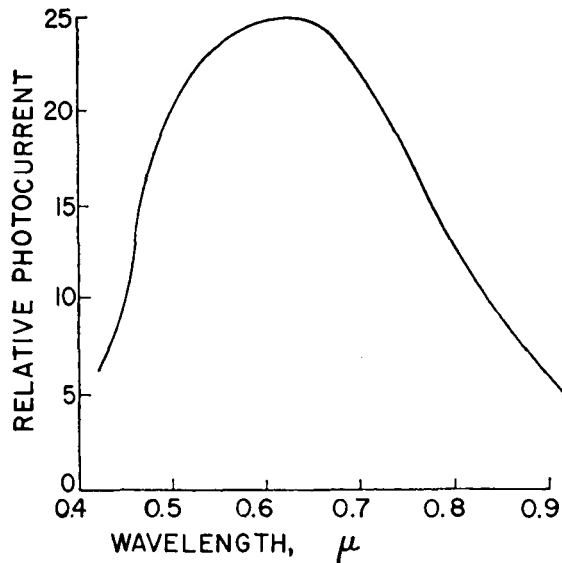


FIGURE 9.—Spectral distribution of the photocurrent in diamond DS-2 at room temperature. The ordinate is proportional to the photocurrent per incident photon/sec.

peaks were found at 1.46 and 1.75μ . The peaks are very weak at room temperature but assume a major role at a low temperature as will be discussed later. A weak photocurrent is obtainable in the region 3.0 – 3.5μ . No photoconductivity was detected in the regions 5.1 – 12μ , or 0.245 – 0.350μ .

Increasing the temperature increases the photocurrent up to approximately 400°K . At higher temperatures the photocurrent decreases as the temperature increases. It should be noted that the increase is an increase in the photocurrent of the broad region of photoconductivity from 0.35 to 3μ , and the half-width is not appreciably increased. The noise level rises sharply at 405°K and increases rapidly with increasing temperature.

The maximum value of the photocurrent decreases in magnitude and shifts to shorter wavelengths upon decreasing the temperature below 300°K . The shift is already observable at a temperature of 250°K (Fig. 10) with the predominant peak appearing at 0.590 instead of 0.600μ . At 223°K an additional peak is resolved at 1.24μ , and the predominant peak shifts to 0.58μ (Fig. 11). The ultraviolet induced photocurrent is no longer measurable from approximately 240°K and lower. When the temperature is reduced to 150°K the predominant peak shifts to 0.52μ . In addition, new peaks appear at 1.02 , 1.62 and

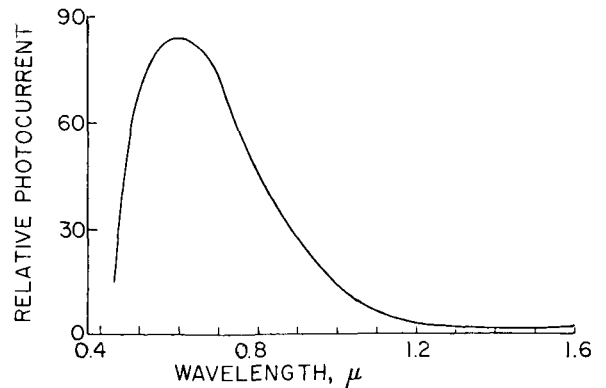


FIGURE 10.—Spectral distribution of the photocurrent in the visible and near infrared regions at 250°K for diamond DS-2.

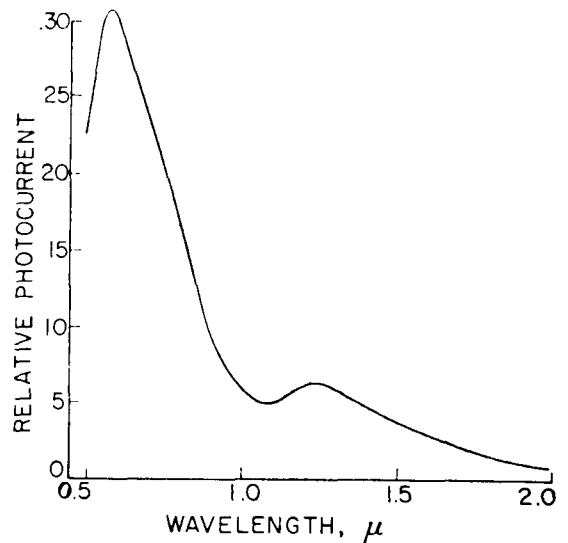


FIGURE 11.—Spectral distribution of the photocurrent for diamond DS-2 at 223°K .

2.18μ (Fig. 12). There is a slight indication of a peak at 1.24μ in the curve which does appear at other temperatures. At a temperature of 127°K , the lowest temperature attained in this study of photoconductivity, the largest peak occurs at 0.88μ . The peak which appears at 0.6μ at 300°K and which shifts to 0.52μ at 150°K has now disappeared; therefore, the region is not shown in Fig. 13. The photoconductivity in the ultraviolet region has been previously reported (last paper under reference (3)).

2.2(d) *Surface vs. bulk photoconductivity.* Measurements were made to determine whether the conduction process is a bulk or a surface phenomenon. The diamond was masked and irradiated

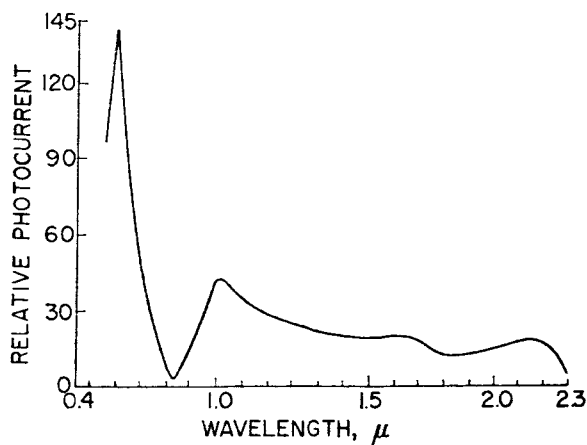


FIGURE 12.—Spectral distribution of the photocurrent for diamond DS-2 at 150°K.

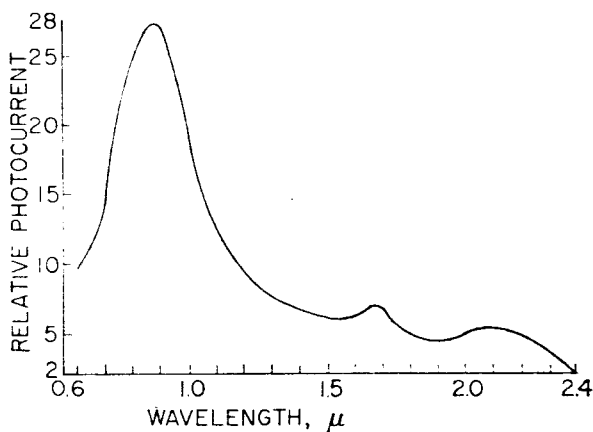


FIGURE 13.—Spectral distribution of the photocurrent for diamond DS-2 at 127°K.

in such a manner that the ratio of the volume to the surface area was different. The photocurrent was the same regardless of which face was illuminated thus indicating a bulk effect. Next, the diamond was masked except for small areas on different faces, and these areas were irradiated. The ratio of the photocurrents was in closer agreement with the ratio of the volumes irradiated than the ratio of the surface areas. This again indicates a bulk effect, but the ratios were not exact. Next, the diamond was masked except for thin strips along the edges. In this case the ratio of the photocurrent was in closer agreement with the ratio of the surface areas rather than the ratio of the volumes. This can possibly be explained by internal reflections from the side surfaces. Although surface conduction cannot

be ruled out, from the data taken one must consider that the conduction process is primarily a bulk phenomenon.

2.3 Lifetimes and trapping of carriers

The generation and decay of carriers responsible for photoconductivity in semiconducting diamonds were studied under varying conditions. The general methods which have been previously employed to study lifetimes and trapping in semiconductors were tried. These include the photoelectromagnetic effect, diffusion length techniques, and photocurrent growth and decay methods. Of the methods tried, only the photocurrent growth and decay gave useful results. Diamond, with an energy gap of 5.6 eV, provides a broader range in which to study intermediate states than do other semiconductors; however, the large energy gap leads to other difficulties. In particular, it is difficult to obtain good data in the fundamental absorption range because of the lack of a suitable intense source of radiation in the ultraviolet region.

From the curves of photocurrent vs. excitation wavelength (Figs. 12 and 13), it is seen that several centers may be responsible for the photocurrent. A study was made of the rate of rise and decay of the photocurrent as a function of the wavelength. In the visible region of the spectrum a tungsten lamp in conjunction with a monochromator was used to excite the photocurrent. The excitation ranged from steady illumination used to determine the steady-state build-up and decay to very short pulses obtained with a camera shutter and chopping disc arrangement. The ultraviolet region was studied using a mercury arc, a hydrogen arc, and short duration sparks. The spark source consisted of a high voltage power supply and a pulse-forming network. The pulse-forming network was discharged between metal electrodes producing an intense short pulse of light which contained a large amount of radiation in the fundamental absorption region of the diamond. The pulse duration was varied by changing the delay line and adjusting the high voltage power supply.

When the sample was exposed to the whole spectrum of the tungsten light decay times of 30 and 39 min were obtained. Using short pulses of light, various recombination times were obtained.

The response and decay times of the photocurrent were found to be independent upon the exciting wavelength. This is attributed to the many intermediate energy levels present in the energy gap of the diamond. Recombination time as used here is the time required for the photocurrent to attain $1/e$ of its steady state value from the initial time of cutoff of the radiation source. Excitation with $222\text{ m}\mu$ radiation resulted in a 12 and an 84 min decay; $440\text{ m}\mu$ produced a 30 min decay and $630\text{ m}\mu$ a 6 min decay. With $800\text{ m}\mu$ radiation, decays of 30 sec and 7 min occurred.

The response time of the photocurrent produced by a short intense spark of light was $10\ \mu$ sec. The response of a photocell to the same light source was approximately $1\ \mu$ sec. The difference between the build-up time of the photocurrent of the diamond and the response time of the photocell is taken to be very close to the free carrier lifetime in the diamond. A free carrier lifetime of $9\ \mu$ sec is therefore obtained in the particular diamond (DS-2) used in this study. The lifetime obtained in this case is believed to be largely independent of trapping.

A detectable photoelectromagnetic effect could not be produced in the sample. This is believed to be due to a very high surface recombination rate and the lack of an intense source in the ultraviolet region. Similarly, diffusion length experiments gave no useful results.

2.4 Hall effect studies

Studies of the Hall effect were undertaken in order to obtain information on the density and mobility of the carriers and also to get some idea of an effective mass of the carriers which would be consistent with a speculative model for the system. The conductivity was due to holes with a room temperature value for the mobility of $1300\text{ cm}^2/\text{V}\cdot\text{sec}$.

Hall effect measurements were carried out with specimen DS-2 in the temperature range from room temperature to 527°C . There was no measurable Ettingshausen effect, and the voltage measured was essentially the Hall voltage. Figure 14 is a plot of the logarithm of the product of the Hall coefficient and the absolute temperature to the three-halves power versus the inverse of the absolute temperature. From the slope of the curve in the low temperature range an activation

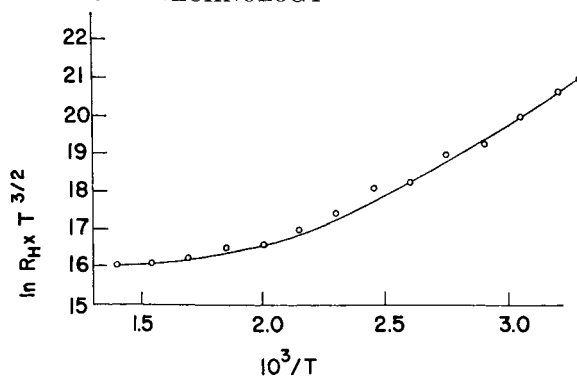


FIGURE 14.—Hall coefficient times $T^{3/2}$ as a function of the inverse absolute temperature for diamond DS-2.

energy of 0.35 eV is obtained. It should be remembered that this cannot be taken with certainty as the activation energy because a specific model for the imperfection states was assumed, and therefore the true position of the Fermi level is uncertain.

A simple model was considered which contains one acceptor and one donor level with the acceptor level located at 0.35 eV . Substituting values for temperature and the corresponding density of holes in the equation

$$p = 3 \left(\frac{N_a - N_d - p}{N_d + p} \right) \left(\frac{2\pi m^* kT}{h^2} \right)^{3/2} e^{-\epsilon/kT} \quad (1)$$

gives three equations in three unknowns, N_a , N_d , and m^* . The symbols are the standard ones. The problem was solved by graphical means. Values for N_a , N_d and m^* were substituted in equation (1), and the density of the carriers was plotted as a function of the absolute temperature. A series of the theoretical curves was compared with the curve obtained experimentally until a good agreement was reached. Good agreement was obtained when the effective mass was 0.3 electron mass.

Equation (1) is based on the condition that the density of electrons in donor states is negligible compared to the density of holes. The factor three in the equation arises from the degeneracy of the levels at the top of the valence bands. In the case of diamond the valence band edge occurs at the center of the Brillouin zone and consists of three bands touching at the central point if spin-orbit coupling is neglected.⁶ The spin-orbit interaction should be very small ($\sim 0.005\text{ eV}$).

The assumption of threefold degeneracy at the top of the valence band and spherical energy surfaces introduces the factor three in the equation.

3. DISCUSSION

From the optical absorption, photoconductivity, and other studies it is apparent that there are a host of energy levels of various types in the energy gap of a natural semiconducting diamond. The optical absorption shown in Figs. 1-6 is for specimen DS-2 which has proved to be interesting in many respects. The specimen being a large parallelepiped is well suited for optical measurements. One end of the specimen is light blue, and the other end is clear to the eye, but the specimen is a single crystal from X-ray analysis. Both regions are semiconductors with the blue region having a room temperature resistivity of $65 \Omega\text{-cm}$ and the clear region $10^5 \Omega\text{-cm}$. Figures 1 and 2 give the optical absorption for the two regions separately. It can be seen that the optical absorption in the clear region approaches that of a type IIb diamond, whereas the blue end shows the extra absorption at 2.43 , 3.40 , 3.56 and 4.07μ along with other small peaks which is characteristic of type IIb diamonds. The demarcation line between the blue and the clear regions is quite sharp, and the birefringence patterns are different in the two regions. It is conceivable from observations of the birefringence patterns that single slip has occurred in the clear region of the diamond and double slip in the blue region. Evidence of slip has been observed in diamonds by TOLANSKY.⁷ Double slip could be more effective in generating vacancies than single slip. In the case of the alkali halides, upon the onset of double slip the rate of generation of vacancies with strain is greatly increased.⁸ This could indicate that the imperfections responsible for the conductivity could be related to a lattice imperfection or possibly an impurity atom associated with a lattice imperfection other than as a direct substitutional impurity. The optical absorption at 3.56μ which corresponds to 0.35 eV does agree with the acceptor activation energy of 0.35 eV obtained from the temperature-resistivity curves.² The determination of activation energies from the temperature-resistivity curves does involve an assumption about the location of the Fermi-level which is actually unknown. WENTORF and BOVENKIRK⁹

produced semiconducting diamonds by diffusing in small amounts of boron, beryllium, or aluminum. Of the three only boron produced a blue color. However, they obtained a value of 0.17 – 0.18 eV from the slope of the resistivity-temperature curves rather than 0.35 eV obtained for diamond DS-2 and also obtained by others in natural semiconducting diamonds. As such, possibly a lattice defect associated with a Group III impurity is responsible for the 0.35 eV level.

The low energy side of the absorption edge of diamond DS-2 (Fig. 6) shifts 0.23 eV between 150 and 360°K compared with an expected change in the energy gap of 0.04 eV . This is indicative of phonon assisted transitions in agreement with CLARK.¹⁰ The optical absorption beyond 0.6μ (Fig. 1) is greater in the region which appears blue to the eye and accounts for the visually observed blue color. The broad region of absorption from 0.6 to 2.8μ is not resolved in the optical measurements, but the photoconductivity measurements (Fig. 13) show a number of maxima in this region.

Photoconductivity studies were undertaken to aid in establishing the positions of energy levels in the energy gap. Since the conductivity is *p*-type, photoionization levels using visible and infrared radiation would correspond to ionization of a hole from a bound state. The presence of a number of maxima in the photoconductivity is similar to the case of germanium.¹¹ The present problem of compensating acceptor levels in diamond with known donors makes it difficult to identify a level with a particular imperfection.

The observations on the temperature dependence of the photocurrent show that the photocurrent increases with temperature until approximately 400°K after which it decreases with increasing temperature. Photocurrents in other materials have been observed both to increase and decrease with temperature depending on the temperature range and the light intensity.⁵ The temperature can be expected to influence the lifetimes of free carriers through the capture cross sections and the number of ground states. Since the distances of the quasi Fermi-levels from the valence and conduction bands depend on the temperature, the higher temperature would bring the levels closer together. If one class of bound states were predominant, this would reduce the number of ground states and result in an increase

in photocurrent with increasing temperature. When there is more than one class of bound states, the photocurrent can increase or decrease with change in temperature as different bound states move in or out of a ground state category.

Part of the optical absorption in the 2–4 μ range may be due to transitions of free holes between the valence bands which, although degenerate at $k=0$, separate for $k \neq 0$. This absorption along with lattice absorption could compete with the transitions producing photoconductivity and thus produce the minima observed in the photoconductivity. Although there is observable photoconductivity in the 2.8–3.5 μ range it is weak compared to the photoconductivity in the 0.6–2.8 μ range. The long wavelength side (2.4 μ) of the photoconductivity in the range 0.4–2.4 μ (Fig. 13) may correspond to an acceptor level at 0.52 eV. This also appears in the optical absorption (Fig. 3) as a peak at 2.43 μ (0.51 eV). As seen from Fig. 3 the absorption at 2.43 μ is temperature dependent with the absorption increasing with decreasing temperature. A thermoluminescence level at 0.52 eV was obtained by HALPERIN and NAHUM.¹²

The photocurrent increases linearly with the applied electric field which is the more usual case with other photoconductors. Even though a potential barrier must exist between the blue and the clear regions of the diamond because of their different densities of holes, the barrier is such as to not affect the linearity of the curve in the measured range.

The photocurrent produced by radiant energy in photoconductors in general may vary as a linear function of the light intensity or as a power different from unity. In some cases the increase in intensity decreases the lifetimes of the carriers, and therefore the photocurrent increases as a fractional power of the intensity.⁵ Figure 8 shows that the photocurrent increases linearly with the intensity of the radiation, thus indicating the possibility of a uniform distribution of bound states or the presence of bound states with different cross sections.

The studies of lifetimes and trapping times show the presence of many long lifetime traps. Therefore, care must be taken in interpreting experiments involving equilibrium conductivity data. Although the longest trapping time measured was

84 min there is evidence of trapping times of over one day.

The determination of an activation energy of 0.35 eV from the Hall data is in agreement with the 0.35 eV obtained from temperature-resistivity measurements, and it also corresponds with the optical absorption at 0.35 eV. The value 0.3 electron mass obtained from Hall measurements for the effective mass of a hole can be compared with the 0.7 electron mass for the effective mass of a hole obtained from cyclotron resonance measurements.¹³ The determination of an effective mass from the Hall data is dependent upon the particular model which is chosen. General agreement as such is helpful in establishing reasonable models.

ACKNOWLEDGMENTS

We thank Dr. J. F. H. CUSTERS, Research Consultant, Industrial Distributors (1946) Limited, South Africa, for the assistance he has given us in this work and for making several semiconducting diamonds available for our studies. We also thank Dr. GEORGE SWITZER of the U.S. National Museum for the loan of a semiconducting diamond.

REFERENCES

1. CUSTERS, J. F. H., *Physica* **18**, 489 (1952); **20**, 183 (1954).
2. LEIVO, W. J. and SMOLUCHOWSKI R., *Phys. Rev.* **98**, 1532 (1955).
3. BROPHY, J. J., *Phys. Rev.* **99**, 1336 (1955); CUSTERS, J. F. H., *Nature, Lond.* **176**, 173 (1955); AUSTIN, I. G. and WOLFE, R., *Proc. Soc., Lond.* **B69**, 329 (1956); CLARK, C. D., DITCHBURN, R. W. and DYER, H. B., *Proc. Roy. Soc. A234*, 363 (1956); WOLFE, R. and WOODS, J., *Phys. Rev.* **105**, 921 (1957); CHARETTE, J. J., *Physica* **25**, 1303 (1960); BELL, M. D. and LEIVO, W. J., *Phys. Rev.* **111**, 1227 (1958).
4. RAMAN, C. V. and JAYARAMAN, A., *Proc. Ind. Acad. Sci.* **A32**, 65 (1950).
5. ROSE, A., *Proc. Inst. Radio Engrs.* **43**, 1850 (1955).
6. HERMAN, F., *Phys. Rev.* **88**, 1210 (1952).
7. TOLANSKY, S., *Phil. Mag.* **44**, 514 (1953).
8. VAUGHN, W., LEIVO, W. and SMOLUCHOWSKI, R., *Phys. Rev.* **110**, 652 (1958).
9. WENTORF, R. and BOVENKIRK, H., *J. Chem. Phys.* **36**, 1987 (1962).
10. CLARK, C. D., *J. Phys. Chem. Solids* **8**, 481 (1958).
11. BURSTEIN, E., DAVISSON, J., BELL, E., TURNER, W. and LIPSON, H., *Phys. Rev.* **93**, 65 (1954).
12. HALPERIN, A. and NAHUM, J., *J. Phys. Chem. Solids* **18**, 297 (1961).
13. RAUCH, C. J., *Phys. Rev. Lett.* **7**, 83 (1961).

A RUTILE TRAVELING WAVE MASER*

C. CURTIS JOHNSON

This paper introduces a new type of traveling wave maser, in which the active material is chromium doped rutile (TiO_2) and the slow wave structure is a meander line. The basic components of a maser, i.e. the active material, the slow wave circuit and the reverse isolation requirements are discussed. This device has a 10°K noise temperature. Data are presented on the gain realizable over a 250 Mc tunable bandwidth.

INTRODUCTION

The name *maser* is an acronym describing the device's operating principle: Microwave Amplification by Stimulated Emission of Radiation. The mechanics of stimulated emission was predicted by Einstein¹ in 1917; and a publication by Weber² in 1953 proposed that the phenomenon could be utilized as a gain mechanism. Since the conception of a practical three-level solid-state maser by Bloembergen³, maser technology has progressed rapidly. The impetus for this accelerated research arises mainly from two characteristics of the maser: it is the most sensitive low-noise microwave amplifier available today, and it is a reliable device which can be ruggedly constructed.

Stated very briefly, the maser principle is as follows: When a paramagnetic ion is placed in a magnetic field, its electron spin systems may assume only the discrete energy states allowed by the principles of quantum mechanics. Under equilibrium conditions more electron spins will be in the lower than in the higher energy states, and to obtain maser action, this situation must be reversed. This *population inversion* is accomplished by *pumping*—that is, providing radiated energy from an external source of the proper frequency to elevate the electron spin systems in the lower energy levels to higher energy states. Such an inverted population can be used as a gain mechanism, since electromagnetic waves of the proper frequency, interacting with these excited electrons, will stimulate them to return to the lower energy levels, emitting photons of the same

frequency. The latter add to the energy of the incident wave, which is thus amplified.

The most complex problem which had to be solved in the development of masers was the retention of the excited energy states, since atoms tend to return to their lowest energy states through spontaneous emission of radiation. The mechanism utilized in the microwave maser is the thermal isolation of the paramagnetic ions. At very low temperatures (e.g., that of liquid helium, 4.2°K) these ions are substantially free of thermal agitation. This effective isolation extends the time that an ion may remain in an excited state.

The first cavity masers had many disadvantages such as non-tunability, relatively poor gain stability, and the necessity of a circulator, which resulted in insertion loss before the maser and hence a higher noise figure. The traveling wave maser provides a workable solution to these problems. The first traveling wave maser (TWM) was developed by Bell Telephone Laboratories⁴. The majority of TWMs designed to date have utilized, as the active element, ruby (Al_2O_3) with an impurity doping of trivalent chromium. Many other materials have been investigated; for example, chromium-doped potassium cobalticyanide⁵ and cerium-doped gadolinium ethyl sulfate⁶ have operated successfully.

This paper is concerned with the active material *rutile* (TiO_2) doped with trivalent chromium as the impurity. The paramagnetic spectrum of Cr^{+++} in rutile was investigated by Gerritsen.⁷ Preliminary investigations of rutile as a maser active material were carried out by Geusic⁸ and continuing effort by L. Morris⁹ has resulted in additional information on this material.

*Published NASA Technical Note D-2490, September 1964.

TRAVELING WAVE MASER, GENERAL CONSIDERATIONS

The final design of a TWM is a complex combination of the active material, the slow wave circuit, and reverse isolation to provide the desired gain and bandwidth of the maser.

Active Material

The characteristics of the active material will, to a great extent, determine the maser's performance. The most obviously significant properties are:

1. zero-field splitting (should be large);
2. spin lattice relaxation time (should be long);
3. density of spin systems (should be large);
4. recovery time (should be short).

Important properties of less obvious significance are:

5. ability to withstand multiple temperature cycling from room temperature to 4.2°K;
6. availability in reproducible crystalline orientations;
7. correct and reproducible impurity concentrations;
8. machineability to accurate dimensions;
9. chemical inertness.

Slow Wave Circuit

Amplification takes place in a TWM when the electron spin systems of the paramagnetic impurity in the host material are acted upon by an electromagnetic wave as it travels along a periodic structure, or *slowing structure*. Since the gain of the maser is proportional to the interaction time between the input signal and the excited spin system, the characteristics desired in the slow wave circuit should be a large amount of slowing over a wide passband.

Nonreciprocal Forward Gain

One of the most important characteristics of the TWM is the ability to obtain nonreciprocal forward gain and nonreciprocal reverse isolation. The reverse isolation must be great enough that

the system will be unconditionally stable against any combination of input and output mismatch.

The transition between spin energy levels is strongly stimulated when the incident electromagnetic field is approximately circularly polarized. The propagating circuit is such that this field will have the circular polarization orthogonal to the direction of propagation. These waves exhibit an opposite sense of polarization above and below the circuit plane. Therefore, if an absorbing material of the proper polarization (a ferrite slab) is appropriately placed in this field, the backward wave is attenuated. In addition the ferrite must be of such composition and dimensions that it will "track" (that is, when the magnetic field, and hence the operating frequency, is changed, the ferrite must provide absorption of the backward wave at the new field and frequency).

Microwave Energy Source

This area of the maser is the simplest problem to solve. Analysis of the spin hamiltonian of the active material will provide the exact frequency or range of frequencies necessary for inversion of the spin system population. For rutile maser operation at 2-3 Gc this requirement is for a signal source in the 50 Gc range, which is a commercially available unit.

CHARACTERISTICS OF CHROMIUM DOPED RUTILE

Chromium doped rutile (TiO_2) is an extremely hard crystal with a tetragonal structure. Its dielectric constant is large (120 perpendicular to and 220 parallel to the C-axis at 4.2°K). When the chromium is added and the crystal annealed, the material becomes opaque and X-ray analysis is necessary to provide crystalline axis information. There are two titanium (or chromium) ions per unit cell (double ion site), and these ions are related by a 90 degree rotation about the Y-axis. Operation in this double ion site where the chromium ions in each unit cell are acted on simultaneously will result in an increase over the gain normally obtained by a factor of two.

Figures 1, 2, and 3 show the computer analysis of the spin hamiltonian of the Z, X, and Y-axis energy levels for chromium-doped rutile.⁹

One way to achieve enhanced maser performance is to increase the idler to signal frequency

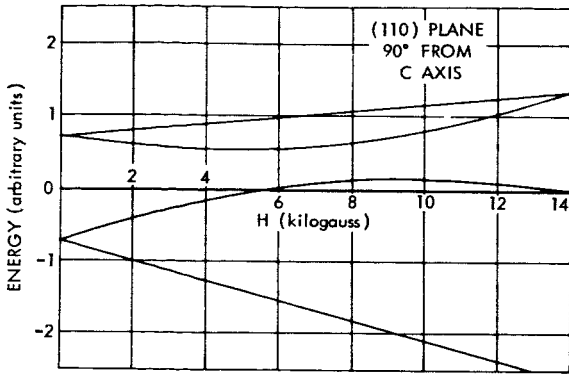


FIGURE 1.—Z-axis energy levels for chromium doped rutile.

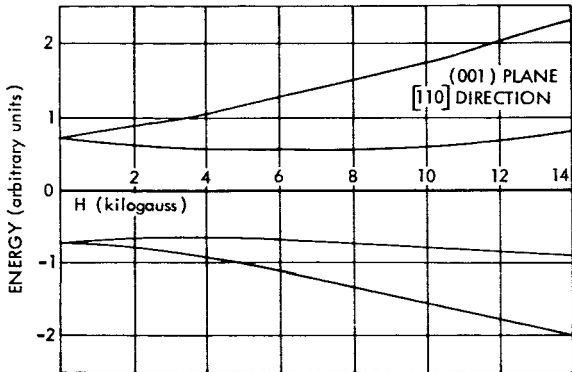


FIGURE 2.—X-axis energy levels for chromium doped rutile.

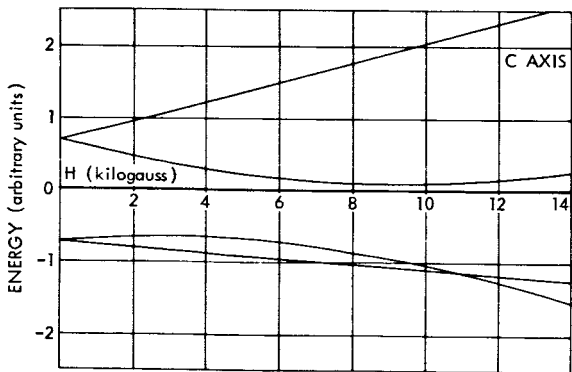


FIGURE 3.—Y-axis energy levels for chromium doped rutile.

ratio. Since the zero-field splitting for rutile⁷ is 43.5 Gc and is only 11.5 Gc for ruby, greater idler-to-signal ratios are obtainable for rutile. It is apparent from Equation 1, since the inversion ratio is a function of F_i/F_s , that higher zero splitting results in higher F_i .

The inversion ratio, which represents a positive means for evaluating maser crystal performance,

is defined as the ratio of the maser's gain with the pump power applied to its absorption with the active material in thermal equilibrium. The inversion ratio takes the form

$$I = \frac{\frac{T_s}{T_i} \cdot \frac{F_i}{F_s} - 1}{\frac{T_s}{T_i} + 1} \quad (1)$$

where

T_s = signal relaxation time,

T_i = idler relaxation time,

F_i = idler frequency (pump frequency minus signal frequency),

F_s = signal frequency.

The maser gain takes the form

$$G(\text{db}) = \frac{27.3 N S}{Q_m} \quad (2)$$

where

S = slowing factor c/v_g (v_g is the group velocity of the propagating wave),

N = number of free-space wavelengths in the total length of active material, and

$Q_m = \omega(W_s' dz)/dP$ = magnetic Q of the system

or the reciprocal of the ratio of the power P emitted by the spin system in a length dz to the energy W_s' stored per second in the length dz , for the frequency ω . The efficiency of coupling the electromagnetic energy of the traveling wave to the maser system is termed the *filling factor* and is incorporated into the term Q_m . The detailed expression for Q_m is:

$$\frac{1}{Q_m} = \frac{2 g^2 \beta^2 \mu_0 h f N \sigma_n^2 I \eta}{\hbar K T n \Delta \ell} \quad (3)$$

$\Delta \ell$ = 3-db line width of the maser material paramagnetic absorption; this is the measured one-half value of the loss component of susceptibility.

η = filling factor,

N = spin density (spins/cm³),

- n = number of energy levels available,
 I = inversion ratio,
 μ_0 = permeability of free space,
 β = Bohr magneton,
 g = spectroscopic splitting factor,
 σ_n = diagonalized portion of transition probability tensor,
 K = Boltzmann constant,
 T = temperature in °K,
 h = Planck's constant/ 2π .

From these equations it follows that for optimum conditions the slowing should be high, Q_m should be as small as possible (the greater I the smaller Q_m), and the filling factor large. The natural 3 db line width of the maser crystal should be reasonably large to assure adequate instantaneous bandwidths. This line width criterion appears to the condition that Q_m be as small as possible; however, this is overshadowed by the condition for bandwidth given by*

$$BW = \Delta\ell \sqrt{\frac{3}{G-3}}, \quad (4)$$

where

$\Delta\ell$ = natural line width,

G = electronic gain in db.

From the foregoing equations the significance of high inversion ratios is apparent: The gain per unit length is thereby increased while the insertion loss is decreased simultaneously. This improvement can provide additional bandwidth, increased gain, reduced size, operation at elevated temperatures, or a combination of these advantages.

SLOWING STRUCTURE

The object of a slow wave structure in a TWM is to hold the incident microwave energy in contact with the paramagnetic ion spin system for a

*This equation is derived on the basis of a Lorentzian shape of the paramagnetic absorption.

sufficient time to maximize the interaction of the microwave energy and the spin system. If the spin system has been elevated to the radiating state by a microwave energy source of the proper frequency, exponential gain will result along the length of the structure.

The two methods of slowing normally considered are resonant and geometric slowing. In the former, a resonant structure is utilized as a means of slowing down the group velocity. The most widely known example of this is the comb structure wherein the slowing is accomplished by forcing the normal waveguide wave to bounce back and forth many times between resonant quarter-wave fingers. Geometrical slowing in its simplest form employs a helix wherein a transmission line of length L is coiled up in such a manner that its actual length is A . In this case the wave slowing factor is L/A .

The slowing structure utilized in this maser is a meander line of 5-mil copper film (see Figure 4)

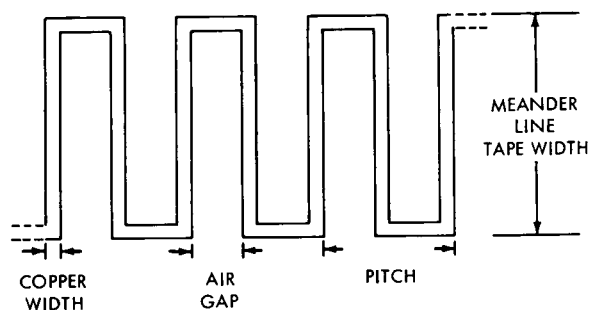


FIGURE 4.—Meander line used as a slowing structure in the rutile maser.

which was first suggested by Siegman⁹ and analyzed by Butcher¹⁰. This structure supports TEM waves for finger lengths of approximately quarter wavelengths. These TEM waves can be thought of as traveling back and forth along the tape at right angles to the slow wave direction of propagation. The technique thus combines resonant and geometrical slowing.

The ratio of air gap width to copper width is a critical parameter for the meander line, since it determines whether the structure is geometric, resonant, or a combination. For the limiting case of narrow tape and large air gap the slowing would be geometrical, and for a very narrow air gap the result is resonant slowing. Proper design of the

meander line combines both techniques, and good slowing over a wide frequency range is accomplished. By adjusting the copper-to-air-gap ratio the maximum tunable bandwidth may be achieved.

The active maser material applies a dielectric loading to the meander line and the exact analysis of slowing as a function of dielectric loading, especially for an anisotropic dielectric, is very difficult. This factor is determined empirically and the broadest bandwidth is achieved with a copper-to-air ratio of 1:1. By trading air for copper, the bandwidth is reduced and the slowing is increased. Therefore, the designer has four variables with which to determine the passband of the structure: (1) the copper-width-to-air-gap ratio, (2) the pitch, (3) the tape width, and (4) the dielectric loading. The latter can be present in several forms: width, thickness, and location of the dielectric with respect to the meander line.

REVERSE ISOLATION

To provide amplifier stability, some form of reverse isolation is necessary. In this device the mechanism is a thin bar of gadolinium doped yttrium iron garnet (YIG). By shaping the ferrite the resonance may be adjusted to track the same magnetic field as needed for the maser crystal. The ferrite must operate at liquid helium temperatures and possess line widths of sufficient magnitude to enable proper tracking, that is, as the operating frequency of the maser is changed by adjusting the magnetic field, the ferrite absorption resonance must follow the magnetic field and continue to provide reverse isolation. Figure 5 shows a number of cross-sectional dimensions for ferrite rods and their change in tracking⁹ as the pure YIG is cooled from 290K to 77K, and then to 4.2K.

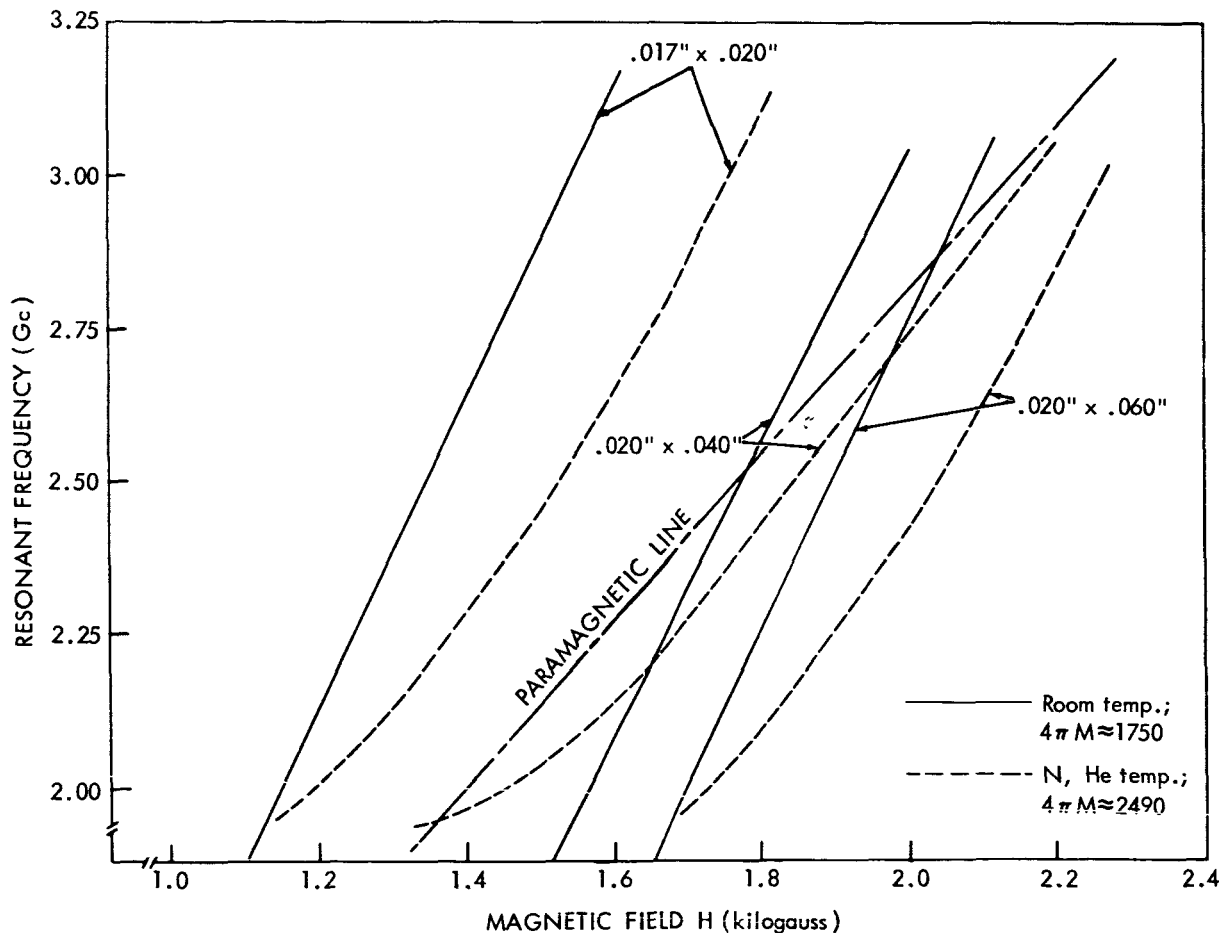


FIGURE 5.—Ferrite "tracking" curves.

The resonant frequency of a ferrite sample is determined by the magnetic field, the inherent saturation magnetization ($4\pi M_s$), and the shape of the sample used. The equation describing these factors is the Kittel equation:

$$f = \gamma \left\{ \left[H + (N_x - N_z)M \right] \left[H + (N_y - N_z)M \right] \right\}^{1/2}$$

where

f = resonant frequency (Mc),

γ = gyromagnetic ratio

H = applied dc magnetic field (gauss),

M = magnetization

N_x, N_y, N_z = demagnetizing factors

Figure 6 is a diagram based on Kittel's formula.⁹ The dashed line shows the Y-axis curve for rutile. The curve is based on pure YIG (yttrium iron garnet) which has a saturation magnetization of $4\pi M_s = 2400$ gauss at 4.2°K . The forbidden regions indicated in the graph are regions where the ferrite sample is not fully saturated and hence lossy.

The saturation magnetization of 10 percent gadolinium doped YIG is approximately 700 gauss. The pure YIG saturation magnetization of 1750 gauss (both values for 4.2°K) would have been too high and resulted in an unsaturated, or lossy, mode for the lower frequencies. It can be seen from Figures 5 and 6 that 10 percent gadolinium doped YIG can be shaped to track rutile and hence provide the necessary reverse isolation.

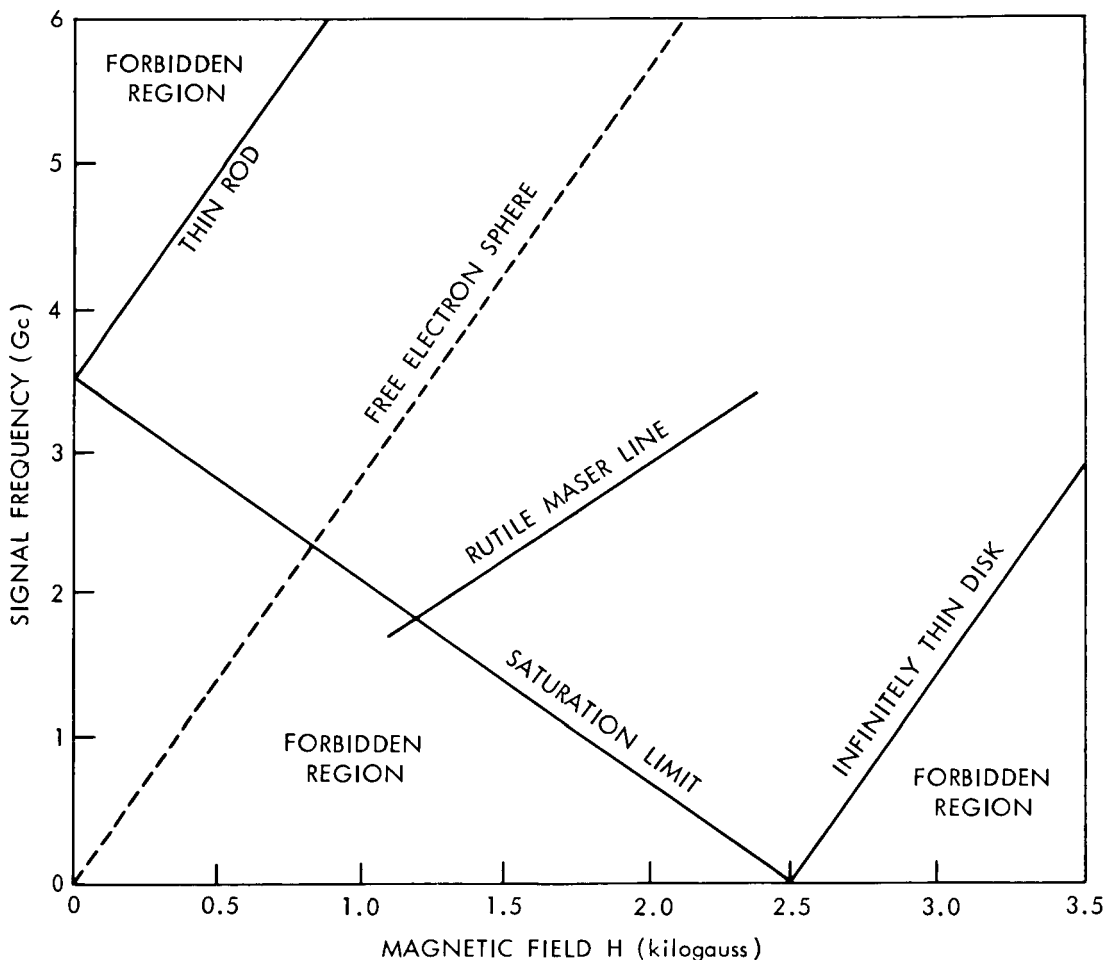


FIGURE 6.—Plot based on Kittel's equation, for pure yttrium iron garnet (YIG) at helium temperature; $4\pi M \approx 2490$.

OPERATING PARAMETERS

The most significant contribution of the maser is its low noise operation. This maser has a noise temperature of 10°K, of which less than 1° can be attributed to the structure. The remaining 9° is encountered in input losses.

The magnetic fields necessary for maser operation in our frequency range (2-3 Gc) are in the vicinity of 1.5 kilogauss and the pump frequency is approximately 50 Gc.

An example of the gain data for a particular crystal orientation is given in Table 1.

Table 1

Example of Gains Obtained with the Maser with a Particular Orientation of the Rutile Active Element.

Operating Frequency (Gc)	Gain (db)
2.05	18
2.1	18
2.15	25
2.2	33
2.25	37
2.3	24

Programs presently underway include an analysis of the inversion ratio as a function of impurity concentration. In addition, magnetic shimming is being investigated in order to provide wider bandwidths. Another phase of the effort to obtain additional bandwidth is an analysis of the results of purposely misaligning portions of the active material and thereby "smearing" the energy levels over a wider range for the same magnetic field.

Experiments with a folded structure will begin soon. This structure allows the rutile to be so placed that, effectively, 12 inches of material require only 6 inches of magnetic field. Optimizing the filling factor, reducing the insertion loss, and providing simpler means of coupling in and out of the meander line with optimum VSWR are further areas under study. A future program under consideration is the evaluation of iron

doped rutile as a maser material. This impurity provides spins of 5/2 and thereby two additional levels for utilization. The program will involve an impurity concentration evaluation and multiple pumping schemes.

CONCLUSIONS

Experimental results indicate that rutile is in fact superior to ruby as an active element. The higher inversion ratios provide higher gain per unit length which may be traded for bandwidth. Bandwidths of 200 Mx and gains of 30 db appear attainable with this material. Considerable investigation will be necessary to optimize this maser configuration and fully realize its potentialities.

REFERENCES

1. EINSTEIN, A., "Zur Quantum Theorie der Strahlung Physik," Leipzig: *Physik. Zeit.* 18:121-128, March 15, 1917.
2. WEBER, J., "Amplification of Microwave Radiation by Substances not in Thermal Equilibrium," *Trans. IRE, Professional Group on Electron Devices-3*, June 1953, pp. 1-4.
3. BLOEMBERGEN, N., "Proposal for a New Type Solid-State Traveling Wave Maser," *Phys. Rev.* 109:324, October 1956.
4. R. W. DEGRASSE, et al., "Three-Level Solid State Traveling Wave Maser," *Bell Syst. Tech. J.* 38:305, May 1959.
5. MACWORTER, A. L., and MEYER, J. W., "Solid-State Maser Amplifier," *Phys. Rev.* 109(2):312-318, January 15, 1958.
6. SCOVIL, H. E. P., FEHER, G., and SEIDEL, H., "Operation of a Solid-State Maser," *Phys. Rev.* 105(2): 762-763, January 15, 1957.
7. GERRITSEN, H. J., and LEWIS, H. R., "Chromium-Doped Titanium as a Maser Material," *J. Appl. Phys.* 31:608, 1960.
8. GENSIC, J. E., "Rutile as a Potential Maser Material," in: *Solid-State Maser Research, First Quarterly Report*, Bell Telephone Laboratories, Whippany, N. J. (ASTIA, AD252544).
9. MORRIS, L., "New Type of Rutile Maser," in: *Conf. Proc. of 7th Nat'l Conv. on Military Electronics*, I.E.E.E., September 9, 1963, p. 174.
10. BUTCHER, G. N., "A Theoretical Study of Propagation Along Tape Ladder Lines," *Proc. IRE* 104:169, March 1957.

WELDED ELECTRONIC MODULE FABRICATION*

JOHN C. LYONS AND DAVID R. DARGO

A pilot facility for the fabrication of welded electronic modules has been developed at the Goddard Space Flight Center for the purposes of: (1) providing a quick-reaction group in the research and development phases of module fabrication, and (2) determining in detail the requirements in terms of facilities and trained personnel to insure the production of high quality welded modules. The experience gained in the successful application of the welding technique to modules in Ariel I (the International Ionosphere Satellite, 1962 o 1) and Explorer XVIII (the Interplanetary Monitoring Platform, 1963 46A) has provided a firm base for the utilization of this concept in future spacecraft programs.

INTRODUCTION

This report presents a general discussion of the activities of the Goddard Space Flight Center in the area of welded electronic module fabrication during the past 2 years. The experience acquired during progress of the work on the International Ionosphere Satellite, Ariel I (1962 o 1) and the Interplanetary Monitoring Platform, Explorer XVIII (1963 46A) forms the basis of the report.

The welding technique offers great advantages over other methods of module fabrication. Since interconnection wire can be welded to component leads immediately adjacent to the component body, greater packing densities are achieved. Since the strength of the weld joint closely approaches the yield strength of the wire itself, mechanical integrity of the module is enhanced. Welded interconnections are not limited to connections on one or two surfaces. Therefore, layout design is simplified and packing density increased by the addition of layers of interconnections as required. The successful operation of the encoders in Ariel I has effectively demonstrated the value of the welding technique in module fabrication. These encoders comprised approximately 2000 components and were about 90 percent welded. A study of the effects of random and single frequency sinusoidal vibration on a test unit encoder for the Explorer XVIII program has been completed. This unit, shown in Figure 1,

contains approximately 1300 components and 8000 welded interconnections. It was found to be unaffected, electrically and mechanically, after being subjected to 27g (rms) random vibration within the frequency spectrum of 20–2000 cps for 4 min. duration along each of 3 mutually perpendicular axes. Sinusoidal vibration sweeps at the rate of 2 octaves per minute at levels of 60g (peak) over the frequency range of 2000–3000 cps, 50g (peak) over the range of 44–3000 cps, and 0.5 in. (double amplitude) over the range of 10–44 cps likewise left no effects upon the unit.

In the interest of greater packing density, the matrix film technique is favored at the Goddard Space Flight Center over the point-to-point technique. Because crossovers are not possible with the point-to-point technique, it does not permit as many connections to be made on a single layer as does the matrix technique. In addition, the point-to-point method requires the fabricator to bend and form the connecting wire as he assembles the unit. This practice produces inconsistent weld characteristics. The matrix film technique simplifies both layout and inspection. Connections can be made on both sides of the film, reducing the number of connecting layers. Individual matrix films may be inspected as they are completed; this permits step-by-step quality control. The problem of bending and forming wire does not occur since all wire runs are straight. This eliminates operator responsibility for their arrangement.

*Published as NASA Technical Note D-2321.

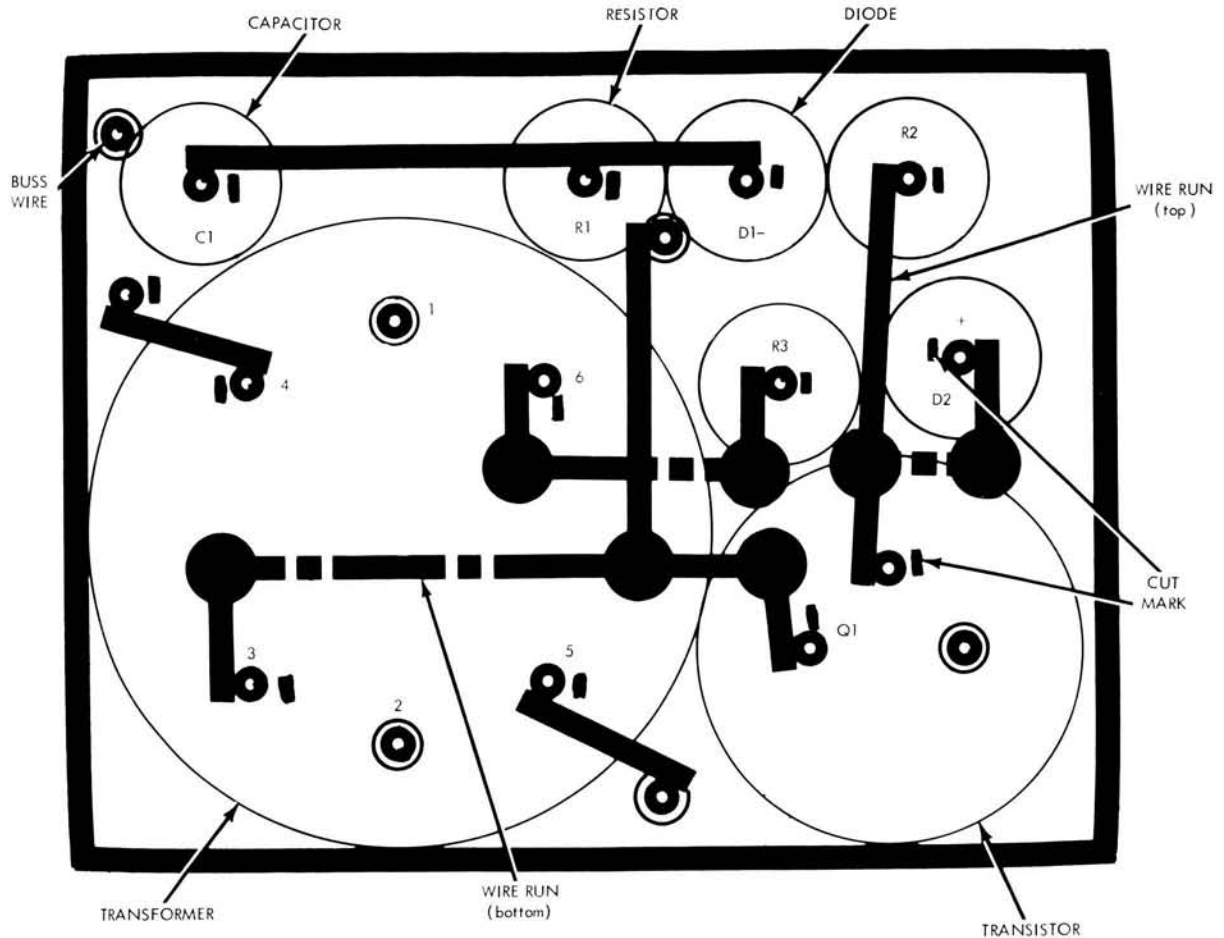


FIGURE 2.—Typical layout for welded modules.

MATERIALS AND EQUIPMENT

It is important that due consideration be given to the choice of materials for component leads, interconnecting wires, and matrix film. Components may be obtained whose leads are either weldable or nonweldable and either magnetic or nonmagnetic. For welded module applications, suppliers should be required to provide components with weldable leads, where possible. Weldable metals should be required to provide components with weldable leads, where possible. Weldable metals which have been found to be satisfactory for component leads and interconnecting wires are nickel, Kovar, Dumet, and Alloy 180. Since the magnetic properties of most commonly used weldable metals would have an influence on the magnetometers included in many satellites, a nonmagnetic weldable metal, such as Alloy 180,

should be used for component leads and interconnecting wires in these applications. Standard stable-base photographic film is adequate for the matrix film.

The shop equipment necessary for a welded module fabrication facility consists of welding machines, microscopes, punch presses, a pull tester, and assorted hand tools. This equipment is shown in Figure 3. The welding machines comprise two major components, a weld head and a power supply. The weld head is a two-electrode device which passes current through the work-piece to make the weld. The power supply is basically a dc supply employing a large capacitor at its output which is discharged through the primary of a step-down transformer. The secondary of the transformer is a low voltage, high current winding connected to the two electrodes.

WELD SCHEDULES

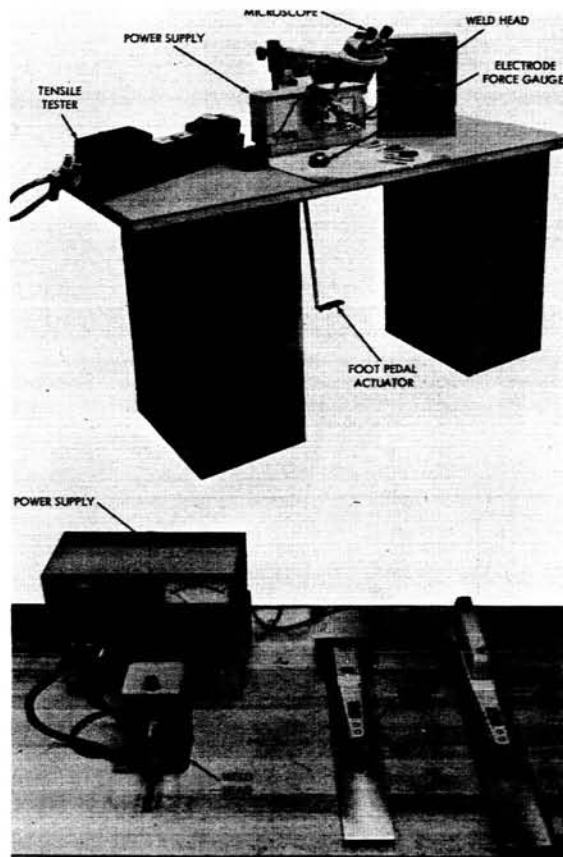


FIGURE 3.—Shop equipment.

The microscopes are of the adjustable focus, variable power, binocular type and are used in observing the welding process and inspecting the welds for quality control purposes. The pneumatically operated pull tester is also an essential quality control tool. A close correlation has been established between external appearance of the welds and the pull-test results.

ENVIRONMENTAL CONDITIONS

It has been effectively shown that welded electronic modules can be successfully fabricated under ordinary laboratory conditions. Therefore, although the industrial practice of providing special workrooms and regulations for welding operations should be encouraged, it is not mandatory to maintain "white room" conditions

Weld schedules specify the equipment to be used, the material and configuration of the electrodes, the electrode pressure to be applied, and the weld energy in watt-seconds for each combination of materials to be welded. A typical schedule is shown in Figure 4. It might be expected that strict adherence to the schedule would enable the welder to produce consistently acceptable welds. It has been found, however, that because of minor daily variations in the performance of machines, coupled with the effect of the basic differences between operators, a standard weld schedule does not insure consistently acceptable welds. Therefore, the pressure and energy settings specified in weld schedules should be considered as rough estimates, defining the range in which acceptable welds can be produced. Figure 5 shows this range.

Visual inspection is the most valuable quality control tool. Metallographic analyses and pull tests may be used to determine the quality of welds in establishing the basic schedule. An analysis requires the examination of photomicrographs of the cross-section of welds, by a metallographer, to determine whether the crystalline structure of the joint is satisfactory. Since neither the external appearance nor the pull strength of a weld can be closely correlated with the results of metallographic analysis, and since welds cannot be sectioned during flight package fabrication, metallographic analysis, although of academic interest, is of little practical value in the workshop. The results of pull tests, flexure tests, etc. may be very closely correlated with external

■ — ACCEPTABLE WELD DIFFICULT TO INSPECT
 ▨ — ACCEPTABLE WELD EASY TO INSPECT

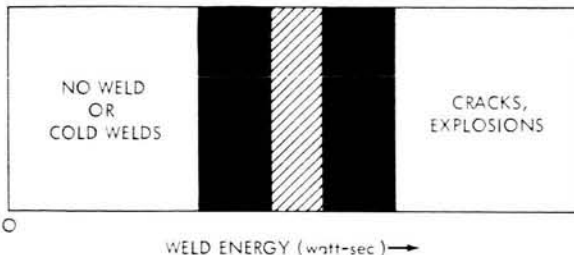


FIGURE 5.—Range of acceptable welds.

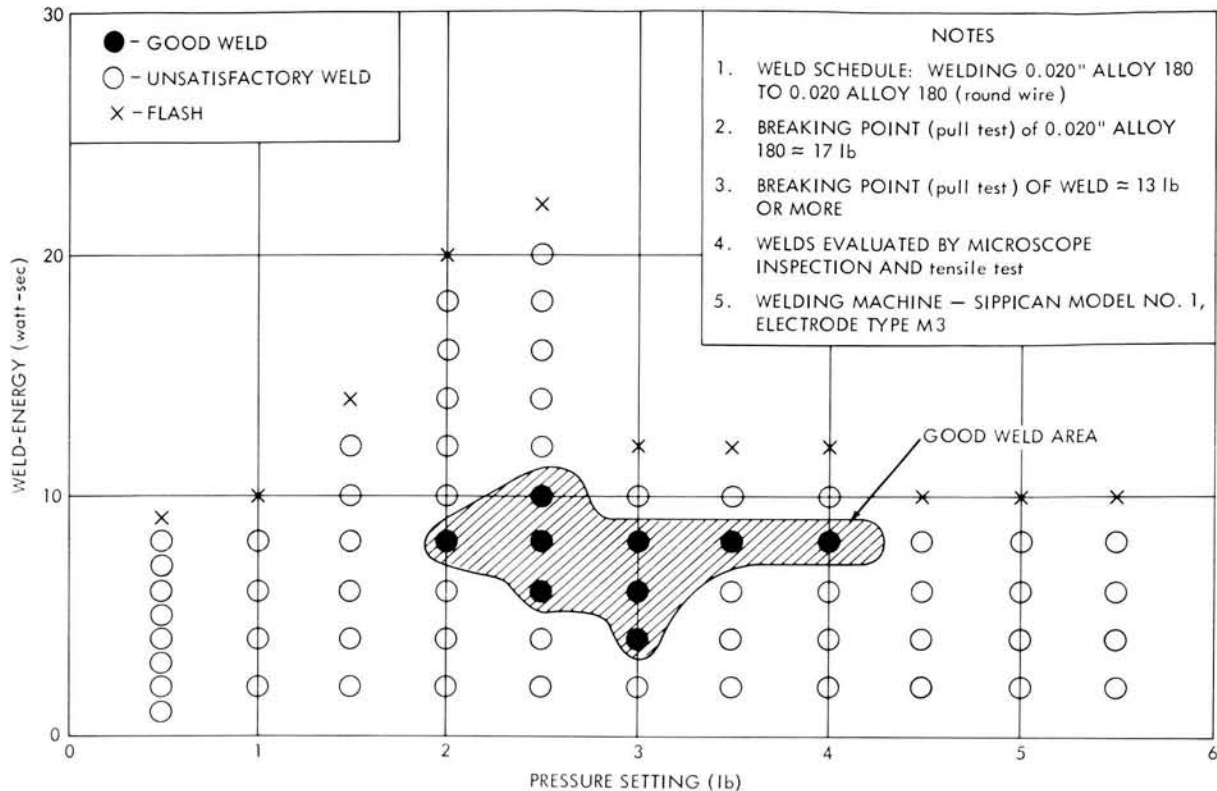


FIGURE 4.—Typical weld schedule.

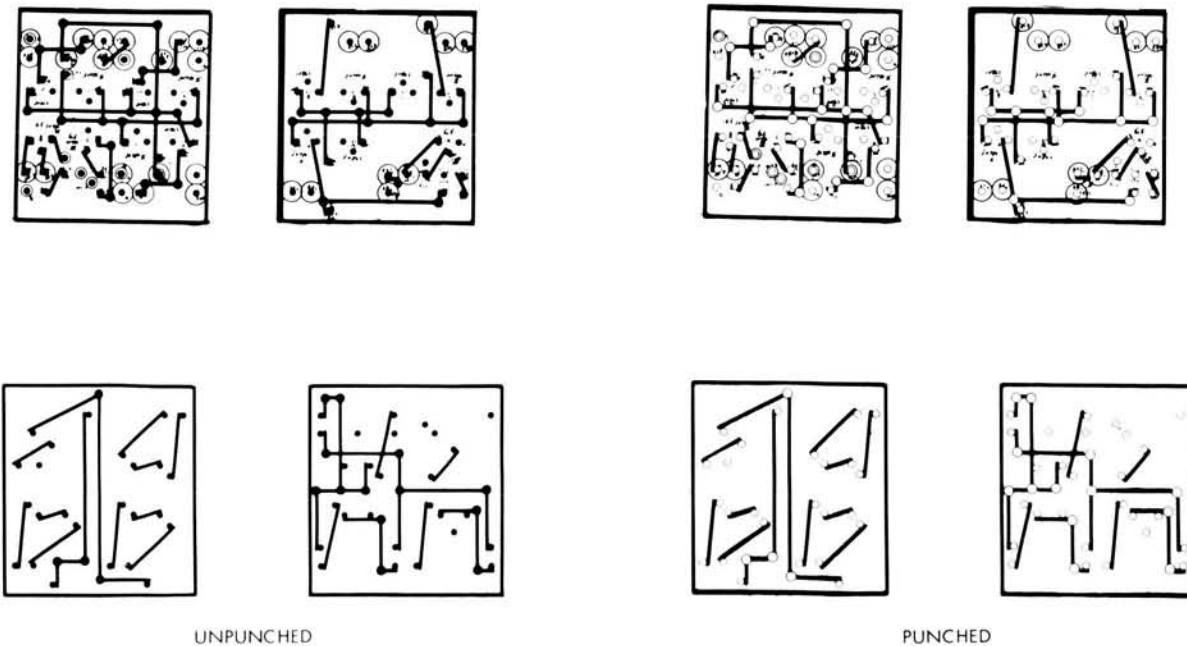


FIGURE 6.—Matrix films for a module.

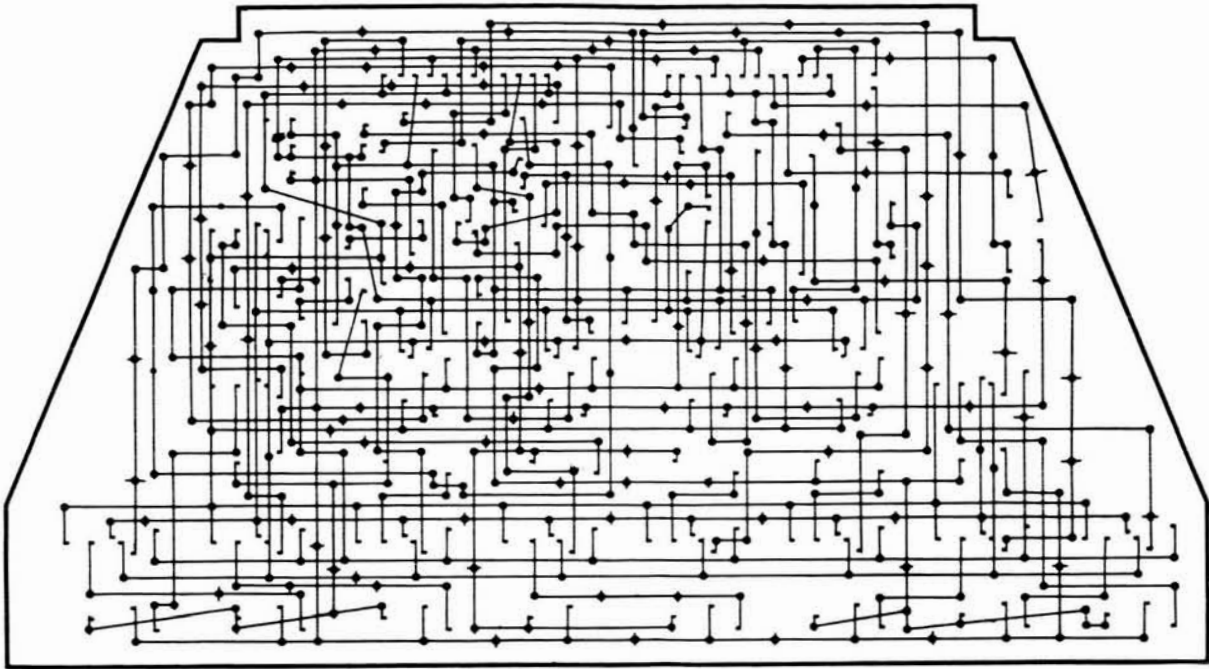


FIGURE 7.—Unpunched matrix film of a package interconnection layer.

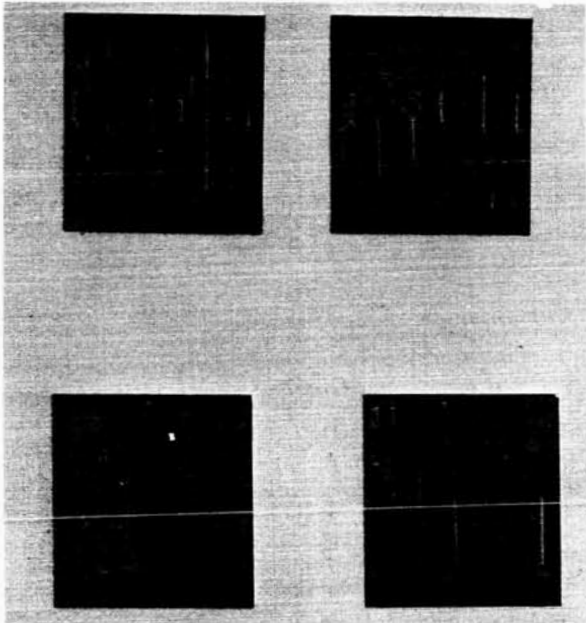


FIGURE 8.—A welded matrix.

appearance of the welds. For example, the color, deformation, and bead configuration (where applicable) have been found to have a definite relationship to the strength of the weld. Visual

criteria have been successfully employed in training personnel to produce acceptable welds.

ASSEMBLY AND INSPECTION

The assembly operation proceeds from the photographic reduction of circuit layouts to the installation of the final layers of the module. The



FIGURE 9.—The welding operation.

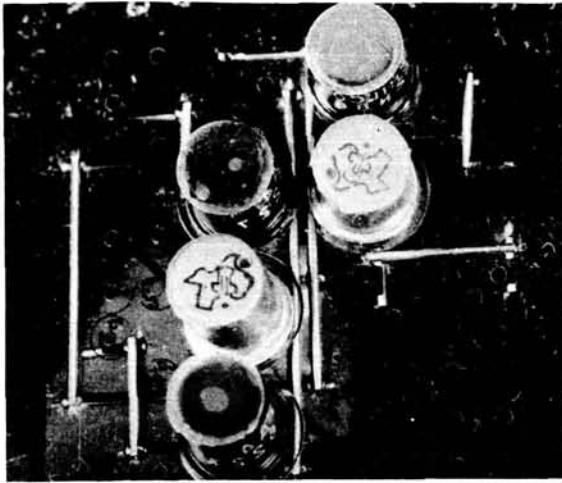


FIGURE 10.—Matrix, transistors installed.

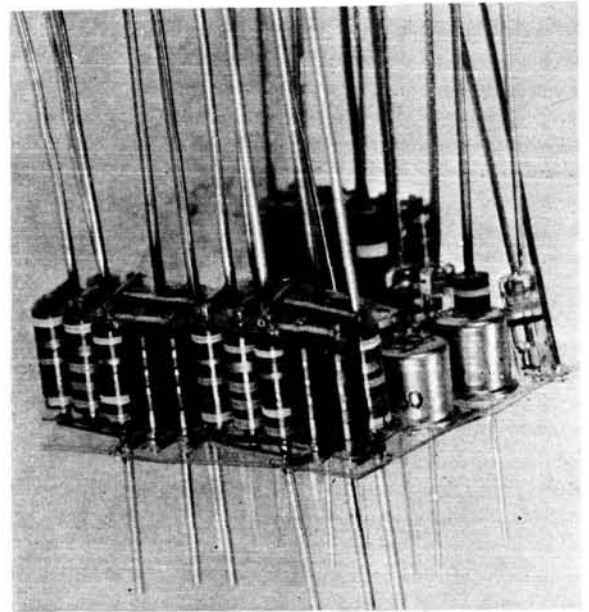


FIGURE 12.—Submodule-transistor assembly.

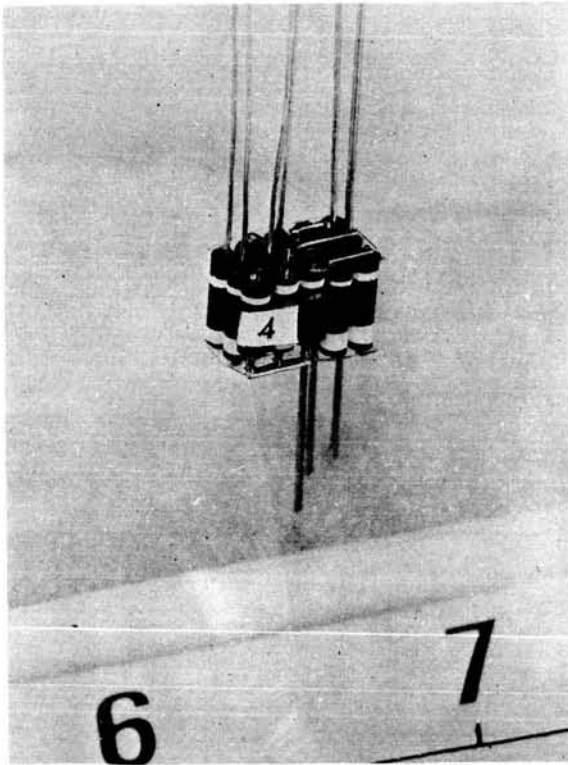


FIGURE 11.—A typical submodule.

major stages of assembly are illustrated on the following pages: Typical matrix films are shown in Figures 6 and 7. Matrix wires are welded into place (Figure 8); note the parallel arrangement of the wires on each side of the film. Figure

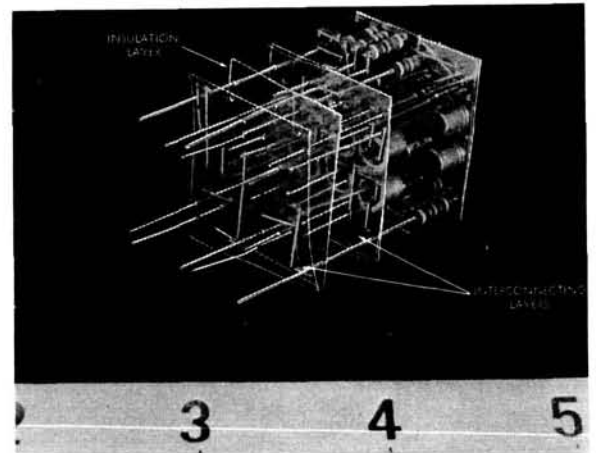


FIGURE 13.—Bottom layers.

9 shows an operator welding a component lead to the matrix. Transistors are installed in the matrix (Figure 10). Submodules similar to the one depicted in Figure 11 are fabricated and assembled to the matrix (Figure 12). When the bottom layers (Figure 13) are welded into place the module is ready for potting. Figure 14 shows a completely assembled module.

REPAIR

Prior to potting, welded modules are completely repairable. Any component that fails can be replaced in a manner that maintains the quality of the unit and assures its acceptability for flight. In most cases potted units can also be repaired. Faulty potted units must be judged on an individual basis in determining the feasibility of accomplishing necessary repairs. The quality of all reworked potted units should be thoroughly evaluated to insure that the repair processes have in no way impaired the reliability of the unit.

CONCLUDING REMARKS

The research and development phases of the welded electronic module fabrication program have been concluded in most respects. However, all work in this field will be monitored for developments which appear to advance the state of the art. An up-to-date facility and capability in welded electronic module fabrication will be maintained at the Goddard Space Flight Center.

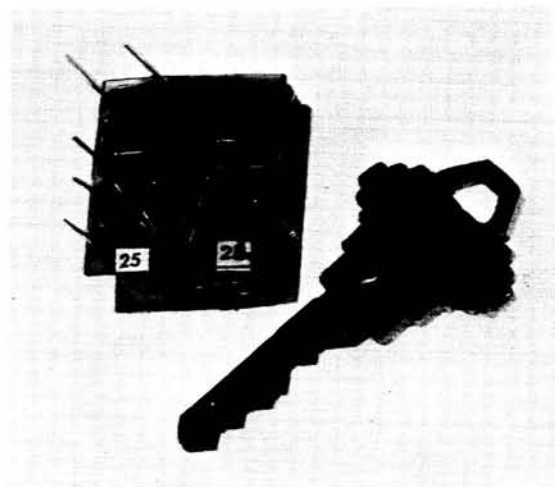


FIGURE 14.—Welded electronic module (ready for potting).

Operators and inspectors should examine welds microscopically, using established visual criteria to determine the quality of welds. One important criterion for inspection is consistency. It is important that all welds in any given package be consistent in appearance.

WIDE RANGE PHASE DETECTOR*

GEORGE B. ROBINSON

A 4-diode bridge-type phase detector of constant cosine response from 20 cps to 100 kc, was developed as a part of an experimental phase-lock filter intended for operation in that range. A 6-transistor circuit drives the diode bridge. A wide-band resistive-T adder network, rather than the usual center-tapped transformer, supplies signal and reference inputs to the bridge. The output is balanced to ground; the internal resistance of the output circuit is 600,000 ohms. Total power consumption is 135 mw at 50 v. An analysis is made which shows that the cosine response at the output terminals to be: $E_{OUT} = -1.28 E_s (\cos n\alpha) / \pi n$, for n odd; $E_{OUT} = 0$, for n -even, where E_s = RMS signal input, α = phase angle of the input signal relative to the input reference voltage, and n = integer ratio of signal frequency to reference frequency. Typical maximum cosine response is 450 mv at 1 v signal level and 5 v reference level. The measured maximum translation between any two cosine curves does not exceed 20 mv throughout the operating range of 30 cps to 100 kc. The measured cosine curve translation at 50 kc does not exceed 20 mv in the temperature range of 0° to 50°C.

INTRODUCTION

The design requirements of a low frequency phase-lock filter system recently investigated required the operation of the phase detector over a wide frequency range. The following characteristics were considered paramount in the design of the phase detector:

1. Cosine response independent of frequency throughout a 20 cps to 100 kc range.
2. Reference and signal balance constant throughout the above frequency range.
3. Signal-reference balance stable in the 0 to 50°C range.
4. Cosine response independent of temperature in the above temperature range.
5. Freedom from spurious outputs.

This paper describes a four-diode phase detector of the double-balanced, bridge-circuit configuration whose salient feature is a push-pull reference source which does not use a transformer. The circuit design is described; the static cosine response is derived in a general form which includes the response when the signal frequency

is an odd integer multiple of the reference frequency. The following characteristics are presented as measured data: cosine response in the 20 cps to 100 kc range; spurious response; reference and signal balance; temperature effects on the cosine response; effects of temperature on signal and reference balance.

CIRCUIT DESIGN

The circuit of the wide-range phase detector shown in Figure 1 is divided operationally into three sections: push-pull reference amplifier; signal amplifier; diode complex.

Reference Amplifier

The function of the reference amplifier is to supply a push-pull signal to the diode complex. The principle design criterion is equality of magnitude and constant 180-degree phase opposition of the individual outputs over the operating frequency range. This characteristic is required so that the phase detector will have a minimum or zero output when the reference is applied singly to the diode complex.

The reference amplifier consists of a degenerative voltage amplifier stage with transistor Q-1 capacitively coupled to a paraphase transistor Q-2 whose collector is again capacitively coupled

*Published as *NASA Technical Note D-2269*, April 1964.

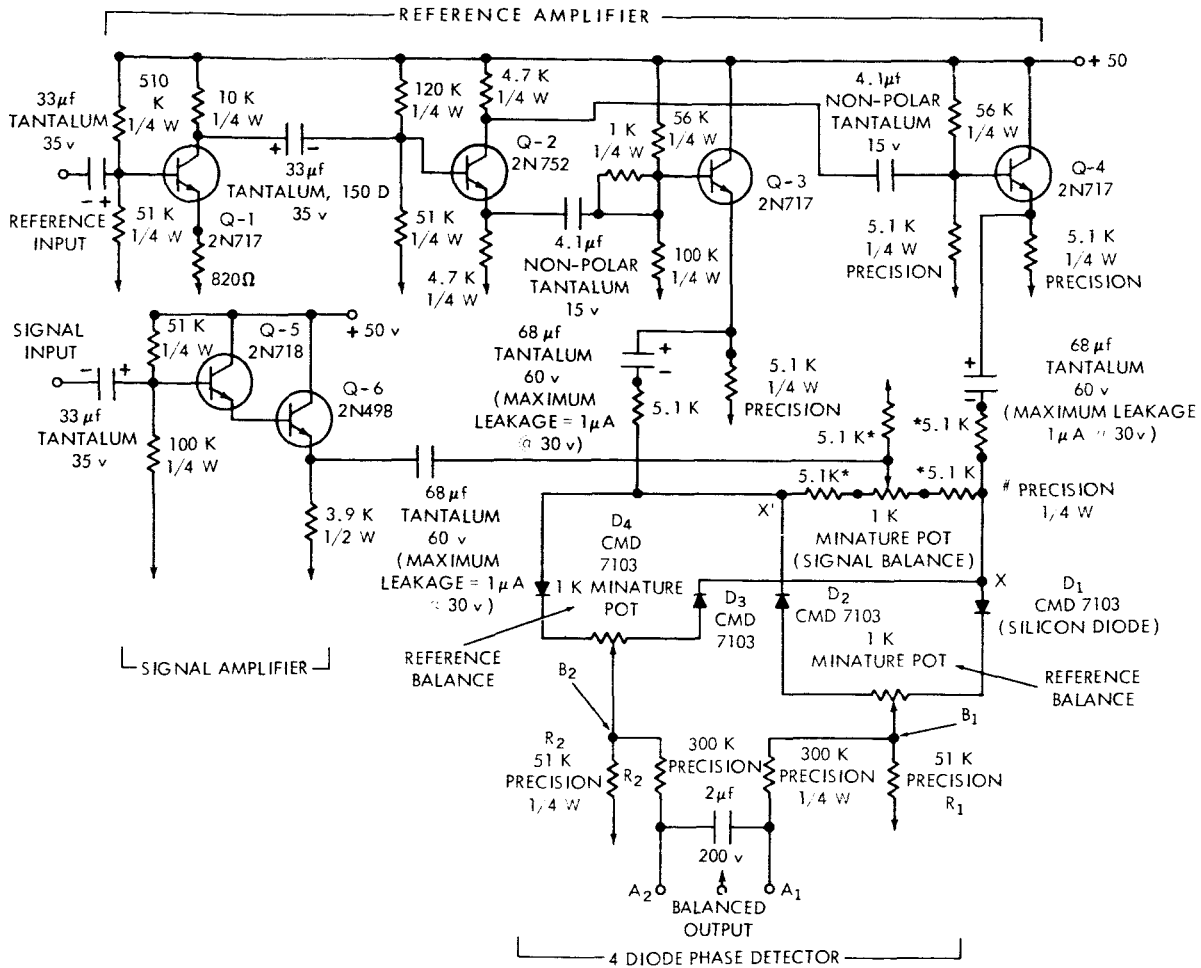


FIGURE 1.—Phase detector circuit diagram.

to the base of an emitter-follower Q-4; the emitter of Q-2 is coupled to the base of the emitter-follower Q-3. Wide frequency range is achieved by the use of tantalum capacitors and low over-all gain. The paraphase amplifier (Q-2) is the most critical stage in that the equality in magnitude of the individual push-pull output voltages and phase opposition are determined principally by this transistor. Collector impedance is considerably higher than the emitter impedance in this type of circuit; thus loading by the emitter follower stages tends to cause amplitude and phase differences. Phase error was reduced to no greater than 2 degrees by the insertion of a 1000-ohm resistor in the emitter output lead. Compensation was not complete in that a small difference in amplitude remained.

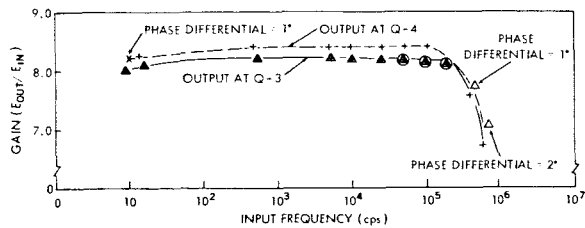


FIGURE 2.—Reference channels-gain vs frequency.

Gain and phase characteristics of the reference amplifier are shown in Figure 2. Other pertinent measured data is shown in summary form in Table 1.

Table 1

Reference-Signal Amplifier Chart.

Transistor and Function	Collector Current (Ma)	Maximum Undistorted Output Voltage (in circuit)	Magnitude of Impedance
Reference Q-1, 2N717 Voltage Amplifier	3.7	22 v, p-p	Input: 7 k-ohms at 1 kc 2 k-ohms at 100 kc
Reference Q-2, 2N752 Phase-splitter	2.7	15 p-p { collector or emitter	-
Reference Q-3, 2N717 Emitter-follower	5.5	30 p-p	Output: 53.6 ohms at 1 kc, 10 v, p-p
Reference Q-4, 2N717 Emitter-follower	5.5	30 p-p	Output: 102 ohms at 1 kc, 10 v, p-p
Signal Q-5, 2N718 } Q-6, 2N498 } Darlington emitter-follower	.4 8.2	- 12 p-p	Input: 32 k ohms at 1 kc 30 k ohms at 100 kc Output: 13.7 ohms at 1 kc 13.7 ohms at 100 kc
Total Current (Power Supply Voltage = 50 v)	26.4		

Signal Amplifier

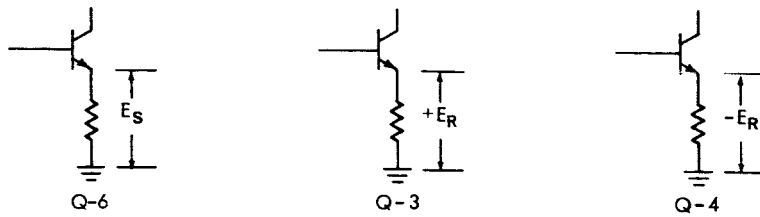
The signal amplifier consists of transistor Q-5 direct coupled to Q-6 in a Darlington configuration. This configuration increases the input impedance at the signal amplifier terminals which would otherwise be low because of the characteristics of the 2N498 power transistor employed as the output stage. The power transistor was found necessary to minimize signal suppression resulting from negative clipping when noise was present. The gain versus frequency response of this amplifier was not plotted; however, the gain is 0.975 at 600 kc and unity at 10 cps. Data pertaining to the signal amplifier is shown in Table 1.

Diode Complex

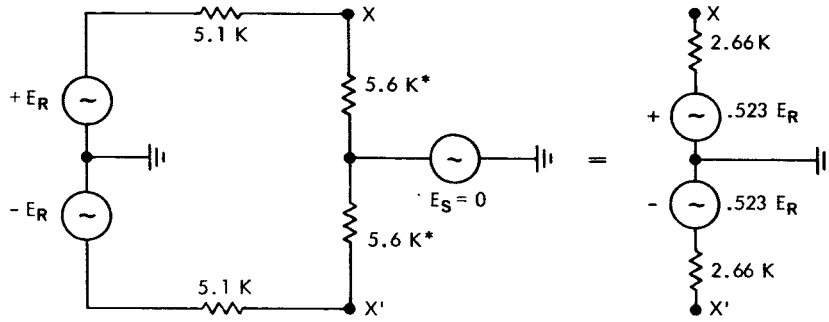
The diode portion of the phase detector is comprised of two independent sections; diodes

D-1 and D-2 with output at A₁ and diodes D-3 and D-4 with output at A₂ (Figure 1). Each section synchronously rectifies a half-wave portion of the signal. The two sections are independent in operation and produce identical outputs of opposite polarity displaced by 180 degrees. Filtering is provided by a 2 μ f capacitor across terminals A-1 and A-2. The dc output at these terminals is balanced with respect to ground. The particular diode configuration used here is known as a double balanced bridge.* The type of diode chosen for this circuit is a silicon mesa type with a maximum recovery time of 4 nanosec. The measured characteristics of this particular diode will be discussed in a later section.

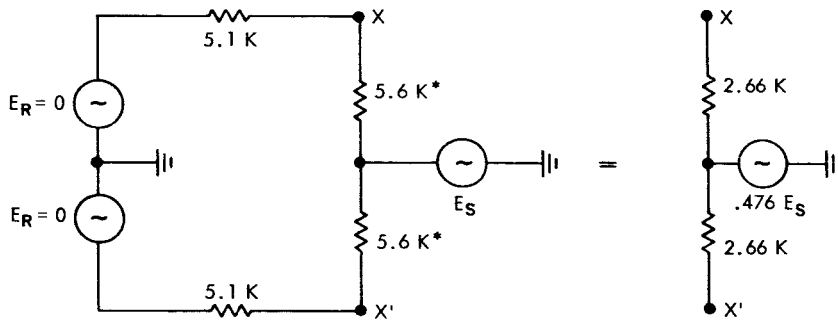
*Blickensderfer, J. A., "Diode Phase Detectors," Electronics Research Laboratory Technical Report No. 4, June 15, 1953.



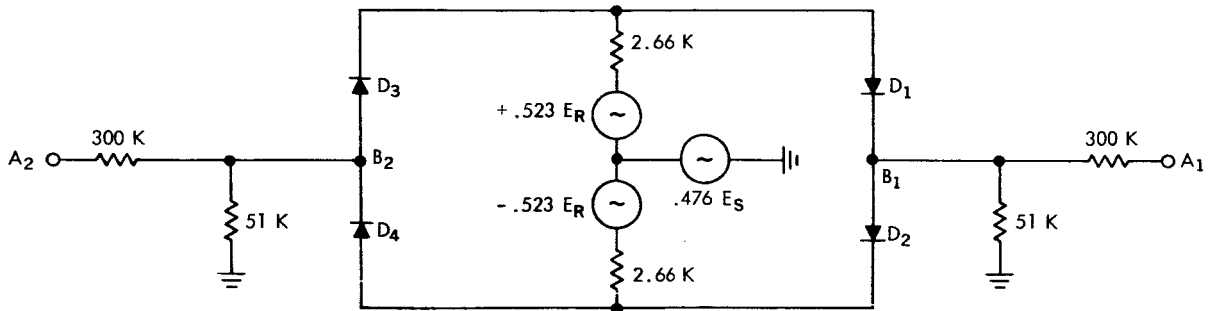
(a) Voltage and source impedance designation



(b) Derivation of reference equivalent generators & source impedance



(c) Derivation of signal equivalent generator & source impedance



(d) Equivalent circuit of phase detector

* Includes balance potentiometer resistance—arm centered

FIGURE 3.—Derivation of equivalent circuit of phase detector, $2\mu\text{f}$ filter capacitor omitted.

The diode complex also includes a resistive network which provides the signal-reference sum and difference to the diode pairs. This network consists of a T configuration (Figure 1) with 5100-ohm resistances as the arms. The signal is fed into the vertical arm of the 'T' from the output of transistor Q-6. A low-leakage 68- μ f tantalum capacitor is used for dc isolation. This capacitor is rather critical in that the leakage must be very low. The type chosen showed a maximum leakage current of 1 μ amp at 30 v. The horizontal arms of the 'T' are fed through series isolating resistors (5100 ohms) from the push-pull reference outputs of Q-3 and Q-4. The isolating resistors prevent signal voltages from appearing across the emitter-follower (Q-3 and Q-4) load resistors. Low-leakage tantalum capacitors are also employed here as dc isolating elements. The network functions in the same manner as the transformer customarily used to form the signal-reference sum and difference.

CIRCUIT ANALYSIS

Phase Detector Equivalent Circuit

The equivalent circuit of the phase detector is shown in Figure 3d. Figures 3a, 3b, and 3c show a step by step derivation of the equivalent generator and equivalent source impedances which represent the T network and emitter-follower drivers. Figure 3a defines E_r and E_s (the reference and signal outputs) and shows their location in the circuit diagram (Figure 1). In making the derivation shown in Figures 3b and 3c, it is first necessary to locate points X and X' in Figure 1. The open circuit voltage and source impedance at these points with the diodes disconnected are then determined by means of Thevenin's Theorem. For this derivation E_s is set equal to zero (Superposition Theorem). The final result is the diagram on the right (Figure 3b). In Figure 3c the process is repeated, but in this case E_r is set equal to zero. Thevenin's Theorem then gives the equivalent signal source shown on the right. Finally, in Figure 3d, Figures 3b and 3c are combined; the diode load and resistive filter network (without the filter capacitor) are also shown. For this derivation, the impedance of the reference and signal emitter followers, Q-3, Q-4, and Q-6, has been arbitrarily

taken as zero to simplify the derivation. Actual values of these impedances taken from the chart shown in Table 1 and used in a more exact derivation show that the error in Figure 3b is about 1 percent, and in Figure 3c about .1 percent.

Phase Detector Operation with Reference Input Only

The equivalent circuit is helpful in formulating a qualitative description of the phase detector operation. First, consider operation with only the reference input with ideal diodes which are matched with respect to the following characteristics: forward and reverse resistance, and shunt capacitance. The configuration of the reference sources and diodes with respect to the ground point is that of a double bridge. Each section of the bridge will conduct during alternate half-wave portions of the reference source cycle. If E_r is positive with the diode polarity shown in Figure 3, D-1 and D-2 will conduct, but because of equal forward resistance, the instantaneous potential at B-1 will remain zero. During this portion of the cycle D-3 and D-4 are nonconducting and because of postulated equal reverse resistance the output at B-2 will also be zero. A similar argument can be made for the negative portion of the reference cycle during which time D-3—D-4 will conduct and D-1 and D-2 will be nonconducting. The conclusion is the same as formerly: B-1 and B-2 remain at zero potential. The output wave forms with reference input only are shown in Figure 4. The center trace shows

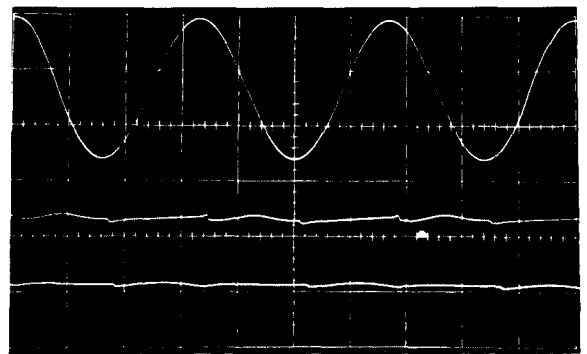


FIGURE 4.—Phase detector reference balance: $E_r=5$ v RMS; signal input=0. Upper trace—1 v RMS for amplitude comparison only; center trace—output at B₁, 1 v RMS=1.8 cm. Lower trace—output at B₂, 1 v RMS=1.8 cm.

the output wave forms of diodes D-1 and D-2 at point B-1, while the lower trace shows the output of diodes D-3 and D-4 at point B-2 (Figure 1). The upper trace is a 1 v signal for amplitude comparison. The output was taken after balancing inequalities in the diode characteristics by means of reference balance potentiometers (Figure 1). The function of these potentiometers will be discussed in greater detail.

Phase Detector Operation with Signal Input Only

Operation of the phase detector with signal input in the absence of the reference leads to the direct transmission of the signal to the terminals, B-1 and B-2, preceding the integrator. This bi-directional conduction results from the polarity of the diodes in the circuit with respect to the signal source. An inspection of the equivalent circuit shows the positive portion of the signal would be expected to appear simultaneously at B-1 and B-2 because of the conduction of diodes D-1 and D-4. Similarly, the negative portion appears at these points because of conduction through D-2 and D-3. This can be seen from



FIGURE 5.—Phase detector bi-directional conduction in the absence of reference input. $F=50$ kc. Upper trace—1 v RMS signal at Q-6; 1 v p-p=.9 cm. Center trace—output at B₁, 1 v p-p=1.8 cm. Lower trace—output at B₂, 1 v p-p=1.8 cm.

Figure 5. The wave forms at B-1 and B-2 are shown in the center and lower traces, while the upper trace shows the signal output at Q-6. These traces show that the signal is distorted and appears with diminished amplitude. This occurs because of the nature of the current-voltage characteristics of the diodes (Figure 6). The principal effect is a delay in conduction resulting

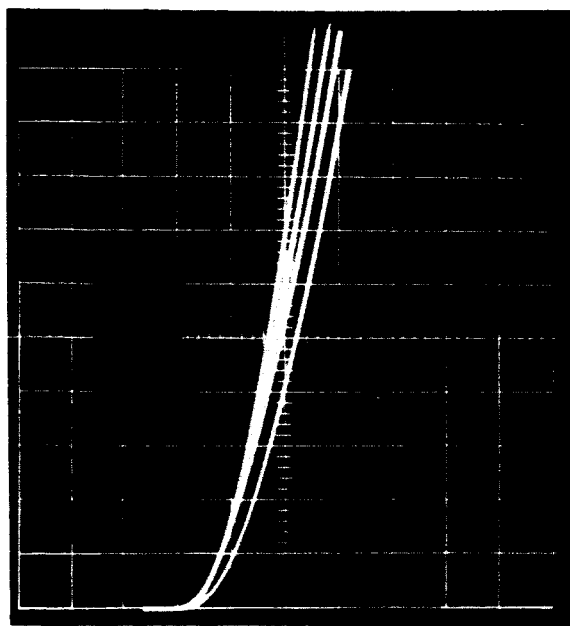


FIGURE 6.—Forward conduction characteristics of diodes, D-1, D-2, D-3, D-4 (left to right). Vertical: 1=50 ma/division; Horizontal: $E=.5$ v/division.

from the knee of the diode conduction curve. It is to be noted that though the phase detector is bi-directional in signal conduction in the absence of reference, the dc potential difference is zero for balanced diodes.

Phase Detector Transfer Function

Thus far the operation of the phase detector has been described with separate reference and signal inputs; the circuit was seen to be bi-directional with respect to signal conduction. Ideally with simultaneous inputs, the reference suppresses the bi-directional conduction and provides a transmission path for the signal over a time interval corresponding to half a cycle of the reference. This is due to the conduction of the diodes D-1 and D-2 during the positive portion of the reference. Hence the output at B-1 will be a 180 degree slice of the signal, the starting point being advanced or retarded depending on the signal-reference phase angle. The polarity of the diodes D-3 and D-4 is such that the transmission path is provided on the negative portion of the reference. Since conduction is retarded 180 degrees the signal output at B-2 will be the mirror image of the output at B-1.

Ideally the balance of the system with respect to the reference is maintained so that the reference does not contribute to the output at B-1 and B-2. The output wave forms for signal-reference phase angle of 0, 45, and 90 degrees are shown in Figures 7, 8, and 9. The central trace shows the output at point B-1 in Figure 3 (diodes D-1 and D-2) while the lower trace shows the output at B-2. The upper trace shows the signal-voltage output at transistor Q-6. Deflection sensitivity of the signal presentation (upper trace) is one half that of the lower traces so that the relative amplitude between the signal and the rectified portions of the signal is maintained. The lower sensitivity was required to maintain proportion because the signal output at transistor Q-6 is attenuated by approximately one-half by the T network (Figure 3c). Because the half-wave signal outputs are

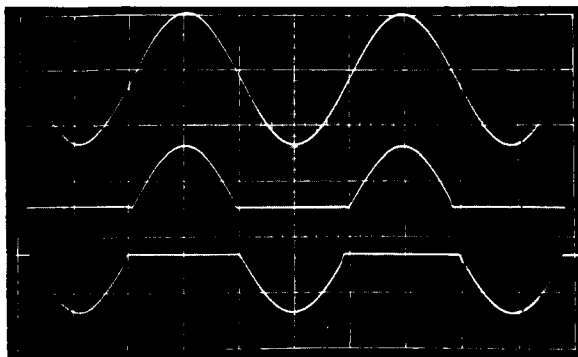


FIGURE 7.—Phase detector output for $\theta=0^\circ$, $E_r=5$ v, $f=1$ kc. Upper trace—1 v RMS signal at Q-6, 1 v = .85 cm. Center trace—output at B₁; 1 v = 1.7 cm. Lower trace output at B₂, 1 v = 1.7 cm.

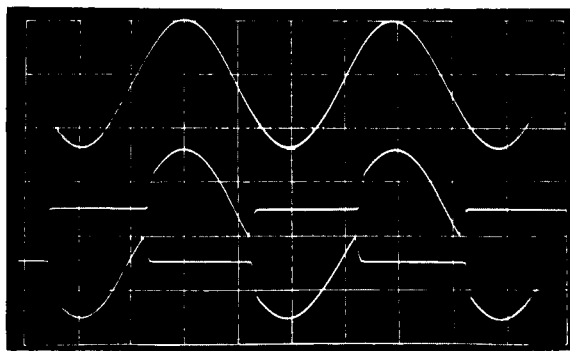


FIGURE 8.—Phase detector output for $\theta=45^\circ$, $E_r=5$ v, $f=1$ kc. Upper trace—1 v RMS signal at Q-6, 1 v = .85 cm. Center trace—output at B₁; 1 v = 1.7 cm. Lower trace output at B₂, 1 v = 1.7 cm.

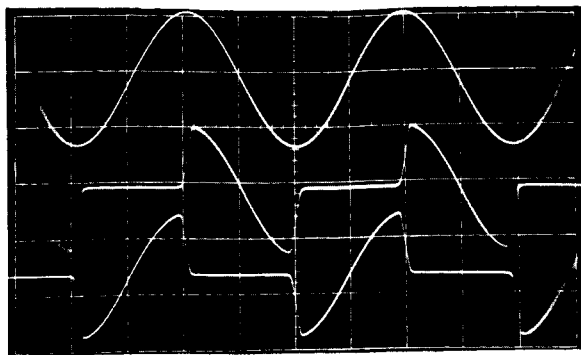


FIGURE 9.—Phase detector output for $\theta=90^\circ$, $E_r=5$ v, $f=1$ kc. Upper trace—1 v RMS signal at Q-6, 1 v = .85 cm. Center trace—output at B₁, 1 v = 1.7 cm. Lower trace output at B₂, 1 v = 1.7 cm.

of opposite polarity relative to ground, the total potential difference available at the output terminals A-1 and A-2 is twice the individual potentials. Integration of the individual outputs is obtained by 300,000-ohm resistors feeding a 2- μ f capacitor placed across A-1 and A-2.

Quantitatively the dc output at the A-1 terminal is the integral of the signal voltage, with the lower limit corresponding to the incidence of the on-time (diodes D-1 and D-2 conduct) and the upper limit corresponding to the incidence of the off-time (diodes D-1 and D-2 cease conducting). The resulting definite integral is then divided by a time corresponding to a complete reference cycle to obtain an average. In describing the dc output, quantitatively the following assumptions are made: (1) Diodes D-1 and D-2 conduct when $E_r > 0$ with a forward resistance of zero independently of the signal voltage; similarly diodes D-3 and D-4 conduct for $E_r < 0$ with zero forward resistance. (2) E_r does not appear in the output; i.e., the system is perfectly balanced with respect to the reference and balance is maintained when the signal is present.

The integral giving the dc output at A₁ is determined as follows: The instantaneous voltage output of the equivalent signal source (Figure 3) at angular frequency ω is $\sqrt{2}$ (.476) $E_s \sin n(\omega t + \theta)$, where θ is the phase angle of the signal measured relative to the reference output E_r , E_s = RMS value at the output of transistor Q-6. The factor n is the ratio of the signal frequency to the reference frequency and is taken to represent an integer. The dc output at A₁ where D-1

and D-2 conduct during the positive half cycle of the reference, i.e. from 0 to π radians, is given by

$$E_{A_1} = \frac{K(\sqrt{2})(.476)E_s}{\frac{1}{f}} \int_{t_1=0}^{t_2=\pi/\omega} \sin n(\omega t + \theta) dt, \quad (1)$$

where $f \equiv \omega/2\pi$. The factor K in the above equation is the fraction of the signal voltage transferred to the 51,000-ohm load resistance

through the source resistance of 2,660-ohms (Figure 3). This fraction is $51/(51+2.66)$ or 0.95. With this value of K Equation 1 integrates

to (Appendix A)

$$E_{A_1} = \frac{(.95)(\sqrt{2})(.476)E_s}{\pi n} \cos n\theta, \text{ for } n \text{ odd}, \quad (2)$$

$$E_{A_1} = 0, \text{ for } n \text{ even}.$$

To determine the dc output appearing at terminal A-2, note that the polarity of diodes D-3 and D-4

is such that conduction occurs during the negative half cycle of the reference, i.e. from π to 2π

radians:

$$E_{A_2} = \frac{(.95)\sqrt{2}(.476)E_s}{\frac{1}{f}} \int_{t_1=\pi/\omega}^{t_2=2\pi/\omega} \sin n(\omega t + \theta) dt. \quad (3)$$

Integrating, we have (Appendix I)

$$E_{A_2} = \frac{-(.95)(\sqrt{2})(.476)E_s}{\pi n} \cos n\theta, \text{ for } n \text{ odd},$$

$$E_{A_2} = 0, \text{ for } n \text{ even}. \quad (4)$$

The total output voltage across the A-1—A-2 terminals is the potential difference $E_{A_1} - E_{A_2}$ (from Equations 2 and 4):

$$E_T = \frac{(.95)\sqrt{2}(.952)E_s}{\pi n} \cos n\theta, \text{ for } n \text{ odd},$$

$$E_T = 0, \text{ for } n \text{ even}. \quad (5)$$

The response predicted mathematically by Equation 5 at odd multiples of n is evident intuitively: The switching interval occurs over a half period of the reference. This interval will contain n alternate positive and negative half cycles of the signal. Each alternate pair cancels in the averaging process, thus a dc output will result only when the switching interval contains an odd number of half cycles of the signal. The

attenuation factor $1/n$ appears in Equation 5 because the remaining odd half cycle is averaged over $2n$ half periods of the signal, whereas when n is unity the averaging occurs over 2 half cycles of the signal frequency since this is a half-wave rectification process.

In deriving Equation 5, the principal assumption made was that the diode pair conduction was determined solely by the reference source and independently of the signal. This assumption of independence is to a large measure satisfied if the reference voltage level is of sufficient magnitude relative to the signal level. However, at low reference levels the above assumption is no longer valid and two effects occur which lead to a reduction in dc output at the A-1—A-2 terminals: (1) The conduction angle becomes less than 180 degrees, (2) Bi-directional conduction occurs. The first effect can be seen in Figure 10 where the reference level has been reduced from 5 to 0.5 v (signal level is 1 v, as in Figures 7, 8, and 9). The upper trace at $f=1$ kc is the signal output at transistor Q-6; The inner trace is the output of diodes D-1—D-2 at point B-1 (Figure 1) with

the base line approximately coincident with the signal base line. The lower trace is the output of diodes D-2—D-3 at point B-2. Deflection sensitivity of the signal presentation is one half that of the lower traces so as to maintain amplitude proportion. It is evident that the reduction in the conduction angle is at least 30 degrees, although bi-directional conduction is not pronounced.

The reduction in conduction angle reduces the interval of integration in Equations 1 and 3 with a corresponding reduction in the value of E_{A_1} and E_{A_2} .

Pronounced bi-directional conduction as well as further reduction in the conduction angle is evident in Figure 11, which is identical with

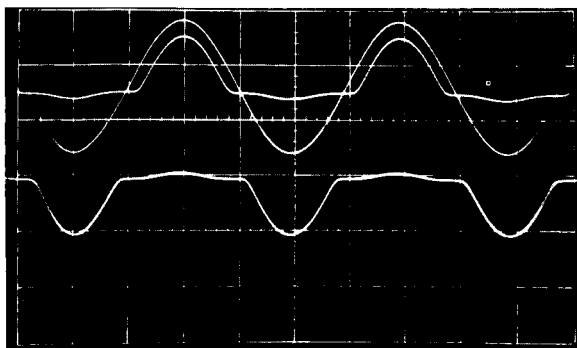


FIGURE 10.—Phase detector output wave forms with reduced reference level: $E_r=0.5$ v RMS, $E_s=1$ v RMS, $f=1$ kc, $\theta=0$: Upper trace—1 v RMS signal at Q-6, 1 v p-p=.85 cm. Inner trace output at B₁, 1 v p-p=1.7 cm. Lower trace—output at B₂, 1 v p-p=1.7 cm.

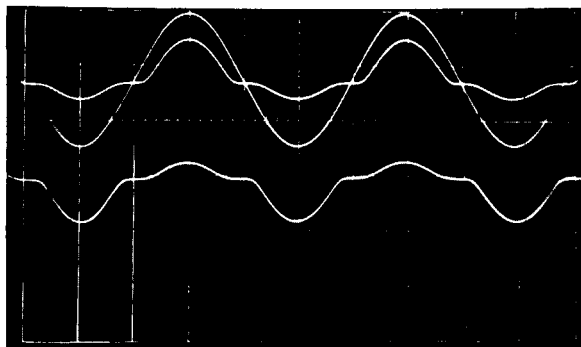


FIGURE 11.—Phase detector output wave forms with reduced reference level: $E_r=.25$ v RMS, $E_s=1$ v RMS, $f=1$ kc, $\theta=0$. Upper trace output—1 v RMS signal at Q-6, 1 v p-p=.85 cm. Inner trace output at B₁, 1 v p-p=1.7 cm. Lower trace output at B₂, 1 v p-p=1.7 cm.

Figure 10 except that the reference has been further reduced to a level of 0.25 v. Bi-directional conduction is evident in the lower trace (the output of diodes D-3 and D-4) below the interval of positive excursion of the signal. Because the phase angle is zero, the reference is in phase with the signal and is therefore positive in this first interval; the polarity of D-3 and D-4 is such that the reference would ordinarily bias them off. Similarly, in the following interval where the signal and reference are negative, conduction is seen to take place in diodes D-1 and D-2, which with sufficient reference would ordinarily be biased off. Comparison of these traces with those shown in Figure 7 where the reference level is 5 v will make the above discussion clearer. A reduction in phase detector output results from bi-directional conduction in that rectification of the signal decreases since conduction takes place during both positive and negative excursions. In the limiting case where the reference is zero (as in Figure 5), rectification ceases entirely.

Quantitatively, both of these effects involve the forward resistance characteristics of the diodes in the region of greatest curvature (Figure 6). This resistance* is given by

$$r_d = \frac{.052}{I_f} + R_s$$

where I_f =forward current, R_s =series resistance in ohms at $T=25^\circ\text{C}$, r_d =forward resistance. Measured values of one diode show that this relationship is followed rather closely for currents greater than 10 μ -amp.

However, the inverse relationship between resistance and current exhibited by the solid state diodes precludes analysis by linear means; consequently no attempt was made to derive the phase detector transfer function at low reference levels. This aspect of phase detector operation has been considered by Blickensderfer† but his analysis appears to be adequate only for thermionic diodes. His analysis shows that at low reference-to-signal levels the output vs phase angle becomes more linear. However, the data

*"General Electric Transistor Manual," 6th Edition, 1962, pp. 281-283.

†Blickensderfer, J. A., "Diode Phase Detectors," Electronics Research Laboratory Technical Report No. 4, June 15, 1953.

shown later in Figure 16 appear to be in conflict with Blickensderfer's analysis in that a cosine response, rather than a linear response was observed at low reference-to-signal levels.

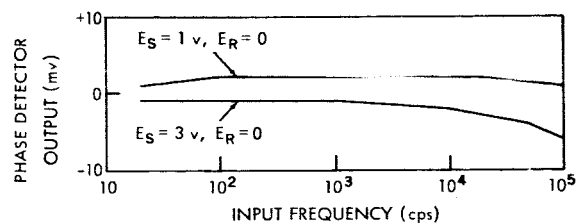
MEASURED CHARACTERISTICS

Reference and Signal Balance

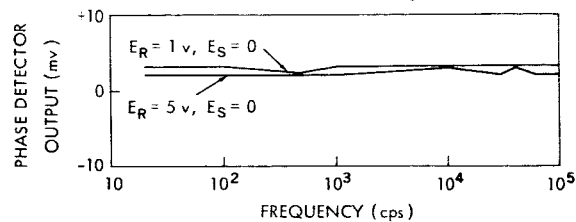
The development of Equation 5 postulated ideal diodes and a balanced push-pull reference source. Operationally these ideal conditions lead to zero output with reference input only, and zero dc output with signal input only. Individual signal and reference balances are desirable in actual operation of the phase detector as lack of balance results in dc offset of the zero point ($\theta=90^\circ$) of the phase detector cosine response as well as lack of symmetry in the positive and negative regions.

The principal cause of deviation from ideal balance probably arises from differences in individual diode characteristics. Although the diodes employed in this circuit were purchased as matched pairs there remains some disparity, particularly in the forward resistance characteristics of D-4. This is evident from the voltage-current traces shown in Figure 6. A secondary cause of lack of balance is the small inequality in the amplitudes of the reference push-pull outputs at transistors Q-3 and Q-4 which is evident from the plot shown in Figure 2. To compensate for these deficiencies, at least partially, three balance potentiometers were included in the circuit. Balance with signal input only was obtained by a 1,000-ohm potentiometer placed in the center of the T network and is designated as "signal balance potentiometer" in the circuit diagram, Figure 1. For balance with reference input a 1,000-ohm potentiometer was inserted in series with diodes D-1 and D-2; An identical unit was placed in series with diodes D-3 and D-4. These two potentiometers are designated as "reference balance" in the circuit diagram.

The designation of individual potentiometers as "signal" or "reference" is optimistic in that their effects on the signal or reference balance are to some extent mutually interacting. In practice, however, these adjustments were not found difficult to achieve, although the balances so obtained were functions of voltage level, temperature, and to a lesser extent, frequency.



(a) Phase detector balanced initially at $E_s = 1$ v, $f = 20$ cps



(b) Phase detector balanced initially at $E_r = 5$ v, $f = 20$ cps

FIGURE 12.—Effect of change in signal or reference voltage on phase detector balance.

Figure 12a illustrates the dependence of phase detector balance versus frequency on signal level. The phase detector output across A-1 and A-2 was minimized by adjusting the signal potentiometer with a 20 cps, 1 v signal input and zero reference input. The output was then plotted over a frequency range of 20 cps to 100 kc. To show the effect of a change in signal level, the signal was increased to 3 v and the output versus frequency again plotted. A similar procedure was followed in Figure 12b to show the effect of a change in reference level. In this case, initial balance was made at a 20 cps reference input of 5 v and zero signal input by adjusting the reference potentiometer.

The effects of ambient temperature variation on signal balance versus frequency are shown in Figure 13a. For this data, initial signal balance was made at $T=25^\circ\text{C}$ with a 20 cps, 1 v signal and zero reference input. The phase detector output versus frequency was then plotted for temperatures of 0°C , 25°C and 50°C respectively. Reference balance versus frequency is shown for these temperatures in Figure 13b. Initial balance was established at 5 v reference input and zero signal input with $T=25^\circ\text{C}$.

One final consideration with respect to balance has yet to be considered; that is, possible leakage through the $68 \mu\text{f}$ capacitors employed to couple the signal and reference to the diode circuitry. It might be expected that the capacitor leakage

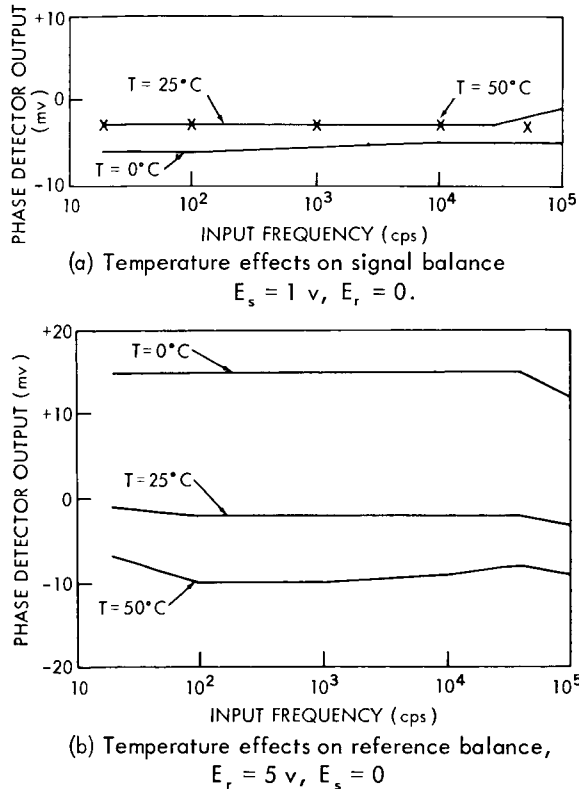


FIGURE 13.—Temperature effects on signal or reference balance.

would produce an output at A-1 or A-2 in the absence of signal or reference inputs, particularly at higher temperatures. However at the highest temperature employed no output attributable to leakage was noted.

Phase Detector Response—Measured and Predicted Values

The solid line plotted in Figure 14 was computed from Equation 5 for a signal voltage of 1 v RMS, $n=1$. The data points indicated by triangles are the measured dc output of the phase detector versus phase angle θ at a frequency of 1 kc for $E_r=5 \text{ v}$ RMS. Data points indicated by crosses were obtained with the additional circuitry shown in Figure 15. With the use of this circuitry the integrating network was isolated so that the potential normally developed across the integrating capacitor was prevented from affecting conduction in the diodes. Ordinarily a small dc potential does appear at the juncture of

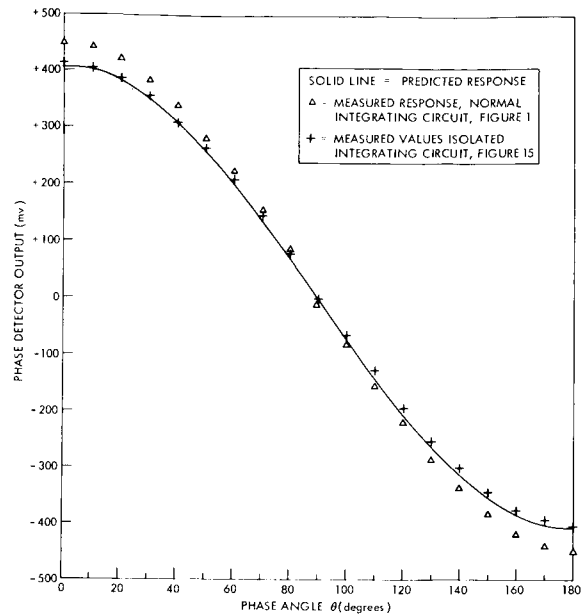


FIGURE 14.—Measured and predicted response. $E_r = 5 \text{ v}$ RMS, $E_s = 1 \text{ v}$ RMS, $f = 1 \text{ kc}$.

each pair of diodes because of transfer by means of the 300,000-ohm resistors to the upper terminals of the 51,000-ohm resistors at the diode junctions.

Cosine Response at Lower Reference Inputs

The dc outputs at the A-1—A-2 terminals versus phase angle with the reference input decreased from 5 v (Figure 14) to 1 v, 0.5 v, and 0.25 v, respectively, are shown in Figure 16. Signal input is constant at $E_s=1 \text{ v}$ RMS. Data points in Curve 1 are plotted as triangles while the values predicted by Equations 5, $E_r=1 \text{ v}$, $n=1$ are shown as the solid line.

The second set, Curves 2 and 3, illustrate the response of the phase detector where bidirectional conduction and pronounced reduction in conduction angle occur because of insufficient reference voltage. The output wave forms at $\theta=0$ for reference voltages 0.5 and 0.25 v, have been shown in Figures 10 and 11. For these curves the solid lines of Figure 16 are best fit cosine functions rather than predicted values.

Wide Band Cosine Response

Wide band cosine response of the phase detector is illustrated by measurements presented in

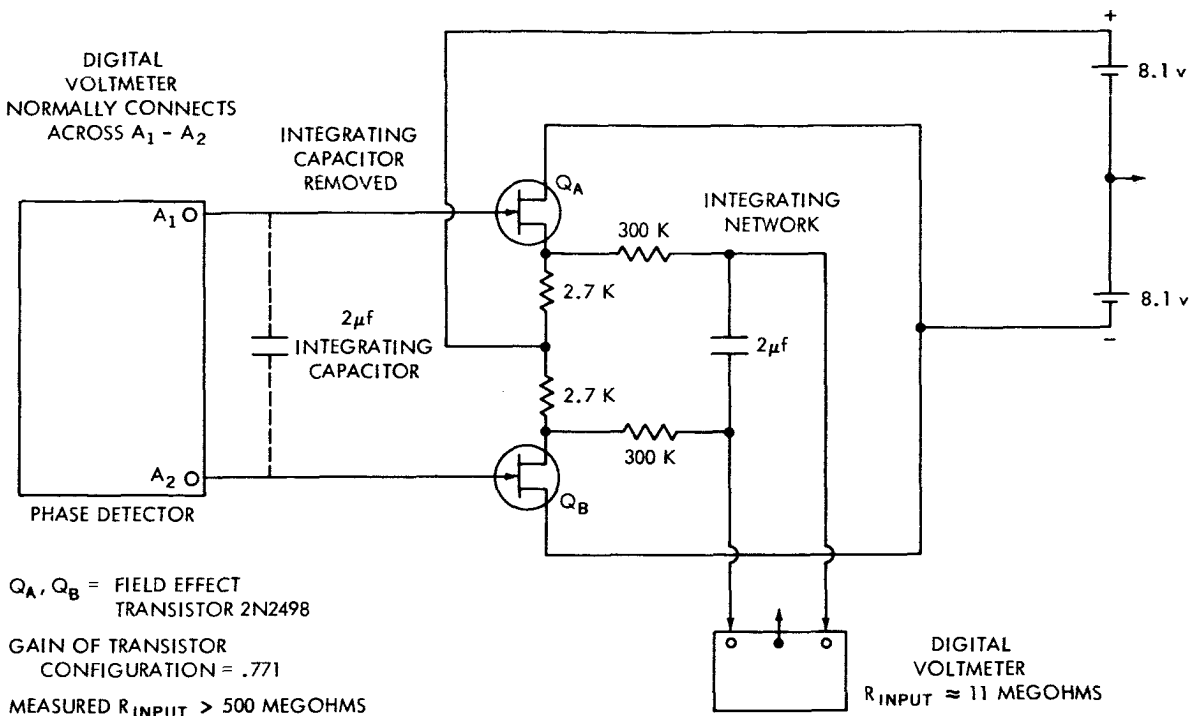


FIGURE 15.—Isolating circuit employed taking data shown in Figure 14.

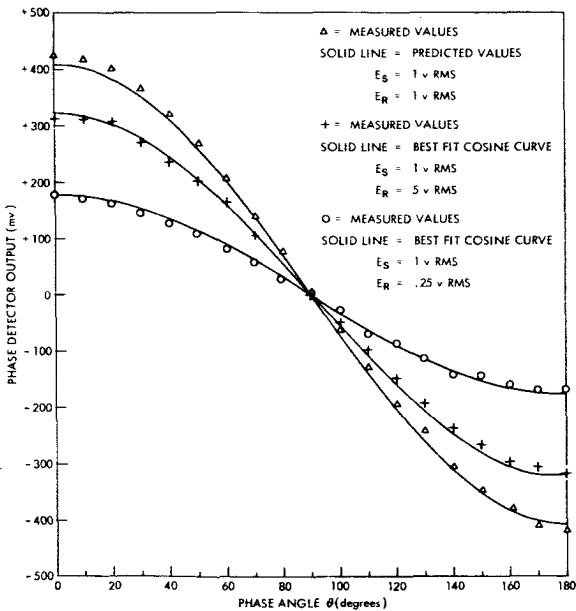


FIGURE 16.—Phase detector output versus phase angle for E_r = 1 v, .5 v, .25 v RMS; E_s = 1 v RMS; f = 1 kc.

Figures 17 and 18. Previous to taking this data individual signal and reference balances were made at 30 cps at levels of 1 and 5 v respectively.

With this fixed balance, cosine curve responses were measured at the following frequencies: 30, 400 cps, 10, 25, 50, and 100 kc with E_r = 5 v RMS, E_s = 1 v RMS. This presentation differs in one aspect from the previous cosine measurements—the phase angle plotted as the abscissa is the *input* phase angle rather than previously defined angle θ . Specifically, the phase angle of the signal at the base of Q-5 was measured relative to the reference input at the base of Q-1. This angle differs by 180° from θ as a result of the phase inversion which occurs in the voltage amplifier stage (Figure 1). Consequently, as a result of the identity $\cos(\theta \pm 180) \equiv -\cos \theta$, E_T in Equation 5 becomes negative for positive θ . Thus, this set of cosine curves is seen to be inverted in polarity when compared with the previous sets. Equation 5 is plotted with this alteration in sign as a solid line in Figures 17 and 18.

Effects of Ambient Temperature Variation on Cosine Response

Figure 19 shows the cosine response with the input phase angle as the abscissa, T = 0, 25, and 50°C, and f = 50 kc. The initial balance was

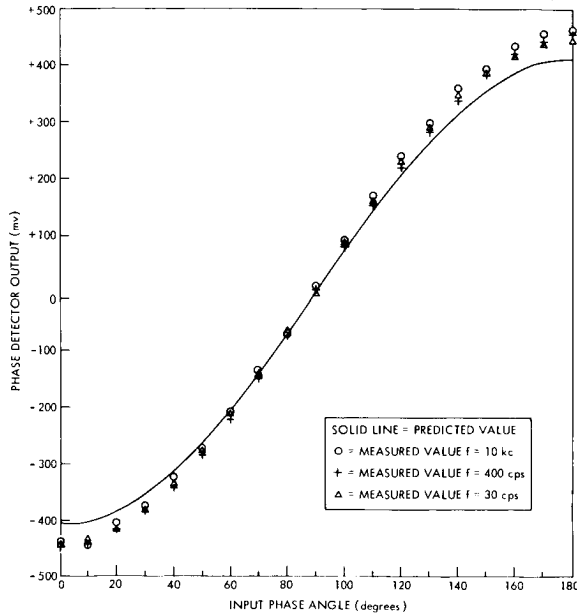


FIGURE 17.—Phase detector output versus input phase angle for $f=30, 400$ cps and 10 kc; input to signal amplifier = 1 v; input to reference amplifier = $.61$ v ($E_r=5$ v RMS).

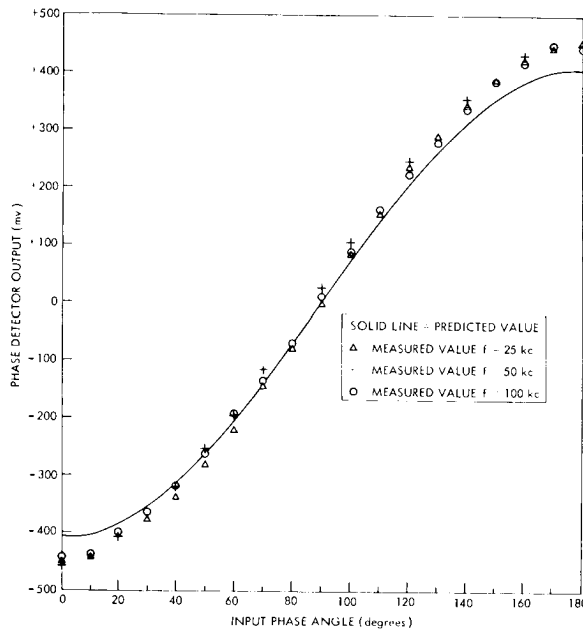


FIGURE 18.—Phase detector output versus input phase angle for $f=25, 50,$ and 100 kc; input to signal amplifier = 1 v; input to reference amplifier = $.61$ v RMS ($E_r=5$ v RMS).

established at $f=30$ cps, $E_r=5$ v RMS, $E_a=1$ v RMS. The lower solid line in Figure 20 is the measured response at $T=25^\circ\text{C}$; Data taken at

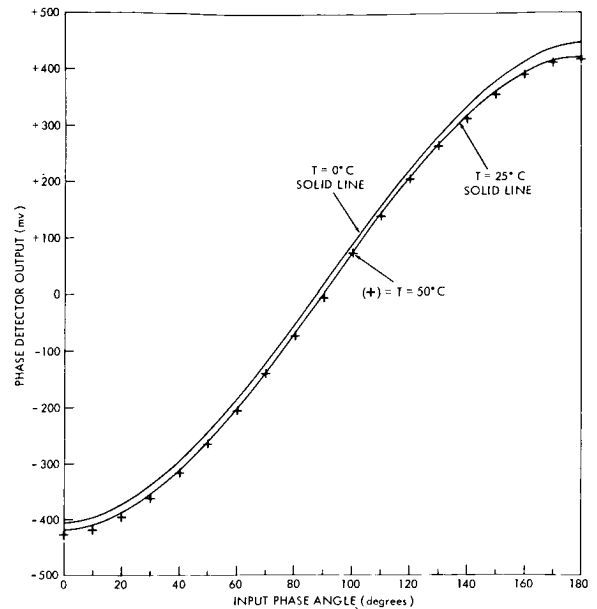


FIGURE 19.—Measured phase detector output versus phase angle for $T=0, 25,$ and 50°C ; $E_r=5$ v RMS; $E_a=1$ v RMS; $F=50$ kc.

50°C are plotted as crosses; and the upper line is the response at $T=0^\circ\text{C}$.

Cosine Response when Signal Frequency is an Odd Multiple of the Reference Frequency

Equation 5 predicts that although the maximum phase-detector cosine response occurs for equal signal and reference frequencies ($n=1$), cosine responses attenuated by the factor $1/n$ will also occur if the signal frequency is an odd multiple of the reference frequency. Experimentally, the cosine response at odd multiples of n is difficult to determine. The type of system required to perform this measurement is shown in Figure 20a. In this system a single source feeds the reference input and simultaneously a calibrated phase shifter. The output of the phase shifter drives an 'N' times frequency multiplier which is fed into the signal input of the phase detector. Coherence between signal and reference inputs is thus assured. Unfortunately most commonly available multipliers produce distorted outputs which would alter the cosine response of the phase detector. The method shown in Figure 20b is much simpler but does not permit actual cosine response measurement. The system does, however, permit the measurement of the $1/n$

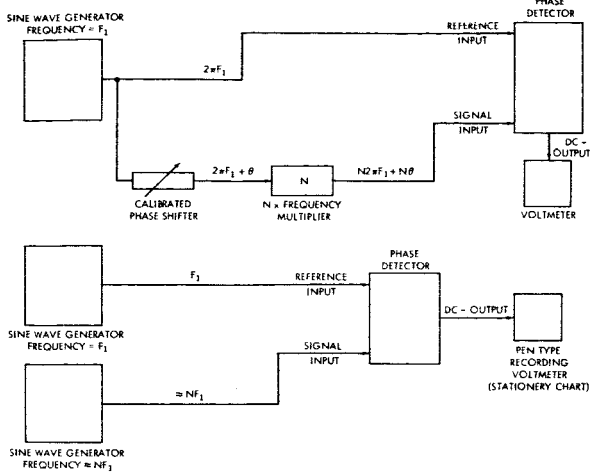


FIGURE 20.—Ideal and approximate methods for determining phase detector response at odd-frequency multiples.

factor predicted in Equation 5. The more simple method uses separate generators as reference and signal sources. The reference source is operated at a fixed frequency F_1 ; the signal source is adjusted to be *nearly* an odd multiple of F_1 . The lack of periodicity between generator frequencies will then cause a cyclic variation in phase angle with a resultant fluctuation in phase detector output. (This fluctuation is, of course, the cosine response, but the phase angle is underdetermined.) With this method, a pen recorder with stationary chart was used to record the output for different

values of n . The length of the traces so obtained are proportional to the maximum to minimum excursion of each particular cosine response. The data obtained from the trace length measurement are shown in the last column of Table 2; the predicted values are shown in the center column.

The two generator system shown in Figure 20 was also used to determine whether cosine responses occurred at fractional values of n (a case not considered in the analysis) as well as non-integer values of n . No responses were observed at these values.

It might be remarked that the cosine response at odd multiples of n , though not measured directly by the two generator method has been observed in a phaselock system which included this type of phase detector.

EXPERIMENTAL PROCEDURE

The two basic quantities which were measured were the A-1—A-2 voltage output and the signal-reference phase angle. The voltage output was measured with a digital voltmeter with an input resistance of 11-megohms; readings made with this instrument were corrected for loading. A value of 600,000 ohms at the A-1—A-2 terminals was employed in making this correction. Phase angle determination was made with a standard commercial instrument which had a coincident slicer preceded by cathode-coupled

Table 2

Predicted and Measured Values of Phase Detector Maximum Output Ratios at Odd Frequency Multiples.

Signal-reference frequency ratio N	Predicted maximum output for $N = 1$ divided by predicted maximum output for indicated N	Measured maximum output for $N = 1$ divided by measured maximum output for indicated N^*
3	3	3.08
5	5	5.2
7	7	7.3
9	9	11.1

*Reference frequency = 1 kc, $E_R = 5 \text{ v RMS}$, $E_S = 1 \text{ v RMS}$.

limiters. Claimed accuracy for this instrument was ± 1 percent. Phase angle variation over a 180° range required three auxiliary phase shifters each covering a different band of frequency. At the upper frequencies a commercial delay line was found adequate. At frequencies below 30 kc and above 1 kc a capacitive rotating field type phase shifter was used. This device has two sets of stator plates mounted orthogonally and excited by quadrature voltages. A third set of plates is mounted so as to form a rotor which is capacitatively coupled to the stator. A two-phase rotating field is produced by the stator plates and therefore the output of the rotor shows a phase variation proportional to its angular position. For frequencies below 1 kc a device similar in principle but magnetic in operation produced variable phase shift from a wound rotor.

CONCLUSION

The phase detector described here has proved to be very reliable in operation over a six-month period. However, the system reliability could probably be increased by the substitution of a type 2N719A transistor in all stages with the exception of Q-6; the 2N719A has an 80 v collector-to-emitter rating which should decrease the likelihood of transistor failure. The major

design goal of constant cosine response in a 20 cps to 100kc frequency range has been successfully achieved. Reference balance stability versus temperature, though adequate, could probably be improved by the selection of more carefully matched diodes. The use of a transistor push-pull wide-band amplifier as the reference source, rather than a transformer, appears to be a good choice, although there is a sacrifice in total power consumption.

The agreement between measured and predicted cosine responses appears to be very nearly perfect when the phase-detector integrating circuit is electrically isolated from the diodes. Where isolation is not employed the response is cosinusoidal as predicted, but the amplitude agreement shows a 10 percent error.

Extension of the present analysis to describe adequately the operation at low reference-to-signal-voltage ratios appears to be difficult with other than graphical means.

The predicted attenuation factor $1/n$ for odd-multiple signal-to-reference-frequency ratios shows good agreement with measured values up to and including 7, but shows an unexplained error at $n=9$.

Measurements show the phase detector to be free from spurious responses throughout the operating frequency range.

APPENDIX I

MATHEMATICAL ANALYSIS OF THE PHASE DETECTOR

Equation 1, page 9, with $K = .95$, is

$$E_{A_1} = \frac{.95(\sqrt{2})(.476)E_s}{\frac{1}{f}} \int_{t_1=0}^{t_2=\pi/\omega} \sin n(\omega t + \theta) dt \quad (\text{I-1})$$

For convenience we let $A = (.95) \sqrt{2}(.476)$. Equation I-1, in integrable form is,

$$E_{A_1} = \frac{AE_s}{n\omega \frac{1}{f}} \int_{t_1=0}^{t_2=\pi/\omega} \sin n(\omega t + \theta) (n\omega) dt, \quad (\text{I-2})$$

which yields

$$E_{A_1} = -\frac{AE_s}{2\pi n} \cos(n\omega t + n\theta) \Big|_{t_1=0}^{t_2=\pi/\omega} \quad (\text{I-3})$$

$$= -\frac{AE_s}{2\pi n} [\cos(n\pi + n\theta) - \cos n\theta] \quad (\text{I-4})$$

But $\cos(\alpha \pm \beta) \equiv \cos \alpha \cos \beta \mp \sin \alpha \sin \beta$; with this identity, Equation I-4 is

$$E_{A_1} = -\frac{AE_s}{2\pi n} [\cos n\pi \cos n\theta - \sin n\pi \sin n\theta - \cos n\theta] \quad (\text{I-5})$$

Let n be an *even* integer; then,

$$\cos n\pi \equiv +1,$$

$$\sin n\pi \equiv 0.$$

With these relationships Equation I-5 gives

$$E_{A_1} = -\frac{AE_s}{2\pi n} [\cos n\theta - \cos n\theta] \equiv 0, \text{ for } n \text{ even.} \quad (\text{I-6})$$

Let n be an *odd* integer, then

$$\cos n\pi \equiv -1,$$

$$\sin n\pi \equiv 0.$$

Consequently Equation I-5 becomes

$$E_{A_1} = -\frac{AE_s}{2\pi n} [-\cos n\theta - \cos n\theta] = \frac{AE_s}{\pi n} \cos n\theta, \text{ for } n \text{ odd.} \quad (\text{I-7})$$

The integral giving E_{A_2} is obtained by a half-period shift of the limits of integration of Equation I-2:

$$E_{A_2} = \frac{AE_s}{2\pi n} \int_{t_1 = \pi/\omega}^{t_2 = 2\pi/\omega} \sin n(\omega t + \theta) n \omega dt. \quad (\text{I-8})$$

The value of the above integral is

$$E_{A_2} = -\frac{AE_s}{2\pi n} [\cos(2\pi n + n\theta) - \cos(n\pi + n\theta)]. \quad (\text{I-9})$$

With the previous cosine double angle identity this can be written as

$$E_{A_2} = -\frac{AE_s}{2\pi n} [\cos 2\pi n \cos n\theta - \sin 2\pi n \sin n\theta - \cos n\pi \cos n\theta + \sin n\pi \sin n\theta], \quad (\text{I-10})$$

but $\sin 2\pi n \equiv 0$, and $\sin \pi n \equiv 0$; where n is any integer. Thus these terms can be eliminated from Equation I-10.

$$E_{A_2} = -\frac{AE_s}{2\pi n} [\cos 2\pi n \cos n\theta - \cos \pi n \cos n\theta]. \quad (\text{I-11})$$

For n even,

$$\cos 2\pi n = \cos \pi n;$$

thus,

$$E_{A_2} = -\frac{AE_s}{2\pi n} [\cos n\theta - \cos n\theta] = 0. \quad (\text{I-12})$$

For n odd;

$$\cos 2\pi n = +1$$

$$\cos \pi n = -1;$$

hence, Equation I-11 gives

$$E_{A_2} = -\frac{AE_s}{2\pi n} [\cos n\theta + \cos n\theta] = -\frac{AE_s}{\pi n} \cos n\theta. \quad (\text{I-13})$$

THE DIGITAL TAPE RECORDER ANALYZER*

JOSEPH A. SCIULLI

The test instrument described herein was designed for the purpose of performing error analyses on spaceborne digital tape recorders. In addition to its two basic functions of detecting errors in a clock track recorded on tape and performing a parity check, under various test conditions, the analyzer's design flexibility will enable simple modification of future models for increased capability. This report describes the theory of operation of the present model and provides operating instructions, together with complete drawings and component listings.

INTRODUCTION

The requirement to test spaceborne digital tape recorders during qualification tests and system integration has generated the need for a test instrument capable of making significant analyses. The instrument described herein has two basic capabilities:

1. Detecting errors in a clock track recorded on tape
2. Performing a parity check

Although the instrument has only two basic functions, it can be set up to make tests under various conditions. In addition, flexibility has been incorporated into the design so that future models of the analyzer may be easily modified for increased capability. The purpose of this report is to describe the theory of operation and to provide operating instructions for the tape recorder analyzer.

THEORY OF OPERATION

Background

The tape recorder analyzer is a test instrument designed to perform error analyses on spaceborne digital tape recorders. The analyzer was designed specifically to test a 200-foot endless loop, 8-channel digital tape recorder for both Tiros and Nimous spacecrafts. The tape recorder has a record speed of about 0.4 inch/second and a playback speed of about 10.4 inch/second, giving a

playback speedup ratio of 26:1. If the tape is assumed to be perfect, the main sources of error in the recorder can be attributed to such things as tape dropout, tape flutter, and tape skew. The specification for the tape recorder subsystem states that the analyzer will be used to record signals on one complete tape. For acceptance of the recorder, there shall be less than 24 total errors (clock and/or data errors) on the tape, not including errors at the splice.

The accuracies of both the clock channel and the data channels can be checked with this analyzer. Approximately 75 percent of the instrument consists of micrologic modules, with the remaining 25 percent consisting of conventional solid-state circuitry. Micrologic modules were chosen for the major portion of the design because of their proven reliability for both spacecraft and ground equipment.

The tape recorder analyzer closely simulates the conditions to which the recorder would be subjected during flight operations. The instrument can be set up for a number of different test situations.

The accuracy of the data channels is determined by continuously forcing one data channel to be in a certain state with respect to the other six data channels during record time. This channel is called the "parity channel" and, during playback, a continuous examination is made to insure that this preselected state is maintained. The accuracy of the clock is determined by generating test signals which permit the continuous examination of the clock signal to insure that it falls within the allowed tolerance.

*Published as *NASA Technical Note D-1863*, October 1964.

The following pages give a more detailed and comprehensive explanation of the operation of the tape recorder analyzer.

General Description

Figure 1 shows the main parts of the system with single-line interconnections. The data input to the tape recorder can be selected at two different points:

1. At the input to the 6-bit shift register
2. At the input to the 6-bit storage register

At the input to the 6-bit shift register, either an external known data pattern, an internal generation of pseudo-random data patterns, or the random code generator can be selected. The 6-bit shift register converts the incoming test data from serial to parallel form. The six output

wires of the shift register are fed to a bank of six 3-position switches, where either a known data pattern, a scaler connection of each bit of the 6-bit storage register, or the six output wires from the 6-bit shift register can be selected.

Odd parity is continuously generated during record time by controlling a parity bit with the initial parity checking circuitry and recording the parity bit on the 7th data channel. Odd parity is defined as an odd number of logical 1's contained in N data bits. In the case of 6 bits, odd parity would be indicated if there were one, three, or five 1's in the 6 bits. For example, if A , C , and E in Figure 2 are 1's, odd parity would be indicated. If, however, only A and C are 1's, even parity would be indicated. If *even* parity exists at the recorder inputs, the parity checking circuitry forces the parity to be odd by making

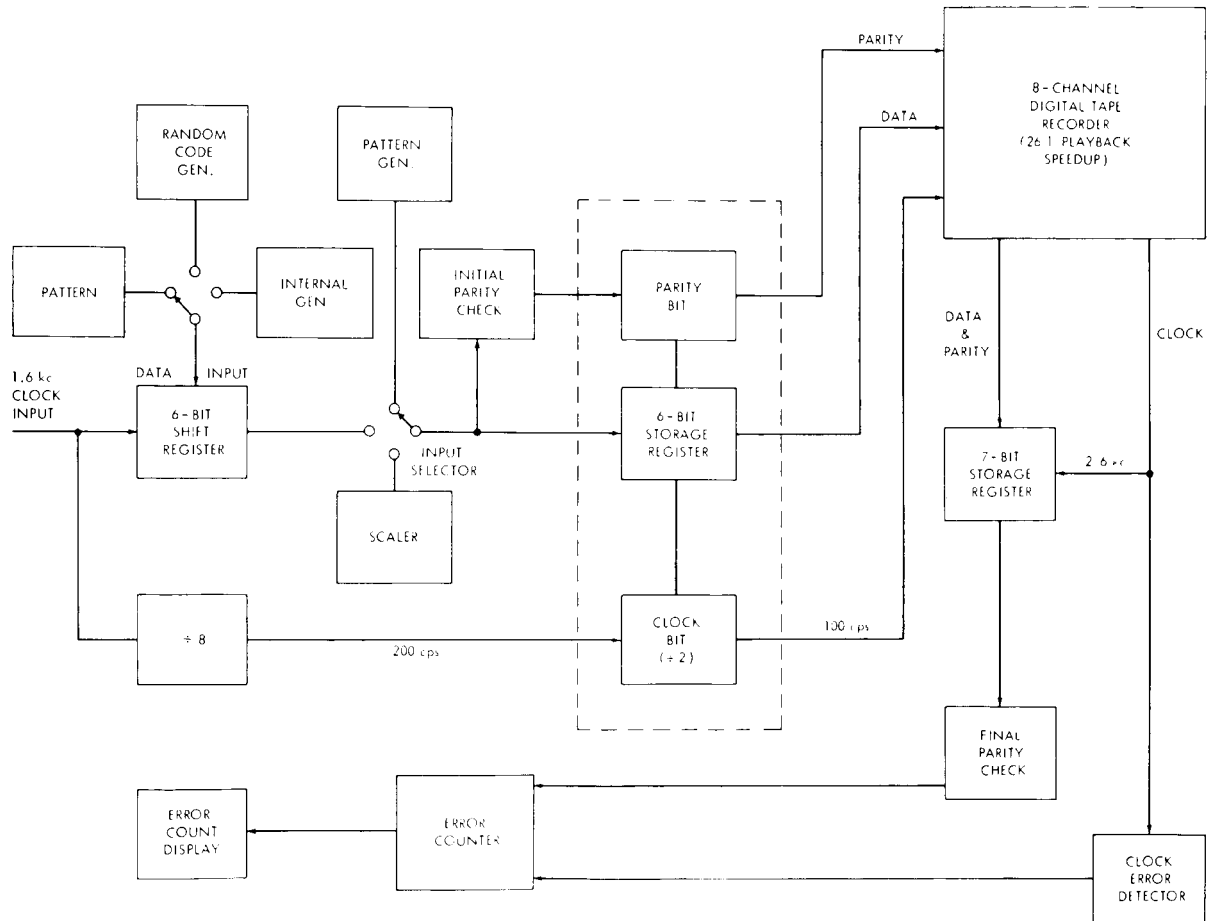


FIGURE 1.—The tape recorder analyzer.

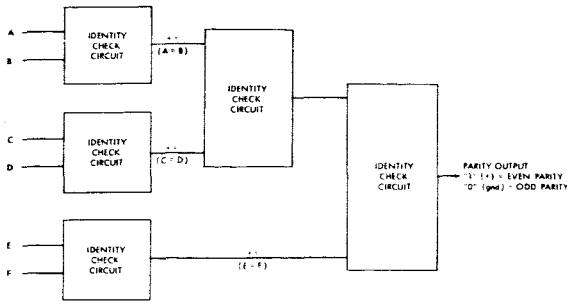


FIGURE 2.—Parity checking network.

the parity bit a 1 (an example is shown in Figure 3). The eight record amplifiers are therefore driven by six data channels, one parity channel, and a 100-cps clock channel which is 90 degrees out of phase with the data so that the data may be strobed at the center of each bit.

During tests the playback outputs of the recorder consist of six data channels, one parity channel, and a 2.6-kc clock channel. The six data channels and the one parity channel are fed in parallel to a 7-bit storage register which is driven by the 2.6-kc clock signal. At the output of the 7-bit storage register the data are checked to insure that odd parity has been maintained. Parity checks are made by a network of identity checking circuits, as shown in Figure 2. The function of the parity checking network is to compare the six inputs and produce an output which indicates that either odd parity or even parity exists. Since the recorder inputs always see an odd number of 1's in the seven data channels, the parity at the outputs also should be odd as long as no error has occurred. If the check made at the output of the recorder indicates

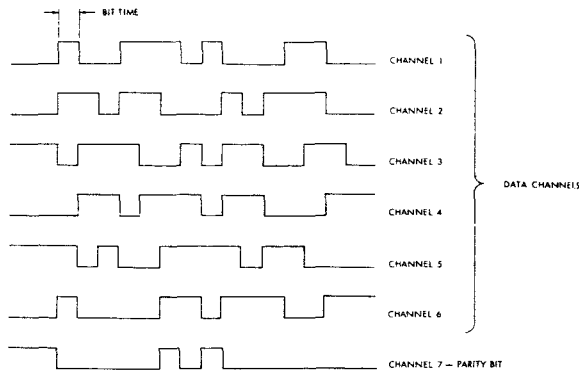


FIGURE 3.—Example: making parity odd.

odd parity, it is assumed that no error has occurred. If even parity is indicated at the output, an error is detected and counted. Under this assumption, it is possible occasionally to count or not count a parity error (data error) falsely; that is, it is possible for a data error to occur without changing parity. This case would occur if, during a certain bit time, one channel changed from 1 to 0 and another channel changed from 0 to 1. However, it has been determined that this event will only occur approximately 0.025 times as often as an error caused by only one channel changing state. It is therefore reasonable to say that the case of the cancelling double error is very remote with respect to the more practical case of a single error. In addition, the tests to be made do not require completely accurate error counts to obtain significant results.

The accuracy of the clock is examined to insure that it falls within the specified tolerance. The data error detector and the clock error detector are both designed so that their outputs provide a pulse each time an error occurs. The outputs of both detectors are then fed to the error counting and error display circuitry.

Clock Errors

Both clock phase errors and missing clock pulses are considered clock errors. The design objective is to develop a countable pulse each time either type of error is detected.

Phase Errors

To develop a phase error detector, each stage of the clock signal must be continuously checked for its accuracy in period within maximum and minimum tolerances. The following discussion refers to the timing diagram of Figure 4. So that both edges of the clock may be tested, a pulse \bar{T} of approximately 10 μ sec in width is generated by each edge of the 2.6-kc clock. Thus, there should be 192.5 μ sec between each developed pulse; a tolerance of 192.5 \pm 30 μ sec has been established. The pulse train \bar{T} is used to trigger one-shot multivibrator A which is cascaded with one-shots B, C, and D. Pulses A and B are each 81 μ sec in width. A and B are gated together to develop a 162- μ sec signal which is used to check the short end (162.5 μ sec) of the tolerance. The pulse

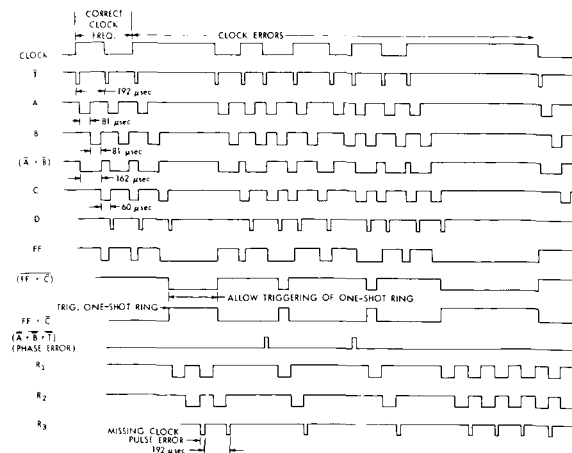


FIGURE 4.—Timing diagram.

train \bar{T} is gated with this 162- μ sec signal to generate $(\bar{A} + \bar{B} + \bar{T})$. An error pulse is developed if the clock signal falls below the 162.5- μ sec lower tolerance limit.

Missing Clock Pulses

To check the long end of the period tolerance, one-shot C is set at 60 μ sec, giving a tolerance coverage—along with A and B—of 222.5 μ sec. As shown in the timing diagram, the falling edge of one-shot C is used to trigger a flip-flop which is reset (through the present input) as soon as the next clock pulse occurs. The output of this flip-flop is in its "1" state (positive voltage) as long as the clock pulses occur within tolerance. If, however, the clock pulses begin to occur out of tolerance, the flip-flop output stays in its "0" state (ground). This allows the triggering of a one-shot ring (3 one-shots cascaded and closed on itself) shown in Figure 5 (also note timing diagram). The purpose of the one-shot ring is to generate a signal which is identical to the expected clock pulse signal \bar{T} . The one-shot ring is turned

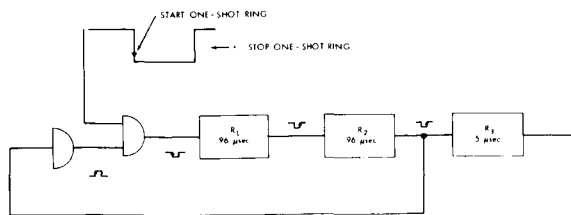


FIGURE 5.—One-shot ring.

on just as the first clock pulse is missed. The pulses generated by the one-shot ring are then counted as errors until the clock is corrected and the one-shot ring is shut off.

Special Circuits

A number of circuits which are utilized in the tape recorder analyzer system require some explanation for a more complete understanding of the entire system.

Random Code Generator

The purpose of the random code generator, shown in Figure 6, is to develop random sets of

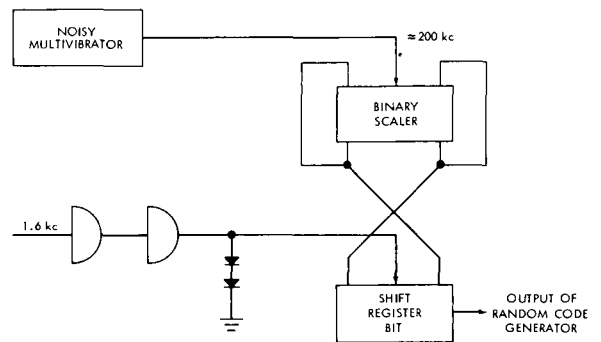


FIGURE 6.—Random code generator.

data in order to test the tape recorder under realistic conditions. A frequency-modulated 200-kc signal is developed by a noisy multivibrator and is fed to a binary scaler whose output is fed to a single-shift register bit, which is strobed by the 1.6-kc clock signal. The noisy multivibrator is simply a standard free-running multivibrator into which 60-cycle noise has been inserted by means of a filament transformer whose secondary is in series with the base-biasing resistor. The output of the multivibrator (Figure 7), a 200-kc frequency-modulated square wave, is then fed to the trigger input of the binary-scaler bit which gives the signal symmetry. Since the triggering (falling) edge of the FM signal is varying in phase and the shift register bit is strobed at a very low frequency compared with the multivibrator frequency, the output of the shift register bit is consequently a random generation of test data bits.

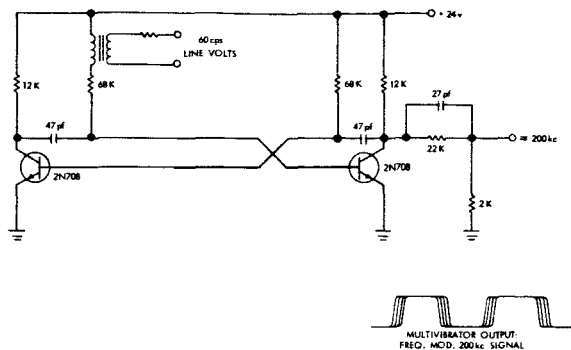


FIGURE 7.—Frequency-modulated multivibrator.

One-Shot Multivibrator

A number of references have been made herein to one-shot multivibrators. All the one-shots mentioned are essentially the circuit shown in Figure 8 with various values of capacitance resulting in various pulse widths.

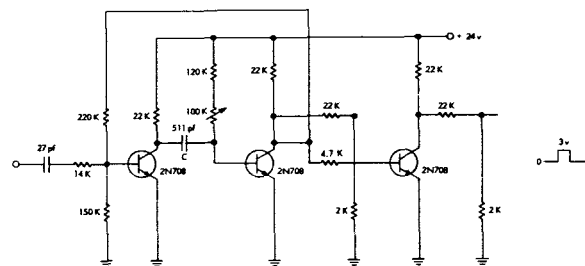


FIGURE 8.—One-shot multivibrator.

Pulse Generator—Frequency Doubler

In the discussion of the clock error detector it was mentioned that, to check both edges of the clock, a 10-µ sec pulse is generated at each edge. The circuit which performs this function is shown in Figure 9. The timing diagram shown in Figure 9 essentially explains the function of the circuit. The falling edge of the 3-kc clock signal triggers flip-flop A, which sets up the proper condition at the R and S inputs of flip-flop B, which is triggered at the much faster 200-kc rate. As soon as flip-flop B switches, its output is fed back to the preset of flip-flop A and resets it, thus resulting in a pulse output at each falling edge. The 200-kc signal need not be phase-coherent with the 3-kc clock signal. Performing an identical process on the opposite phase of the clock and combining

the two pulse outputs results in a frequency doubling.

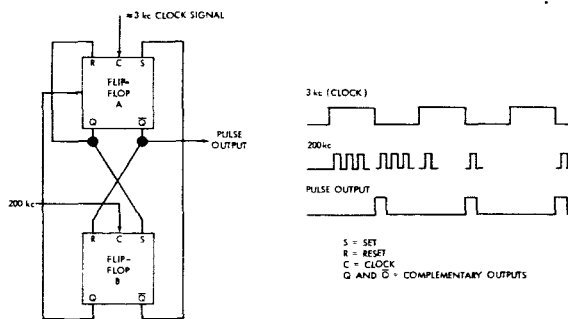


FIGURE 9.—Pulse generator-frequency doubler.

Additional Design Considerations

General

Since the analyzer will be used by a number of individuals under various conditions, substantial mechanical and electronic protection was incorporated into the instrument. At the same time, the analyzer was designed for ease of operation and maintenance.

Mechanical and Electronic Protection

Occasionally, a power supply or signal generator is inadvertently applied to a piece of equipment in a manner that could seriously and permanently damage the equipment. For this reason, every power supply and signal generator input to the analyzer has been provided with special protection circuits which insure that dangerously high signals are not applied directly to the delicate micrologic modules and conventional solid-state circuitry.

The mechanical frame of the analyzer makes it adaptable for either enclosure or rack mounting, in addition to providing protection for the components. Thus the analyzer is a single self-contained unit, except for power supply and signal generators.

All input jacks which take a standard G-R connector have been spaced such that the connection can be made in only one way and thus eliminate the possibility of dangerous shorting of independent signals.

Operation and Maintenance

For ease of operation the analyzer has been designed so that it can be activated regardless of any

switch setting. However, to use the instrument properly, the operator should have a thorough knowledge of the proper switch settings and procedure. Since all the main controls and monitoring points are located on the front panel with appropriate identifications, the operation of the analyzer during tests is very simple.

Maintenance also is easy to perform, since every circuit card and every solder joint can be reached without turning a screw.

Flexibility

As previously mentioned, the analyzer was designed to test the 200-foot endless loop, 8-channel digital recorder. Because of the flexibility incorporated into the design, it would be easy to modify the analyzer to test other digital recorders as well. If the specifications for qualification should change at any time, it is a simple matter to recalibrate a number of one-shots so that the analyzer may be utilized.

Another important consideration is the ability of this instrument to analyze prerecorded tapes which have been recorded with either even or odd parity. This capability means that it is possible to analyze tapes during flight operations. The analyzer therefore could be included as part of the ground station and could be used to determine the accuracy of the recorder and the tape at any time and to determine the reliability and accuracy of

the recorder by making life tests during flight operations. It also is possible to locate defective spots on the tape to some degree. Utilizing these capabilities, the analyzer could conceivably aid in making more significant data analyses.

OPERATING INSTRUCTIONS

General

To utilize the tape recorder analyzer to maximum advantage, one should become as familiar as possible with the mechanics of operating the instrument. During the following discussion, refer to Figure 10 to become familiar with the controls and procedure.

Initial Procedure

Equipment

In addition to the analyzer itself, a few additional pieces of equipment are necessary to perform tests on the 8-channel recorder:

- 3 Power Supplies: +3 v at 75 ma, -3 v at 10 ma, and +24 v at 110 ma
- 1 Frederick Pulse Pattern Generator, Model 201 (or equivalent): capable of 35 bits
- 2 Square-Wave Generators, HP Model 211A (or equivalent)
- 1 Strip Chart Recorder
- 1 Dual-Trace Oscilloscope, Tektronix Type 533A (or equivalent)

Front Panel—Description

On the front panel [Figure 10(a)] are located all the switches and test points necessary for normal operation of the analyzer during tests. The DATA SELECT SWITCH chooses the type of test data pattern to be used. The ERROR COUNT SELECTOR switch chooses the error count to be displayed. The DATA INPUT SELECTOR SWITCHES choose the connection of the data storage register according to the test being made. The PRESET PB (pushbutton) resets the error count display to zero. The patch panel (6×8 array of jacks) is used to cross-patch record the and playback channels. The various test points and toggle switches are all utilized during normal tests.

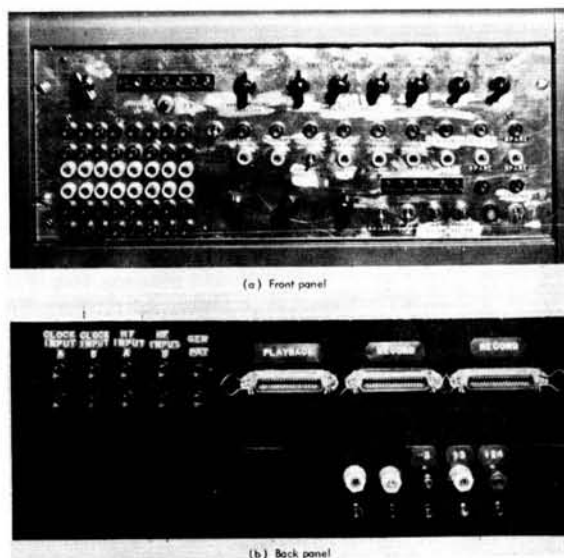


FIGURE 10.—The tape recorder analyzer control panels.

Back Panel—Description

On the back panel [Figure 10(b)] of the analyzer are located all power supply and signal generator inputs used during normal testing. Also on the back panel are located three 36-pin connectors which are used to connect the record and playback channels of the recorder to the analyzer. The two connectors marked RECORD are logically identical. This means that two recorders could be recorded simultaneously. However, only one recorder can play back at a time.

Patch Panel

At the lower left of the front panel a 6×8 array of jacks makes up a patch panel on which any of the eight record channels of the analyzer may be directed to any of the record amplifiers of the recorder. Each pair of vertically opposite, identically colored jacks is a jumper pair in series with one of the record or playback channels. The red set of jacks (upper two rows) is in series with record output connector 1, the yellow set (middle two rows) is in series with record output connector 2, and the green set (lower two rows) is in series with the playback input connector. On normal testing of one recorder, each red pair and each green pair is shorted, with the yellow set of jacks open. To record on two taps simultaneously, the yellow pairs of jacks also must be shorted. Each pair of jacks has been positioned so that it can be shorted with a standard G-R connector. For any other patching arrangement use the patching leads provided.

Power Supply and Signal Generator Connectors

Make the scope trigger and scope ground connections on the front panel. On the back panel, connect *but do not turn on* +3v, -3v, and +24v power supply voltages, CLOCK INPUT HF input (100-kc), and GEN PAT (pattern generator). On the front panel choose proper position for the clock select switch (CLOCK SEL SW), high-frequency select switch (HF SEL SW), and pattern-generator select switch (PAT GEN SEL SW) according to the connections made.

Switches and Controls

On the front panel insure that LOAD PULSE SELECT SWITCH and ONE SHOT SELECT

SW are both in the "normal" positions. Choose PAT (pattern generator), RC (random code generator), or INT (internal generator) at DATA SELECT SWITCH. At the bank of six DATA INPUT SELECTOR SWITCHES, choose SC (scaler), DIR (direct) or EXT (external) inputs.

Power ON

Turn on +3v, -3v, +24v, CLOCK INPUT, HF INPUT (100-kc), and GEN PAT on the back panel. Utilizing the monitoring tests points on both the front and back panels, check power supplies and signal generators for their accuracy. Depress PRESET PB (front panel) to reset the error counting mechanism.

Test Procedure

General

The specification and statement of work for the 8-channel digital tape recorder subsystem specifies that the tape recorder analyzer is to be utilized in tests before vacuum thermal. The specification states that the analyzer will be used to record a complete tape. There shall be less than 24 total errors (either clock or data errors) not including errors at the splice. To perform this test properly, the following procedure should be used.

Test Setup

The initial test procedure described in the section "Initial Procedure" should be performed before beginning tests on the recorder. The operator should be satisfied that the analyzer is performing satisfactorily in the RECORD mode. He should insure that all record channels of the analyzer are providing appropriate test data according to the data input selected. Every record channel of the analyzer also should be checked to insure that the amplitude of a logical 0 is approximately -0.1 volt in amplitude. The record amplifiers are so sensitive that a logical 0 of a few tenths of a volt of positive amplitude could be interpreted as a logical 1. The level of the logical 0's can be adjusted by controlling the -3 volt power supply. The 1.6-kc clock signal and the 100-kc high-frequency signal must be square waves with amplitudes of approximately 3 volts above ground. The external oscillator input

of the pattern generator must be driven by the 1.6-kc clock signal, and the NRZ output of the pattern generator should be applied to the pattern generator input of the analyzer. The amplitude of the pattern generator also should be approximately +3 volts above ground.

When satisfactory operation of the analyzer is achieved, the record channels of the analyzer should be connected to the inputs of the record amplifiers using the harness provided. During recording there should be no connection between the playback channels of the analyzer and the playback amplifiers. A full recording requires approximately 100 minutes. This time can vary somewhat, and therefore at least 110 minutes should be allowed for a full recording.

After a full tape has been recorded, the record connections should be removed and the playback connection made. The D-A converter output on the front panel should be connected to one channel of the strip chart recorder. The purpose of the D-A converter and the strip chart recorder is to give indication of the location of the splice and also the number of errors which are encountered at the splice. The indicator lights on the front panel give a digital readout of the number of errors counted. At the end of the playback cycle the number of errors at the splice is subtracted from the error count indicated by the

lights. Note that it is possible to count parity errors only, clock errors only, or both simultaneously. The preset pushbutton is used whenever the error count is to be reset to zero.

CONCLUDING REMARKS

The tape recorder analyzer is a test instrument which was specifically designed to aid both NASA and contractor personnel in testing and qualifying the 8-channel digital recorder. The instrument has been electronically and mechanically designed to be simple to operate and maintain, and as foolproof as possible. In addition to performing the primary function of making error analyses on the 8-channel digital recorder, it is believed that with careful operation the analyzer can be conveniently utilized to troubleshoot and locate sources of error in the recorder at the same time. This instrument can be easily modified to test future digital tape recorders and can be used to analyze prerecorded tapes.

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Mr. Marvin S. Maxwell, Mr. Sebastian M. Costa, and Mrs. Marjorie R. Townsend for their assistance in preparing this paper.

APPENDIX A

COMPLETE DRAWINGS AND COMPONENT LISTINGS

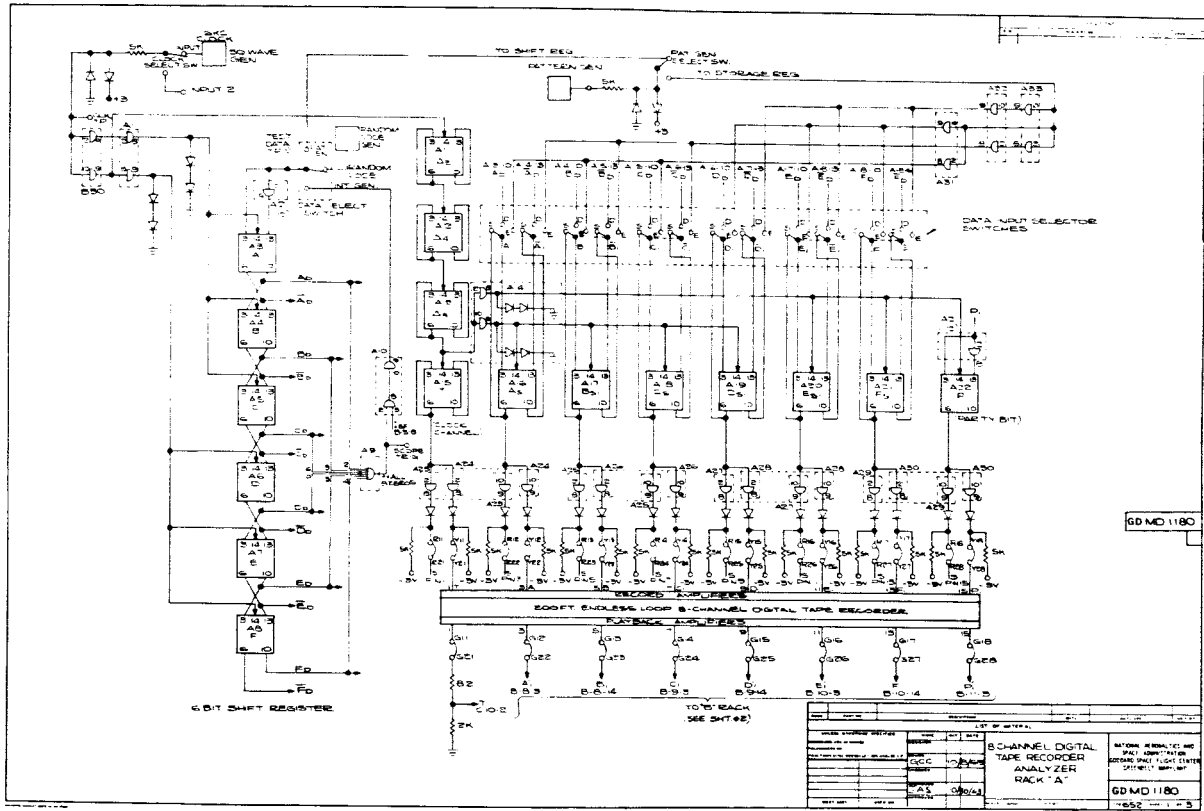


FIGURE A1.—(unnumbered sketch)

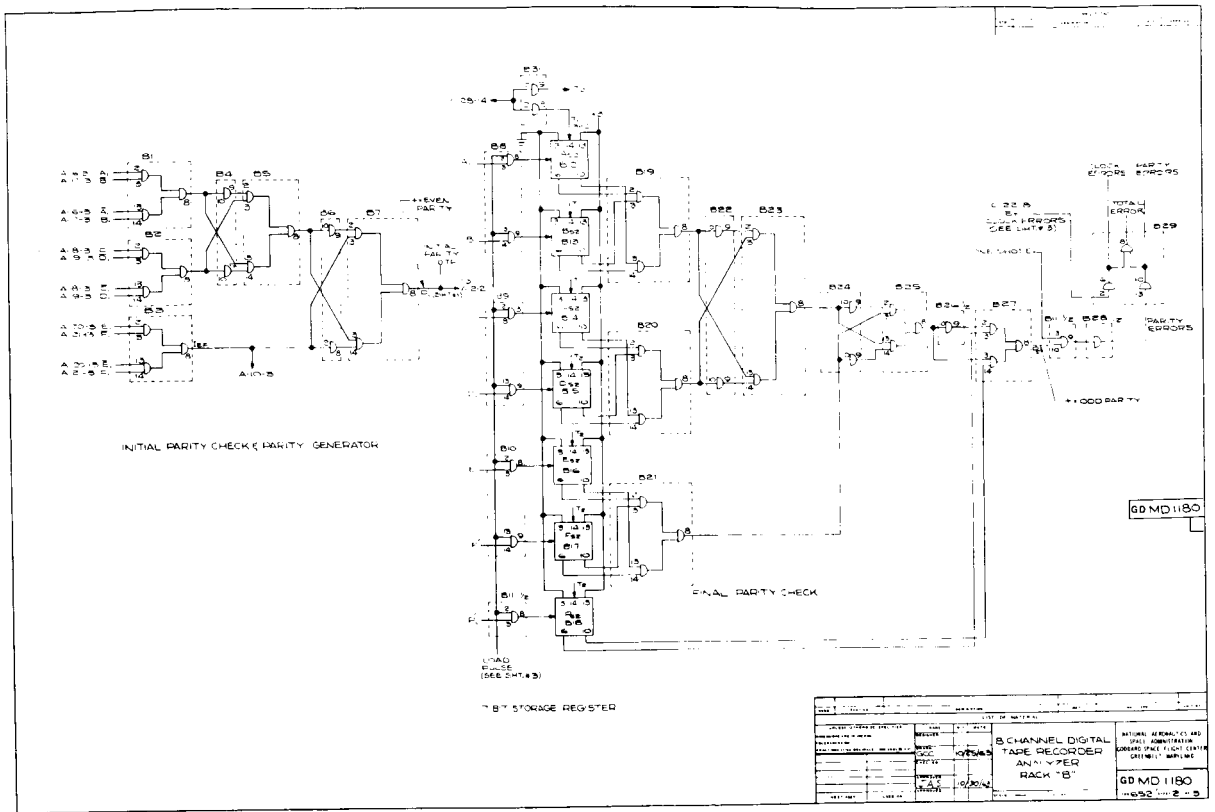


FIGURE A2.—(unnumbered sketch)

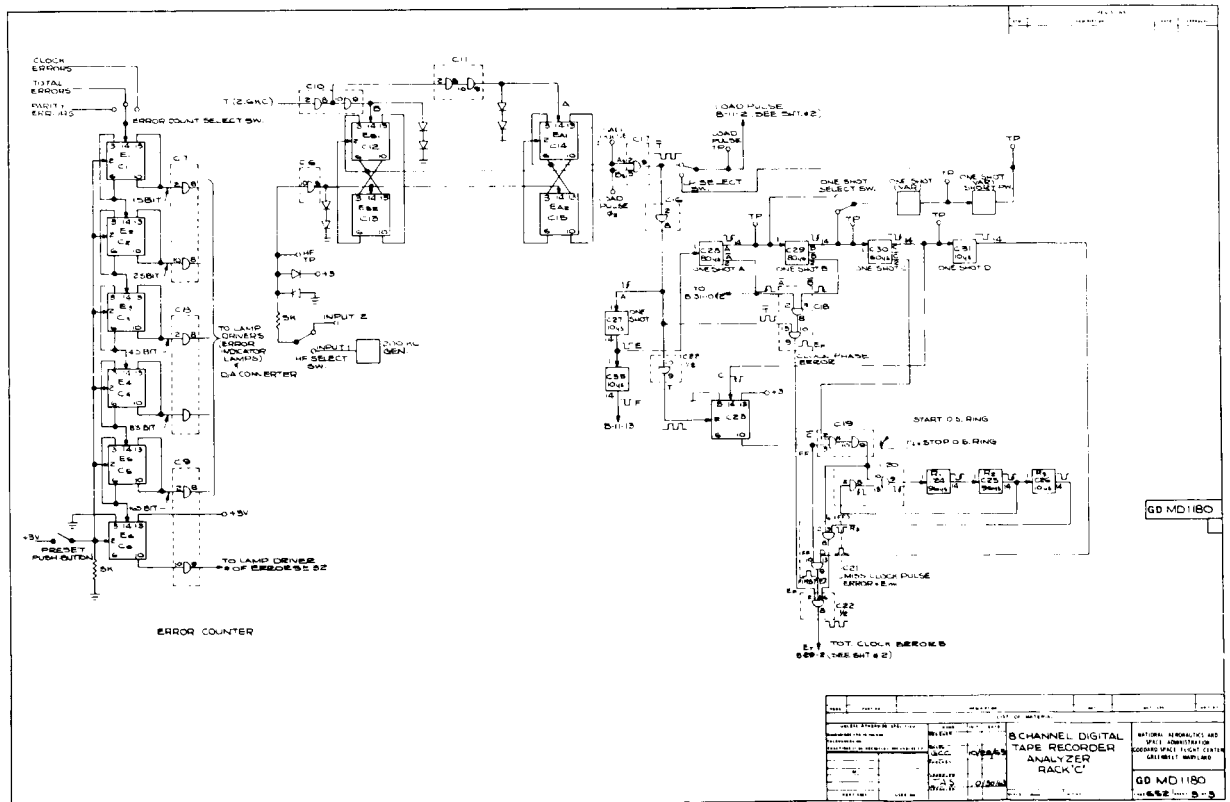


FIGURE A3.—(unnumbered sketch)

Component Listing*

Module Identification	Module Type	Module Identification	Module Type	Module Identification	Module Type
A-1	SN514	B-1	SN515	C-1	SN510
A-2	SN514	B-2	SN515	C-2	SN510
A-3	SN510	B-3	SN515	C-3	SN510
A-4	SN510	B-4	SN514	C-4	SN510
A-5	SN510	B-5	SN515	C-5	SN510
A-6	SN510	B-6	SN514	C-6	SN510
A-7	SN510	B-7	SN515	C-7	SN514
A-8	SN510	B-8	SN514	C-8	SN514
A-9	SN512	B-9	SN514	C-9	SN514
A-10	SN514	B-10	SN514	C-10	SN514
A-11	SN510	B-11	SN514	C-11	SN514
A-12	SN510	B-12	SN510	C-12	SN510
A-13	SN510	B-13	SN510	C-13	SN510
A-14	SN514	B-14	SN510	C-14	SN510
A-15	SN510	B-15	SN510	C-15	SN510
A-16	SN510	B-16	SN510	C-16	SN514
A-17	SN510	B-17	SN510	C-17	SN513
A-18	SN510	B-18	SN510	C-18	SN514
A-19	SN510	B-19	SN515	C-19	SN514
A-20	SN510	B-20	SN515	C-20	SN514
A-21	SN510	B-21	SN515	C-21	SN514
A-22	SN510	B-22	SN514	C-22	SN514
A-23	SN514	B-23	SN515	C-23	SN510
A-24	SN514	B-24	SN514	C-24	One-Shot 96 μ sec (R_1)
A-25	SN514	B-25	SN515	C-25	One-Shot 96 μ sec (R_2)
A-26	SN514	B-26	SN514	C-26	One-Shot 10 μ sec (R_3)
A-27	SN514	B-27	SN515	C-27	One-Shot 10 μ sec (E)
A-28	SN514	B-28	SN514	C-28	One-Shot 80 μ sec (A)
A-29	SN514	B-29	SN515	C-29	One-Shot 80 μ sec (B)
A-30	SN514	B-30	SN514	C-30	One-Shot 60 μ sec (C)
A-31	SN514	B-31	SN514	C-31	One-Shot 10 μ sec (D)
A-32	SN514	B-32	SN510	C-32	5 Lamp Drivers
A-33	SN514	B-33	SN510	C-33	One-Shot 10 μ sec (F)

- *1. Each module shown on the drawings is identified with a letter — number designation. The letter (A, B, or C) denotes the rack in which the module is located. The number (1 thru 33) denotes the particular socket in which the module is located.
2. Each module having a module type designation of 3 digits preceded by an SN prefix is one of the Texas Instrument Series 51 micrologic modules. Additional information on this series of modules is available upon request. Any other module-type designation is a printed circuit card containing the circuitry described.

DESIGNING TOROIDAL INDUCTORS WITH DC BIAS*

G. D. SMITH

A practical method of designing toroidal inductors with dc bias is presented. The toroidal core equations for the inductance, magnetizing force, and core winding factor are used to derive two equations for selecting the smallest possible core. Acceptable requirements are established for the variation of Q with frequency and for the variation of inductance with temperature. The inductor equations and requirements indicate the desirable combinations of core constants and from these the smallest core may be determined.

A step-by-step procedure for establishing the inductor requirements, selecting the smallest core, and designing the inductor is given, explained, and illustrated. Tables and curves employed in this method of design are presented.

INTRODUCTION

Selection of the best magnetic core for an inductor frequently involves trial-and-error type calculations. Comparison of the physical parameters of the core with the basic inductor equations and requirements reveals a clear, accurate, and concise method of designing inductors with dc bias. Equations derived from the basic inductor equations can be used to select the smallest core and auxiliary inductor requirements can be used to obtain an optimum design.

The design of an inductor frequently involves consideration of its effect upon the magnetic field in which it is placed. This is especially true in the design of high-current inductors for converters and regulators used in satellites which employ sensitive magnetic field detectors. For this type of design problem it is frequently imperative that a toroidal core be used, because of the low magnetic effect external to the core. The core should be small to minimize the amount of magnetic material in the inductor. A large number of air gaps should be spaced evenly around the circumference and across the cross-section of the core.

The molybdenum permalloy toroidal powder cores are well suited to this type of design problem. A large variety of sizes have high permeability with high Q at high frequency. A discussion of the characteristics of these cores will be used to illustrate this method of designing toroidal

inductors. All curves and data are for molybdenum permalloy powder cores. They are not intended to cover all commercially available cores,† but only to augment the design discussion.

DERIVATION OF EQUATIONS

The smallest core which fulfills the inductor requirements of inductance and inductance tolerance at peak current can be determined from the basic equations of the core. The inductance of a toroidal inductor is

$$L = \frac{0.4\pi N^2 \mu A}{10^8 l} \quad (1)$$

where

L = the inductance in henries,

N = the number of turns of wire,

μ = the relative permeability,

A = the cross-sectional area of the core in square centimeters,

l = the mean magnetic path length in centimeters.

In selecting the smallest core to meet the requirements of Equation 1, the maximum number of turns of wire which can be placed on the core must

*Published as NASA Technical Note D-2320, June 1964.

†All magnetic core data are taken from Magnetics, Inc., Catalog No. PC-303.

be known. This is a function of the ratio of the core window area (W) to the wire cross-sectional area (A_w). Defining a winding factor (K_w) as the ratio of the total wire area (NA_w) to the core window area (W), we can write

$$K_w = \frac{NA_w}{W} \quad (2)$$

where A_w and W are in circular mils. Experience in winding many cores of various sizes shows that the maximum practicable winding factor is approximately 0.4. This means that approximately 60 percent of the window area is required for wire insulation and space between the wires, leaving only 40 percent ($K_w=0.4$) for the wire. The smallest toroidal core which will meet the requirements of Equations 1 and 2 can be selected by combining these equations. Solving Equations 1 and 2 for N^2 and setting them equal gives

$$LA_w^2 = \frac{0.47\mu AK_w^2 W^2}{10^8 l} \quad (3)$$

where the left side contains the inductor requirements and the right side the core and winding factor constant. Thus, the smallest core which meets the requirements of Equations 1 and 2 is one which has a core and winding factor constant equal to or next greater than the inductor requirements on the left of Equation 3.

The relative permeability and inductance of an inductor vary with the magnetizing force. This relationship is shown in Figure 1 for some powder permalloy cores of different relative permeabilities. This family of curves reveals that excessive variations in the inductance may occur unless the magnetizing force is limited. For this reason an inductance tolerance is imposed; it is

transformed into a maximum magnetizing force requirement by use of manufacturer curves similar to those in Figure 1. The magnetizing force is related to the inductor current by Ampere's law:

$$H = \frac{0.47NI}{l} \quad (4)$$

where H is the magnetizing force in oersteds and I is the current in amperes. Equation 4 is employed to impose maximum requirements, therefore, $H_{\pm\%}$ denotes the maximum value of the magnetizing force for the inductance tolerance indicated in percent from the nominal and I_p denotes the corresponding peak current. The number N in Equation 4 cannot be determined until the core is selected. For this reason N is replaced by its equivalent from Equation 1, and the core constants are placed on the right:

$$LI_p^2 = \frac{\mu A l H_{\pm\%}^2}{0.4710^8} \quad (5)$$

Therefore, the core chosen must have a constant equal to or next greater than the inductor requirements on the left of this equation.

The smallest core which meets the requirements of Equation 1, 2, and 4 is one with the core constants on the right of Equations 3 and 5 equal to or greater than the respective inductor requirements on the left.

DESIGN PROCEDURE

A detailed procedure for selecting the core and designing the inductor follows:

1. Form or obtain a table, similar to Table 1, of all available cores for a suitable inductance tolerance. Table 1 was made for 10 percent inductors since the magnetizing force ($H_{\pm 10}$) was taken from Figure 1 where the inductance is down 10 percent. The core number in column 1 of Table 1 indicates the physical size and permeability of the core.

2. Select the wire size to be used for winding the core. The maximum rms current shown in Table 2 should be used as a guide for this selection. In case a very low dc resistance or very high Q is desired, a slightly large wire should be selected.

3. Compute the product (Equation 3) of the inductance and the square of the wire cross-sectional area (LA_w^2).

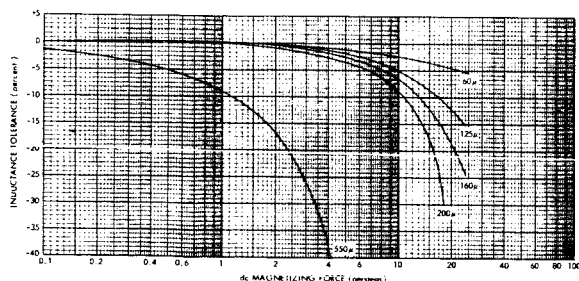


FIGURE 1.—Inductance tolerance vs dc magnetizing force for several powder permalloy cores.

Table 1
Molybdenum Permalloy Powder Core Data.

Core Number	μ	H_{110} (oersteds)	Inside Diameter (in.)	Outside Diameter (in.)	Height (in.)	A (cm^2)	l (cm)	W (cir mil)	$\frac{0.4\pi\mu AK_w^2 W^2}{10^8 l}$ [cm^3 (cir mil) 2]	$\frac{\mu A H_{110}^2}{0.4\pi 10^8}$ [cm^3 oersteds 2]	Frequency Range (kc)	Inductance vs. Temperature (see Table 4)	K_1 (ft./turn)
55894	60	45	0.555	1.09	0.472	0.635	6.35	320,000	1.23×10^3	3880×10^{-6}	10-50	A2, B4, D4, W4	0.135
55930	125	17.5	0.555	1.09	0.472	0.635	6.35	320,000	2.58×10^3	1230×10^{-6}	0-20	B4, D4, E4, W4	0.135
55928	160	13	0.555	1.09	0.472	0.635	6.35	320,000	3.30×10^3	870×10^{-6}	0-10	A2, D4, W4	0.135
55927	200	11	0.555	1.09	0.472	0.635	6.35	320,000	4.12×10^3	774×10^{-6}	0-7	A2	0.135
55926	550	1.2	0.555	1.09	0.497	0.635	6.35	320,000	11.3×10^3	25.6×10^{-6}	0-4	A2	0.135
55071	60	45	0.760	1.332	0.457	0.655	8.10	590,000	3.395×10^3	5120×10^{-6}	10-50	A2, B4, D4, W4	0.146
55548	125	17.5	0.760	1.332	0.457	0.655	8.10	590,000	7.06×10^3	1620×10^{-6}	0-20	A2, B4, D4, W4	0.146
55546	160	13	0.760	1.332	0.457	0.655	8.10	590,000	9.04×10^3	1145×10^{-6}	0-10	A2, D4, W4	0.146
55545	200	11	0.760	1.332	0.457	0.655	8.10	590,000	11.3×10^3	1030×10^{-6}	0-7	A2	0.146
55585	125	17.5	0.888	1.385	0.387	0.449	8.91	810,000	8.3×10^3	1220×10^{-6}	0-20	A2, D4, W4	0.137
55583	160	13	0.888	1.385	0.387	0.449	8.91	810,000	10.6×10^3	864×10^{-6}	0-10	A2, D4, W4	0.137
55582	200	11	0.888	1.385	0.387	0.449	8.91	810,000	13.3×10^3	776×10^{-6}	0-7	A2	0.137
55076	60	45	0.848	1.445	0.444	0.670	8.99	730,000	4.8×10^3	5800×10^{-6}	10-50	A2, B4, D4, W4	0.151
55324	125	17.5	0.848	1.445	0.444	0.670	8.99	730,000	10.0×10^3	1835×10^{-6}	0-20	A2, B4, D4, W4	0.151
55322	160	13	0.848	1.445	0.444	0.670	8.99	730,000	12.8×10^3	1300×10^{-6}	0-10	A2, D4, W4	0.151
55321	200	11	0.848	1.445	0.444	0.670	8.99	730,000	16.0×10^3	1170×10^{-6}	0-7	A2	0.151
55083	60	45	0.918	1.602	0.605	1.056	9.87	860,000	9.55×10^3	$10,000 \times 10^{-6}$	10-50	A2, D4, W4	0.188
55254	125	17.5	0.918	1.602	0.605	1.056	9.87	860,000	19.8×10^3	3180×10^{-6}	0-20	A2, D4, W4	0.188
55252	160	13	0.918	1.602	0.605	1.056	9.87	860,000	25.5×10^3	2250×10^{-6}	0-10	A2	0.188
55251	200	11	0.918	1.602	0.605	1.056	9.87	860,000	31.8×10^3	2020×10^{-6}	0-7	A2	0.188

Table 2

Resistance, Area, and Current Capacity of Various AWG Wire Sizes.

AWG Wire Size	Resistance/Length (ohm/ft)	Wire Area* (cir mil)	Current Capacity† (amps)
8	0.00063	23,000	16.5
9	0.00079	16,500	13.1
10	0.00100	11,700	10.2
11	0.00126	8890	821
12	0.00159	7100	6.52
13	0.00200	5660	5.18
14	0.00252	4520	4.12
15	0.00318	3610	3.26
16	0.00402	2900	2.58
17	0.00505	2330	2.05
18	0.00639	1860	1.63
19	0.00805	1490	1.29
20	0.01013	1197	1.024
21	0.0128	961	0.8123
22	0.0162	767	0.6401
23	0.0203	620	0.5108
24	0.0257	497	0.4040
25	0.0324	400	0.3204
26	0.0410	320	0.2528
27	0.0514	259	0.2016
28	0.0653	207	0.1588
29	0.0822	169	0.1277
30	0.104	134	0.1000
31	0.131	108	0.0792
32	0.162	88.3	0.0640
33	0.206	70.5	0.0504
34	0.261	56.2	0.0397
35	0.331	44.9	0.0314
36	0.415	36.0	0.0250
37	0.512	29.1	0.0202
38	0.648	23.0	0.0160
39	0.847	17.6	0.0122
40	1.074	14.4	0.0096
41	1.32	11.6	0.0078
42	1.66	9.0	0.0063
43	2.14	7.3	0.0048
44	2.59	6.2	0.0040

*Based on nominal diameter of wire with heavy Formvar insulation.

†Based on 1000 cir mil per amp.

4. Compute the product (Equation 5) of the inductance and the square of the peak current (LI_p^2).

5. Determine the acceptable core permeabilities for the maximum inductor frequency. The

approximate frequency ranges for core permeabilities are shown in Table 3. In view of the fact that the frequency ranges are approximations, they should not be followed rigidly. Table 3 is based on the variations of Q with frequency as

Table 3
Frequency Range of Cores.

Core Permeability	Frequency Range (kc)
60	10-50
125	0-20
160	0-10
200	0-7
550	0-4

shown in Figures 2-5 for several different core sizes and permeabilities. Comparison of these curves reveals that Q varies with core size for a certain core permeability, frequency, and inductance. For this reason, these more exact curves cannot be used readily until the core size is determined. Because of the inaccuracies introduced in making the approximations for Table 3, it is frequently more convenient to omit this step and consider the permeabilities of the cores (step 7).

6. Determine the acceptable variation in inductance as a function of temperature. A list of inductance tolerances over specific temperature ranges, similar to Table 4, is usually available from the core manufacturer.

Table 4
Variation of Inductance with Temperature

Core Identification (commercial designation)	Inductance Tolerance vs. Temperature
A2	Not specified
B4	± 0.1 percent from 13° to 35°C
D4	± 0.1 percent from 0° to 55°C
E4	± 0.15 percent from 0° to 55°C
W4	± 0.25 percent from -55° to +85°C

7. Select the core on the basis of the information obtained in steps 3, 4, 5, and 6. That is, select the smallest core from a table similar to Table 1 which: (1) has a constant for the core and winding factor, $0.4\pi\mu AK_w^2 W^2 / 10^8 \ell$, that is equal to or next greater than the inductor requirements LA_w^2 ; (2) has a constant for the core, $\mu A(H_{\pm 10}^2) / 0.4\pi 10^8$, that is equal to or next greater than the inductor requirements LI_p^2 ; (3) and meets the

requirements established in steps 5 and 6. The frequency range from Table 3 (step 5) and the inductance tolerance vs. temperature characteristic available from Table 4 (step 6) are also included in Table 1. Thus, these four columns (10, 11, 12, and 13) of Table 1 should be used for selecting the core.

In view of the knowledge obtained about the core size, the acceptable core permeabilities, determined in step 5, should be reconsidered by use of curves similar to those of Figures 2-5 for the individual cores.

8. Compute the number of turns of wire for the selected core to give the required inductance by using Equation 1.

9. It is frequently desirable to use the largest wire possible on the core to decrease the dc resistance and power loss. This can be accomplished by using Equation 2 with the winding factor K_w set equal to 0.4. The dc resistance of the inductor should be calculated from

$$R_{dc} = K_1 N K_2, \quad (6)$$

where

R_{dc} = the dc resistance of the inductor in ohms,

K_1 = the average length of wire per turn in ft (Table 1),

K_2 = the resistance per unit length of the wire in ohms/ft (Table 2).

10. Using Equation 4, compute the maximum magnetizing force (H_{max}) for peak current. Refer to a curve similar to Figure 1 and determine

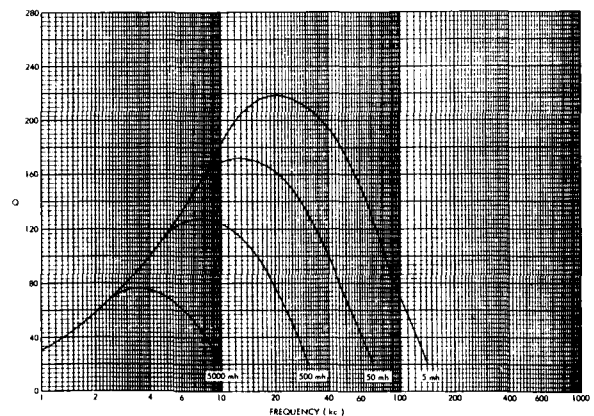


FIGURE 2.—Curves of Q vs frequency for core 55894; $\mu = 60$, $A = 0.635 \text{ cm}^2$, $\ell = 6.35 \text{ cm}$, $W = 320,000 \text{ cir mil}$.

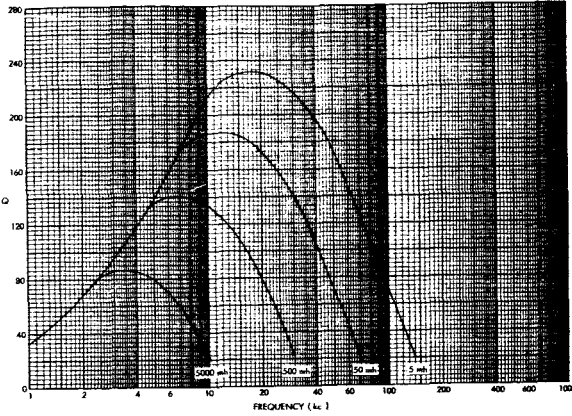


FIGURE 3.—Curves of Q vs frequency for core 55071; $\mu=60$, $A=0.655$ cm², $\ell=8.1$ cm, $W=590,000$ cir mil.

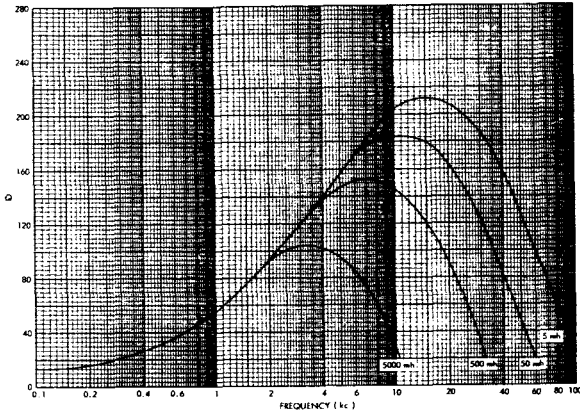


FIGURE 4.—Curves of Q vs frequency for core 55930; $\mu=125$, $A=0.635$ cm², $\ell=6.35$ cm, $W=320,000$ cir mil.

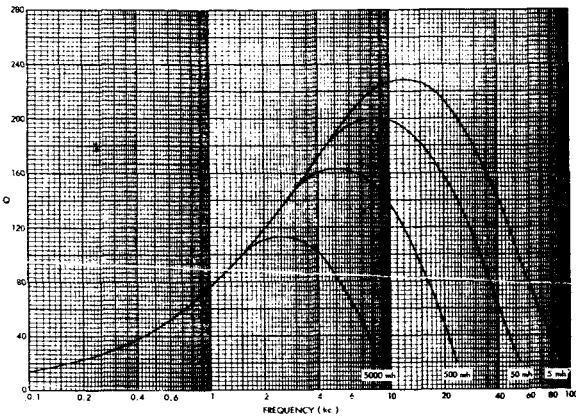


FIGURE 5.—Curves of Q vs frequency for core 55548; $\mu=125$, $A=0.655$ cm², $\ell=8.1$ cm, $W=590,000$ cir mil.

the maximum change in inductance due to the maximum magnetizing force.

EXAMPLE

A toroidal inductor will be considered which has an inductance of 5 mh with a peak current of 0.55 amp, an rms current of 0.5 amp, and a maximum frequency of 10 kc. From the previous section, the following steps are performed:

1. Table 1 may be used in this example.
2. From Table 2, select No. 23 AWG wire for which $A_w=620$ cir mil.
3. Solve $LA_w^2=5 \times 10^{-3} (620)^2=1920$ henries (cir mils)²
4. Solve $LI_p^2=5 \times 10^{-3} (0.55)^2=1.51 \times 10^{-3}$ henries (amps)²
5. From Table 3, permeabilities of 60 and 125 are considered suitable.
6. Since a wide temperature range is desired, the W4 cores from Table 4 should be used.
7. Two cores of the same size in Table 1 meet the requirements of steps 3, 4, 5, and 6—cores 55071 and 55548. The curves of Q vs. frequency for a 5 mh inductance using these two cores are shown in Figures 3 and 5. These curves indicate a high Q well beyond the maximum frequency for this inductor. The highest permeability core (55548) is selected to permit the use of large wire having a low dc resistance.

8. From Equation 1

$$N^2 = \frac{Ll}{0.4\pi\mu A} = \frac{5 \times 10^{-3} (8.10) 10^8}{0.4\pi (125) (0.655)}$$

$$= 3.94 \times 10^4$$

$$N = 198 \text{ turns.}$$

9. The largest wire which can be used can be determined from Equation 2,

$$A_{w,max} = \frac{K_w W}{N} = \frac{0.4 (590,000)}{198}$$

$$= 1190 \text{ cir mil.}$$

From Table 2, No. 20 wire can be used. The dc resistance is given by

$$R_{dc} = K_1 N K_2 = (0.146)(198)(0.01013)$$

$$= 0.292 \text{ ohm.}$$

10. The magnetizing force at peak current is given by Equation 4:

$$H = \frac{0.4\pi N I_P}{l} = \frac{0.4\pi (198)(0.55)}{8.10}$$

$$= 16.9 \text{ oersteds.}$$

Reference to Figure 1 reveals the inductance tolerance to be 9 percent.

The design results for cores 55071 and 55548 are shown in Table 5. This table reveals that the high permeability core yields the lowest number

Table 5

Inductor Design Using Cores of the Same Size.

Core Number	55071	55548
Permeability	60	125
N (turns)	286	198
Q at 10 kc	210	230
Wire Size	No. 22	No. 20
dc Resistance (ohm)	0.675	0.292
Inductance Tolerance (percent)	5	9

of turns of wire, which allows larger wire to be used and in turn yields a considerably lower dc resistance. The high permeability core also gives a higher Q at maximum frequency and a greater inductance change.

A VERTICAL TEST RANGE FOR ANTENNA RADIATION MEASUREMENTS*

JOHN STECKEL AND WILLIAM KORVIN

In order to facilitate the measuring of satellite antenna radiation patterns under a controlled RF environment the concept of a "Vertical RF Test Range" was devised. The vertical test range configuration is shown in Figure 1.

Of primary importance was the test range's capability to adequately absorb and/or suppress undesirable electromagnetic energy as low as 125 Mc which would normally be reflected from the chamber walls and floor. The satellite antennas of concern are of the broad-pattern near-omnidirectional type (i.e., dipole, turnstile, etc.).

The preliminary study, investigation, and 1/9-scale chamber model measurements indicated that a feasible reflection coefficient level of the order of 20 db to 25 db can be expected in an electrically small chamber of approximately $2\lambda \times 2\lambda \times 1\lambda$ in size. Subsequent full-scale measurements in the completed chamber have verified these reflection coefficient levels.

At higher frequencies of operation it was anticipated (and also indicated by 1/9-scale chamber model measurements) that the reflected energy level would be lower. The dominant controlling factor at higher frequencies is the structural weather protective radome. At the lower frequency the radome appears as a very thin wall (approximately 0.05λ thick). But as the frequency is increased, the presence of the radome becomes increasingly significant.

It should be remembered that an anechoic chamber is by no means free space. Therefore, it is of utmost importance that the RF anechoic chamber be "calibrated" not only as a function of frequency, but also as a function of illuminating or source antenna characteristics. Then, indeed, the confidence level at which the anechoic chamber may be used is defined.

I. INTRODUCTION

A. Antenna Test Range

The most important feature of an antenna test range is the control and reduction of reflections to the extent that they do not introduce significant errors in the measurement of the system under test. In order to achieve this, most test range sites are chosen in open fields clear of reflecting objects. Satisfactory operation is obtained as long as the system under test is sufficiently directive to discriminate against ground reflections and reflections from the test tower. Antennas for small scientific spacecraft are operated in the UHF region and designed to be as nearly omnidirectional as possible and thus are a "worst case" for conventional test ranges. The ideal solution is to provide a reflectionless environment for the

system under test that still permits convenient operation.

Recent advances in r-f absorbent material has improved the performance of r-f anechoic chambers to where they are comparable to the best outdoor range sites.

B. Definition of R-f Anechoic Chamber

A r-f anechoic chamber may be defined as an enclosure, suitably lined with an electromagnetic energy suppressing (absorbing and scattering) material, which may be used to measure such electrical characteristics as impedance, radiation pattern, and back scatter of a body within the chamber. Further, the degree of suppression of reflected undesirable electromagnetic energy shall be such as to approximate free space. The allowable divergence from free space is a function of the type of test and the tolerable error which the designer places on the test. For instance, a

*Published as *Goddard Space Flight Center Document X-632-64-243*, July 1964.

higher level of reflections may be tolerable when measuring impedance than when attempting back scattering measurements.

C. Testing Criteria as a Function of Type of Test: Impedance, Back Scatter, and Radiation Pattern

A primary requirement for measuring "free space" impedance of an antenna is the exclusion of all external, foreign sources of reflection. This requirement at first thought appears extremely stringent, if not impossible. To meet the requirement explicitly would mean that an impedance measurement would have to be made at distances miles from the surface of the earth and from within the antenna system. Fortunately, the requirement can be relaxed to a degree governed by the tolerable error permitted in the system impedance expressed as a Voltage Standing Wave Ratio. Consider a perfectly matched antenna system radiating energy. Under a condition of no electrical mismatch from external sources there would be no standing wave and VSWR=1.00:1. In practice, all antenna systems exhibit some mismatch from reflection resulting from electrical discontinuities. The vector addition of the incident and reflected voltages results in a standing wave, the magnitude of which is expressed in the familiar form:

$$\text{VSWR}_{\text{db}} = 20 \log_{10} \frac{E_1}{E_2}$$

when

E_1 = incident voltage

E_2 = reflected voltage

This VSWR is an indication of the power transfer of an antenna system. For instance, a system with a VSWR of 6.0:1 would reflect $\frac{1}{2}$ of the available power which is intolerable in most applications. Below is a brief table indicating the percentage of r-f power transmitted as a function of system VSWR.

VSWR	R-f Power Transmitted
1.00:1	100 %
1.04:1	99.6 %
1.22:1	99.02%

Since a perfect impedance match is seldom achieved the question then becomes how large a reflection due to the environment of the test site is tolerable while still permitting acceptable measurement of reasonable accuracy of the system impedance. Consider first, a worst case condition, i.e., all the incident energy striking a reflecting surface is reflected back to the antenna under test. If the power radiated P_0 from a perfectly matched antenna of unity gain, travels a distance R to a reflecting wall and is completely reflected back to the antenna a standing wave will result. However, the inverse square law for the attenuation of electromagnetic energy will reduce the amplitude of the reflected wave such that the SWR induced will be less than infinite. For example, the attenuation of a 136 Mc signal to a reflecting wall 8 feet away can be computed from:

$$\alpha = 37 + 20 \log f + 20 \log R$$

where

α = attenuation (db)

f = frequency (mc)

R = distance (feet)

then

$$\alpha = 23.53 \text{ db}$$

Assuming perfect reflection from the wall, the signal will be further reduced by 6 db in the return path to the antenna. Thus, the returned signal will be about 29.5 db below the incident signal. The resulting mismatch to the perfectly matched antenna is computed from:

$$|\Gamma| = \frac{r-1}{r+1}$$

where

$$|\Gamma| = \text{reflection coefficient} = 29.53 \text{ db} \\ = .0335$$

$$\therefore \text{VSWR} = r = 1.07:1$$

From the analysis above it is concluded that the impedance of a low gain antenna system may be measured with confidence if the walls of the enclosure are at least eight feet away and are less effective as a reflector than a metallic wall. Thus, the requirement on the absorption of unwanted

reflections from the environment is easily met when making impedance or VSWR measurements on low gain antenna systems.

By definition, σ , the scattering cross section of a body is the ratio of the power scattered per unit solid angle to the power incident per unit area. In terms of the radar equation,

$$\sigma = \frac{P_R 16\pi^2 R^4}{P_T G_T G_R \lambda^2} \text{ square meters,}$$

where

λ = the wavelength in meters.

R = the distance between the source antenna and the scattering cross section body in meters.

P_T = the transmitted RF power.

P_R = the received (reflected) RF power.

G_T = the transmitting antenna power gain with respect to an isotropic radiator.*

G_R = the receiving antenna power gain with respect to isotropic radiator.*

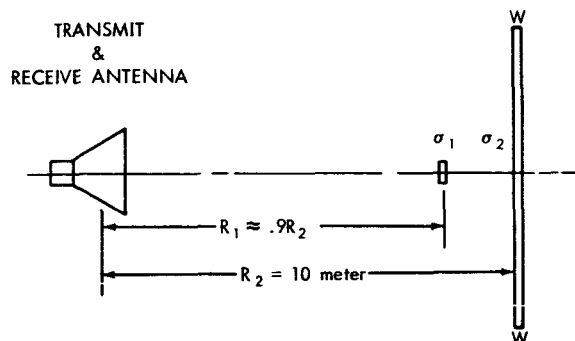
The problem which presents itself in measuring scattering cross sections in an anechoic chamber is the suppression of the energy from chamber walls reflected back to a receiving antenna. The necessary chamber performance, as in the case of VSWR measurements in a chamber, is relative. For example, a large scattering cross section would reflect an energy level which could be many orders of magnitude greater than energy contributed by the reflected energy from the chamber walls (including ceiling). In this case a percentage tolerance error can be established. Unfortunately, many radar cross section measurements are made of objects which have a small scattering cross section. Now, although the walls of the chamber absorb and suppress the RF energy, the fact that the walls are many orders of magnitude in size larger than the object being measured results in a chamber wall reflection coefficient which is as large as or larger than the object to be measured. In other words, the scattering energy of the body

to be measured is hidden by the reflections from the chamber.

There is one technique which is described by Elery F. Buckley (Emerson and Cuming, Inc.) which allows a calibration of a chamber for measuring radar cross sections in a controlled manner.

Very briefly, a transmitting and receiving antenna are placed at one end of a chamber. Two or more conducting spheres of known physical size (and hence known radar cross section) are placed at the opposite end of the chamber one at a time. The spheres, one at a time in the chamber, are rotated eccentrically about an axis producing an in-to-out of phase response. There will be two amplitude components for each of the different size spheres which contribute to the recorded interference pattern. One of the amplitude components is from the sphere; the other is the constant field equal to the energy return from the chamber plus electric field from transmitter-receiver cross coupling. The voltage ratio of these magnitudes produce an ambiguity. That is to say, the chamber scattering cross section will be either of two values, hence, the second measurement is made using a different conducting sphere with a known σ . The second measurement results again in two values of the chamber cross section. The true chamber reflecting cross section is that cross section which is common to both of the measurements.

An example which illustrates the magnitude of energy suppression which an anechoic chamber must exhibit for the measurement of radar cross sections follows:



Let the scattering cross section be that produced by a sphere of radius $\alpha = 0.1$ meter when the wavelength $\lambda = 3\text{cm}$ (i.e., $\alpha/\lambda > 1$). Then the

*An isotropic radiator is a fictitious antenna, used as a reference which radiates energy truly omnidirectional. Further, assuming this antenna is radiating an RF power of one watt, the magnitude of the electric field strength measured at a radius of 1 mile from the source is equal to 3.40mw/meter.

scattering cross section $\sigma_1 = \pi\alpha^2 = 3.14 \times 10^{-2}$ meter².

Now, the return power received by an antenna G_R from a scattering body of cross section σ when the transmitted power P_T is radiated by an antenna G_T is,

$$P_R = \frac{P_T G_R G_T \lambda^2}{16\pi^2 R^4} \sigma$$

Consider first the return power received from the spherical scattering body and letting $P_T = 1$ watt, $R_1 = 9$ meters,

$$G_T = G_R = 63 \text{ and } \lambda = 3 \text{ cm.}$$

$$\frac{P}{P_{R1}} = \frac{P_T G_R G_T \lambda^2}{16\pi^2 R^4} \sigma_1$$

$$P_{R1} = .109 \times 10^{-6} \text{ watts}$$

Next, assume the wall w-w in the figure on Page 551 to be a perfectly reflecting surface 3 meters on a side so that the wall radar cross section is

$$\sigma_2 = \frac{4\pi A^2}{\lambda^2}$$

where A = area of wall and λ = wavelength at the operating frequency. Then,

$$\sigma_2 = \frac{4\pi A^2}{\lambda^2}$$

$$\sigma_2 = 12.56 \times 10^4 M^2$$

The return power received from the back wall w-w is (the side, top and bottom walls may be neglected in this case since the predominant contribution to erroneous back scatter measurements is the back wall), and, $P_{R_2} = .284$ watts. By comparison $P_{R_1}/P_{R_2} = -64.2$ db

Or, the energy normal to the back wall W-W must be suppressed by 64.2 db to allow the spherical body (.1 meter radius) to present an equal amplitude to the receiver antenna G_R . Of course, further suppression of the back wall energy must be attained before the energy from the spherical body is discernible from (i.e., above) the back wall scattered energy.

The example outlined above now allows a feeling for the nature and magnitude of energy suppression and/or absorbing characteristics that may be required of an RF anechoic chamber designed to measure radar cross sections.

Measurement of antenna radiation patterns generally requires a better anechoic chamber than one for the measurement of antenna impedance. But the requirements are not as critical as for radar cross section (back scatter) measurements.

As an illustration, let us examine a typical E-plane pattern of a horn antenna and then show the effects of various magnitudes of reflected energy interfering at various aspect angles in the chamber. Figure 2 shows the E-plane pattern as measured in a controlled environment at an outside antenna range. The pattern approaches the theoretical pattern for a horn antenna exhibiting uniform distribution across the aperture. Figure 2 also shows a ray outline of reflected energy when measuring the E-plane pattern of this antenna in an anechoic chamber. Let sources of reflection be located at angles defined by $\theta = 50^\circ$ and 60° . Then the resulting perturbations from these reflecting sources are shown in Figure 2 as dashed lines. The deviations in the pattern are identified as ①, ②, and ③. The magnitude of the perturbation at 1 is 1.6 db. The magnitude of the reflected energy from point I which produces this 1.6 db is equal to 21 db since the reflection coefficient in db = $20 \log (S-1/S+1)$, where S = standing wave ratio. Referring to the figure, it can be seen that the perturbations take place at -14.2 db on the radiation pattern. Therefore, the actual reflection coefficient (magnitude of the interfering reflected energy) is equal to 35.2 db. At point 2 the deviation from point II is 1.1 db or a reflection coefficient of 25 db. Since the radiated energy is down 15.5 db at this point, the actual reflection coefficient is 40.5 db. Now consider point ③ (i.e., 0 db down on the pattern). A reflection coefficient of 40 db will cause a 0.175 perturbation in the pattern at this point.

From the above exercise one can now see clearly the effect of reflected energy on a typical pattern from arbitrarily chosen points in an anechoic chamber enclosure which is 35 db to 40 db down from the incident energy.

II. DESCRIPTION OF VERTICAL RF TEST RANGE

A. Physical Description

The vertical test range is as shown in Figure 1. The walls and floor of the structure are reinforced concrete. Attached to the chamber through a

common wall is the control room which houses all the electrical measuring devices. Opposite the control room side are two doors which when opened allow the chamber to be used as one end of a horizontal antenna range in conjunction with available antenna towers. The roof of the

chamber is an A-sandwich type RF transparent radome. Also, shown in this figure is the outside azimuth-elevation mount and fiberglas mast approximately 35 feet above the chamber controlled from inside. This facilitates the changing of source antennas. Inside the chamber (see Figure 3) is a wall-mounted antenna mount capable of rotating and allowing measurement of a satellite antenna system at or near the center of the chamber quiet zone.

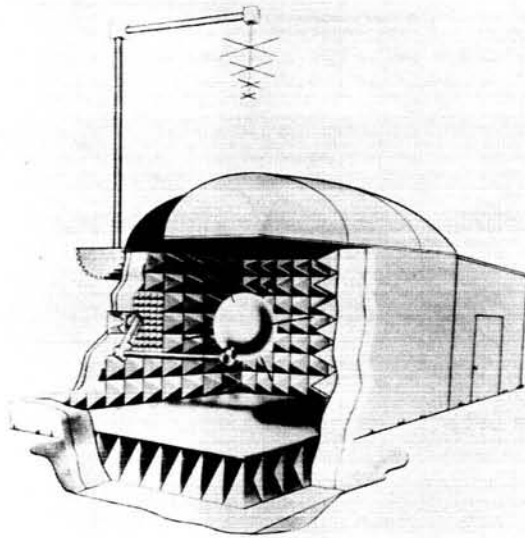


FIGURE 1.—Vertical anechoic chamber.

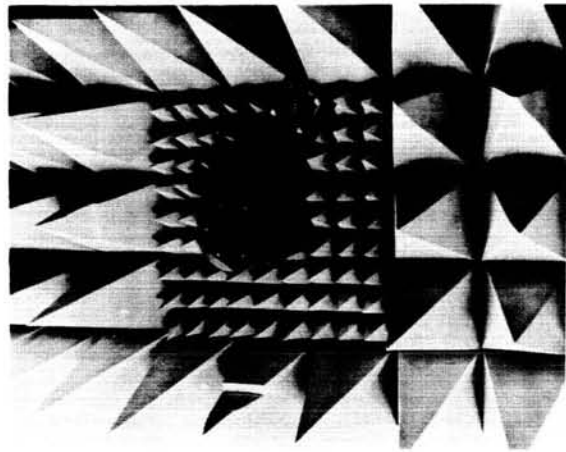


FIGURE 3.—Photograph of wall mount inside chamber.

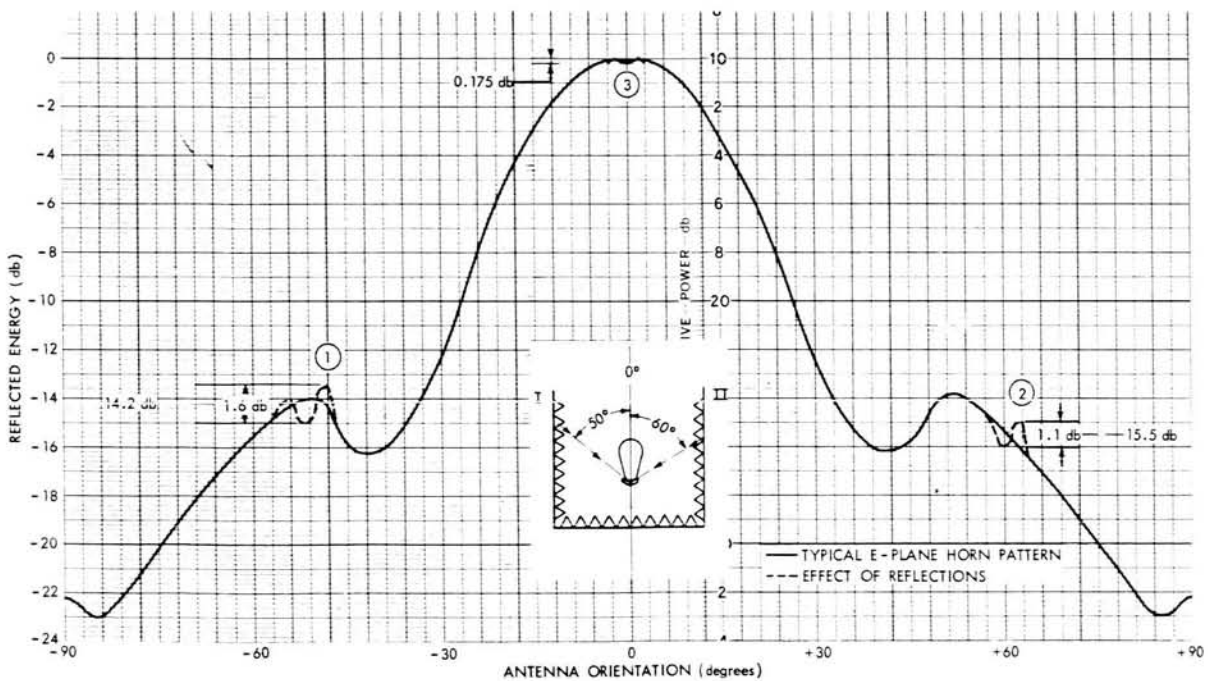


FIGURE 2.—Effect of reflected energy on typical E-plane pattern of horn antenna.



FIGURE 4.—Photograph illustrating size of absorbent material.

The walls are lined with pyramidal absorbent material 70 inches in length on 2 feet square bases (Figure 4). In the area of the side wall mount smaller pyramidal absorbent material is used allowing the rotation of an off-set arm containing the fiberglass mast. The floor of the chamber is lined with foam structure 70 inch pyramids on top of which is cemented a smooth floor—decking material of $\frac{1}{2}$ " thick sheets of semi-rigid vinyl foam which can be walked upon in setting up experiments within the chamber.

The radome is a structural body of the A-sandwich type. It is approximately 4" thick and the 9 sub-pieces are assembled into a continuous weather protective roof (Figure 5).

Uniform chamber lighting is achieved by directing four flood lamps toward the radome. The four lamps are located at the four upper corners of the chamber such that they are well hidden by the absorber on the walls. This technique of lighting results in two very favorable conditions; the lighting is uniform and avoids the visual problem of looking into the lamps; a minimum of RF reflec-

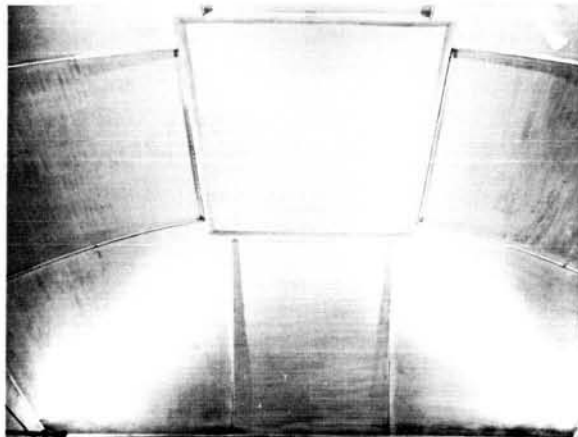


FIGURE 5.—Photograph of 9 panel radome.

tion from the fixtures is obtained since they are well shadowed by the 70" absorbent material.

Finally, the chamber is temperature controlled to prevent large variations of temperature which may effect antenna measurements.

B. Electrical Description

Electrically the chamber was lined on 5 surfaces with an RF absorbent material and the 6th side enclosed by an RF transparent (radome) roof.

The RF absorbent material was supplied by the B. F. Goodrich Company, Shelton, Connecticut. Each individual piece of the material, VHP-70 is a 70" pyramid shaped absorber on a 2 feet square base. All of the absorbent material exhibits a minimum reflection coefficient of 28 db at 120Mc, 40 db at 400Mc and 50 db at 1,000 to 10,000Mc.

The tests on the absorber at the manufacturer's plant were made in a closed loop technique at 120Mc and 400Mc. At 1,000, 5,000 and 10,000 Mc an open loop technique was used. The closed loop technique consisted of measuring four 2' x 2' absorber pieces at a time in a flared waveguide system in which the absorbent material serves as a load under test. The absorber is moved inside the waveguide resulting in a standing wave that moves in conjunction with the physical movement of the absorber. A fixed probe inside the waveguide located between the absorber and RF oscillator detects the standing wave which is converted to power reflected.

The open loop technique of testing consisted of a horizontal version of the NRL type arch

method. A reference was used which consisted of a flat 2'×8' metal plate. The reflection level of four 2×2 pieces was measured and compared to the flat metal plate. The reflectivity of the absorbent material is the difference in return power in db between the metal plate and the absorbent material.

The radome was supplied by Raymond Development Industries, Inc., Huntington Park, California. The electrical characteristics of the A-sandwich radome material were specified as:

- (a) Transmission loss ≤ 7.5 per cent over the frequency range of 120Mc to 10,000Mc.
- (b) Refraction less than 10 percent for angles of incidence from 0° to 45°.
- (c) Loss tangent of glass laminate = .005.
- (d) Dielectric constant of glass laminate = 4.0.
- (e) Loss tangent of foam core = .0005.
- (f) Dielectric constant of foam core = 1.12.

Although tests at the manufacturers' plant confirm the meeting of the specifications on a flat 40"×40" sample of the radome, measurements indicate that the actual radome does exhibit less than specified performance particularly at frequencies above 1,000Mc. This can be easily explained since the actual completed radome possesses a curved surface and the nine pieces comprising the total radome are assembled with beefed-up glass laminate flange sections. These two conditions increase the detrimental effects (increased diffraction and reflection).

As previously stated the frequencies of primary interest are in the 120Mc to 400Mc region. But some tests were performed in the chamber at higher frequencies.

Normally, in evaluating the chamber a quiet zone is defined. That is, a volume within the chamber in which known (measured) reflectivity levels exist. Then in this zone antenna systems can be evaluated being fully aware of the limitation of the chamber. Therefore, by definition, the quiet zone is a 10 foot diameter sphere which is tangent to the chamber floor and centered elsewhere within this chamber.

C. Advantages

Ground reflections are the most serious source of error when using a conventional test range for measurements of low frequency, low-gain antennas. The error may be reduced to a degree by

additional height to the towers, but towers over 100' high are expensive and rather inconvenient to use. Likewise a conventional anechoic chamber designed for operation at low frequency is expensive and requires a large building to provide even a modest size test range. The vertical test range provides an attractive compromise between tall towers and a large anechoic chamber. The chamber portion need only be large enough to prevent reflections from the ground and nearby reflecting surfaces and the tower need only be tall enough to hold an antenna out of the near field of the antenna under test. Furthermore, the length of the test range may easily be varied by adjusting the height of the outside antenna. Operation of a vertical range is especially convenient. The model under test, the test and control equipment, and personnel are all at ground level and thus avoid the need for hoists and elevators.

III. ANALYSIS OF CHAMBER

The magnitude of unwanted reflections that can be tolerated in an antenna test range have been shown to be a function of the parameter being measured. Since the site is never perfect, the results obtained may be interpreted in terms of the known site imperfections provided the reflection coefficients and in some cases phase are accurately specified. However, measurement of reflection coefficient of absorbent material is difficult and techniques for its evaluation have not been standardized. Currently the *Pattern Comparison Technique*¹ and the *Free Space VSWR Technique*² are favored in evaluation of r-f anechoic chambers. In both techniques the significant result is the comparison of the incident to the reflected energy from an absorbing wall to determine its reflection coefficient.

Briefly, in the Pattern Comparison Technique, the pattern of a directive antenna (15 to 20 db gain) is measured successively at closely spaced points along the radii of the chamber quiet zone. The quiet zone may be defined as the volume within an anechoic chamber in which an antenna under test will be measured. Then the patterns are superimposed on each other with pattern peaks coincident. The deviations in the patterns are read at different aspect angles and VSWR curves vs aspect angle are constructed. The

curves may then be converted to reflection levels within the chamber.

The Free Space VSWR Technique is a method of continuously recording the amplitude variations produced by reflections. Two directive antennas are used, with one being moved continuously across the chamber quiet zone at a discrete aspect angle for each recording. The amplitudes recorded are reduced to incident and reflected energy levels, thus allowing the reflection coefficient vs aspect angle of a chamber to be determined.

From the brief summary above (and more so from the referenced literature) it can be seen that both techniques rely on use of directive antennas. This is not a disadvantage in the usual situation where the chamber is large in terms of wavelength and directive antennas are usually evaluated. However, in the vertical test range more emphasis is placed on measurement of nearly omnidirectional antennas in a termination chamber that is electrically small, i.e., approximately 1λ in depth and 2λ on a side. Although evaluation of chamber performance using directive antennas will indicate both a direction and reflection coefficient for sources of reflections, unless considerable effort is made to integrate the reflection levels from all directions the chamber will appear better than when used with an omnidirectional antenna. Therefore most of the analysis of both the scale model and the full scale range was made using a dipole antenna to probe the energy levels within the termination chamber. In this method, a dipole antenna is moved throughout the quiet zone and energy levels vs position recorded. The difference between the levels recorded and calculated free space levels are converted into VSWR and finally a reflection coefficient computed.

It is common practice to try to reduce the large number of different reflection coefficients that are measured in evaluating a chamber to a single number that is then used to define the chamber's performance. In general, it will be found that this number is not really a common denominator and that the relative performance of two different chambers should not be judged on this alone. For instance the quiet zone of chamber A may be only $\frac{1}{3}$ the volume that was included in evaluation of chamber B, yet chamber B may be considerably better than A over the same quiet zone.

Unfortunately, the performance rating of anechoic chambers is like the evaluation of radio receivers in 1940, i.e., not complete unless the test conditions are known as well as the results.

A. Scaled Chamber Measurements

Reflection coefficient measurements were made in a 1/9th-scale chamber. A photograph of the scale chamber is shown in Figure 6. Figure 7 defines the test set-up. No attempt was made to design and test a 1/9th-scale radome because the radome would appear as a thin wall structure in

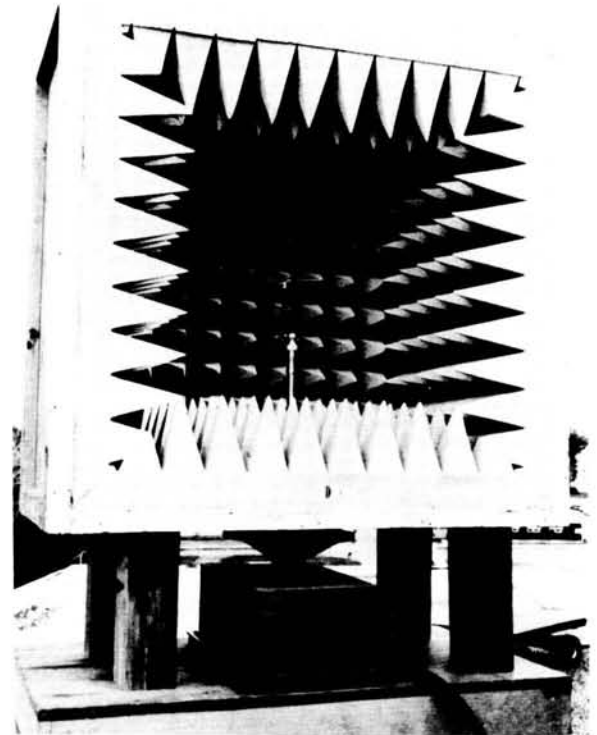


FIGURE 6.—Photograph of 1/9th scale chamber.

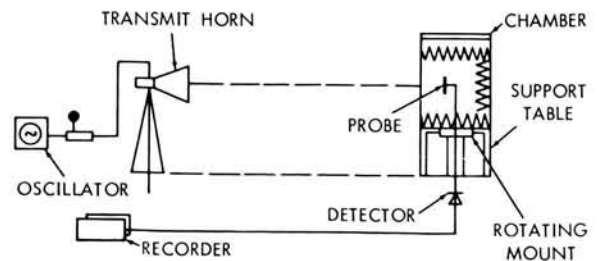


Figure 8

FIGURE 7.—Equipment set up for 1/9th scale measurements.

the frequencies of most interest (up to around 400Mc) in the full scale chamber. Although no scaled radome measurements were made in this particular chamber design it is felt that general comments on scaled radome measurements are in order.

Scaling a multi-panel, sandwich radome is, in general, difficult. Extreme care must be taken in scaling the ribbing, flange design and radome curvature. Serious errors can be expected if this precaution is not taken. Thin wall structural radomes of dielectric constants of approximately 3 or 4 may be scaled for measurements with confidence and also thick wall foam radomes of low dielectric constants ($\epsilon_r = 1.1$ to 1.4) are practical for scaling purposes.

Measurements were made at frequencies from near 1,080Mc to 3,600Mc. This corresponds to full scale chamber measurement of approximately 120Mc to 400Mc. It is to be noted that a dipole probe is used inside the scaled chamber to deliberately prevent discrimination against any reflected energy. A directive type antenna would have been too selective.

The measurements at 1,080 Mc in the 1/9th-scale chamber indicated a reflection coefficient of 20 db and at 3,600Mc 30 db reflection coefficients were measured.

Figure 8 is typical of the standing waves measured by probing the scaled chamber with dipoles. The dipole probes were moved in small increments in terms of wavelengths from the aperture of the

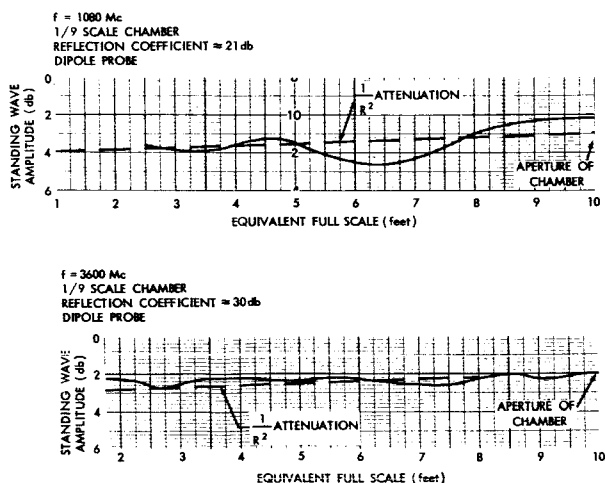


FIGURE 8.—Standing waves in 1/9th scale chamber measured from aperture of chamber floor.

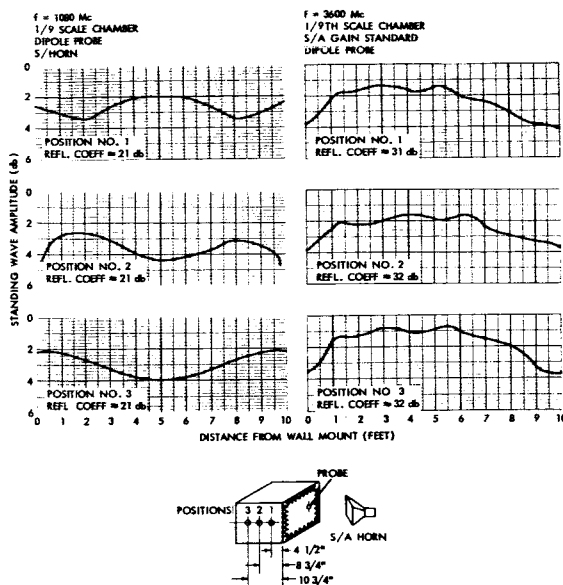


FIGURE 9.—Standing wave curves measured across 1/9th scale chamber aperture.

chamber to the back. Figure 9 depicts standing wave curves resulting from probing across the chamber in small increments. This data is typical of that obtained in the 1/9th-scale chamber.

It was from these measurements that the full-scale chamber performance was predicted.

As was previously mentioned no attempt was made to determine the effect of an A-sandwich type radome on the full scale chamber at scaled frequencies. That the radome effect would be negligible at the low frequencies of interest was predicted on the basis that the radome would appear as a thin wall (approximately $\lambda/8$ thick or less) at 400Mc and lower.

B. Full Scale Chamber Measurements

The quiet zone of the chamber was probed with a dipole using an illuminating antenna which was directional in nature.

Specifically, at 125Mc and 400Mc the magnitude of the reflection coefficient was determined by moving a dipole horizontally inside the chamber in directions normal to each wall and across the diagonals. With the dipole fixed to a cart made of low dielectric constant foam material, the cart was moved along tracks which could be oriented as desired within the chamber. The

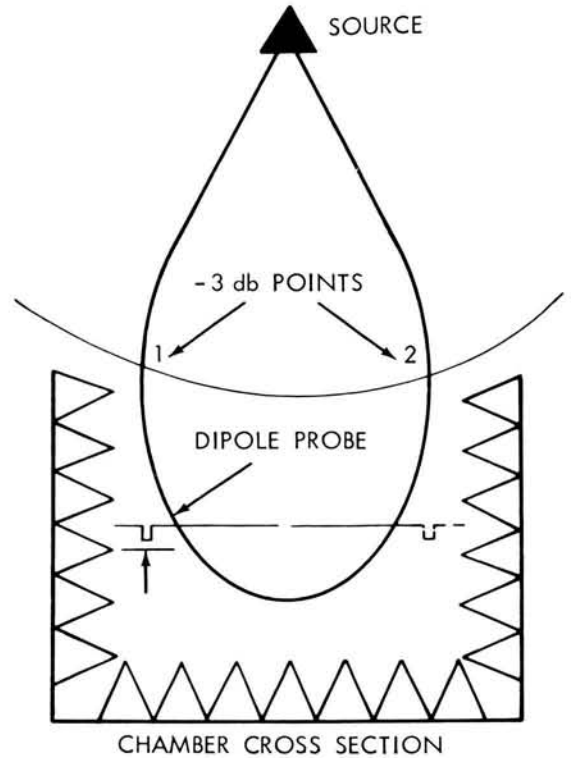


FIGURE 10.—Photograph of track and cart arrangement.

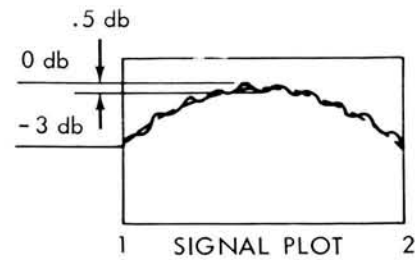
track, cart and dipole support are shown in Figure 10. These horizontal measurements were made at discrete heights of 2 (and/or 3), 4, 5, 6 and 8 feet above the floor of the chamber. The reflection coefficient in the vertical direction was determined by measuring the standing wave as a function of vertical movement of the dipole from 2 to 10 feet above the floor. Resulting standing wave curves are converted to an equivalent reflected energy level (reflection coefficient). Figure 11 is a typical curve. The dashed curve of Figure 11(b) represents the probed energy level as measured in Figure 11(a) in the absence of reflected energy. The solid curve superimposed on the dashed curve represents the effect of the reflected energy. Peak to peak value of the standing wave is 0.5 db. This standing wave of 0.5 db results from a reflected energy level of 30 db below the incident signal level. Or, Reflection Coefficient = $20 \log \frac{VSWR-1}{VSWR+1}$. Measured curves are shown in Figures 12(a) through 12(d).

At 1,200Mc reflected energy levels greater than expected were measured (approximately 32 db average). Also, the reflected energy levels were polarization sensitive (i.e., E-perpendicular vs E-parallel) with differences in reflection coefficients between 6 and 15 db as a function of aspect angle. This obvious deviation from the 45 to 50 db levels anticipated was attributed to the radome for two reasons:

- (1) The chamber at 1,200Mc is electrically large ($15\lambda \times 15\lambda$ 10λ) and therefore the



(a)



(b)

FIGURE 11.—Interpretation of standing wave curves.

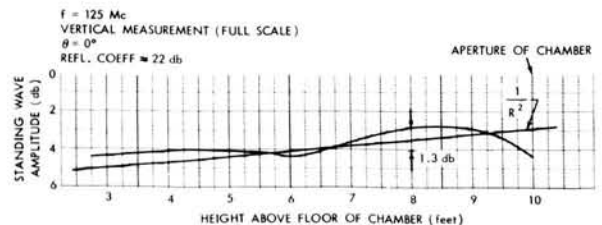


FIGURE 12a.—Full scale chamber standing waves at 125 Mc, 400 Mc.

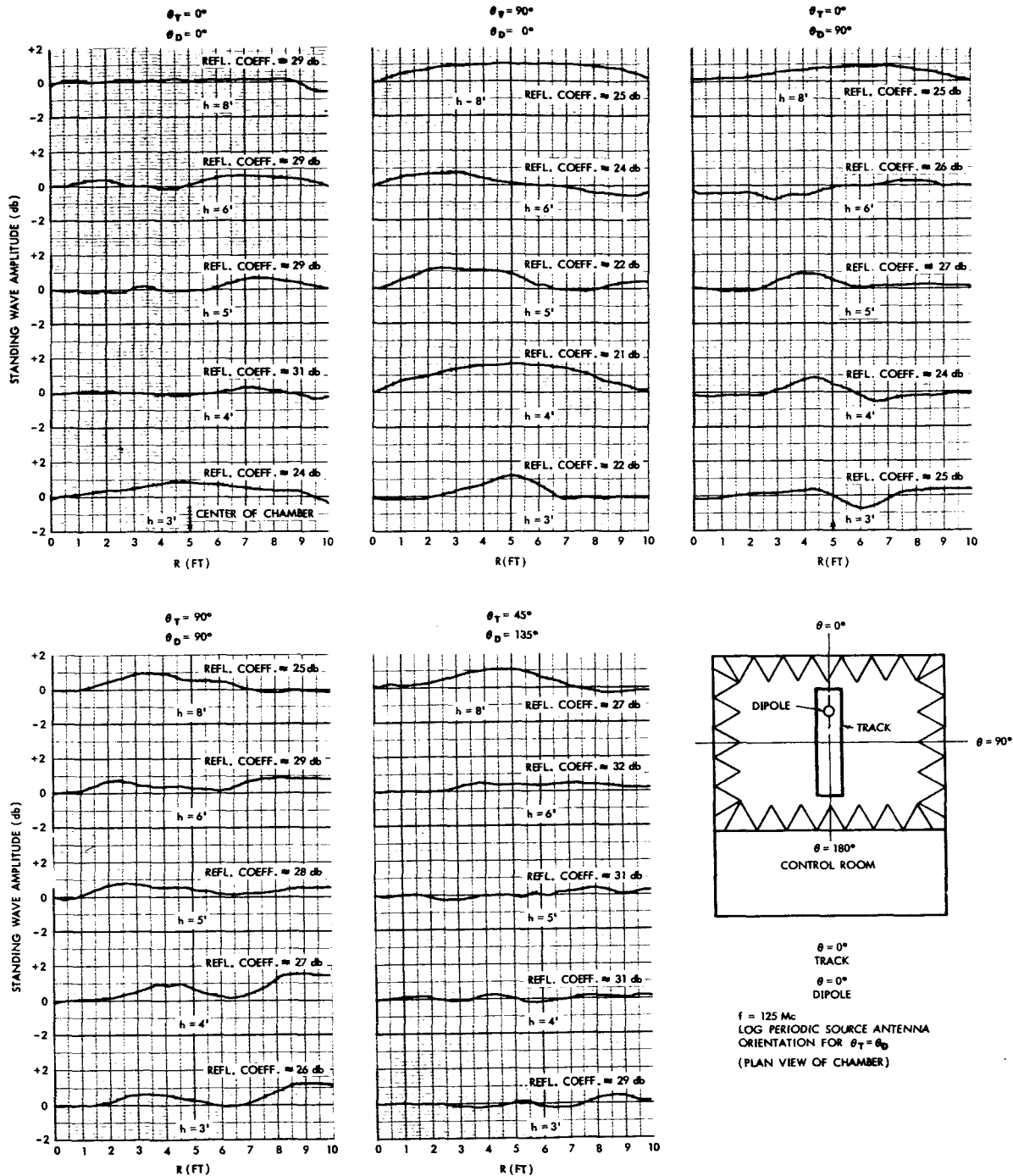


FIGURE 12b.—Full scale chamber standing waves at 125 Mc, 400 Mc.

chamber would be expected to be much better than at 400Mc (reflection levels ≈ 32 db).

(2) This A-sandwich radome (4 inches thick,

with solid fiberglass panel flanges) appears as a relatively large discontinuity.

Measured magnitudes of reflected energy for various linear polarization orientations when

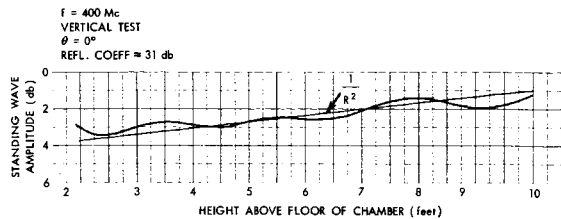


FIGURE 12c.—Full scale chamber standing waves at 125 Mc, 400 Mc.

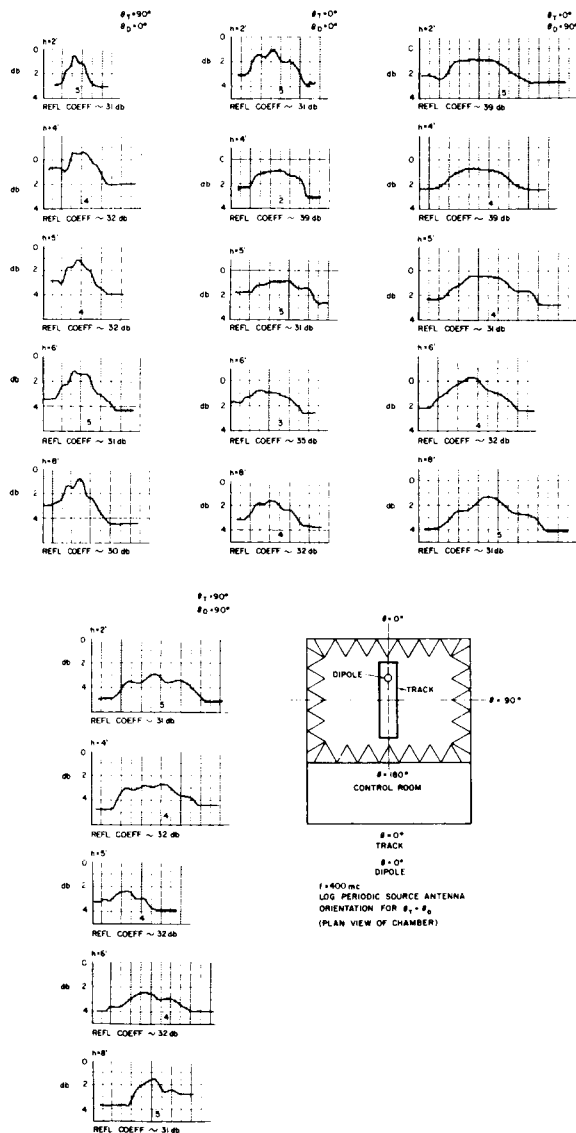


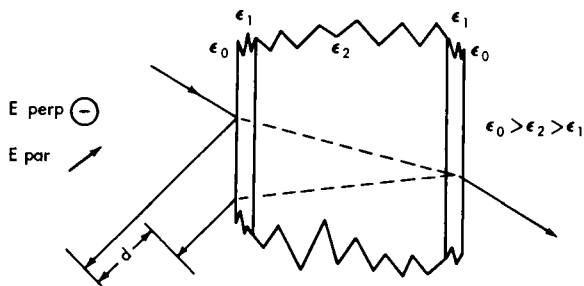
FIGURE 12d.—Full scale chamber standing waves at 125 Mc, 400 Mc.

probing the chamber at 125Mc and 400Mc are tabulated in Tables 1 and 2. Some inconsistency is apparent, especially in the higher coefficient values measured at 125Mc. This is attributed to instability of the measurement equipment; however, retests of selected orientations indicate no serious differences from the data presented or the conclusions drawn from the data.

Data taken at 1,200 Mc are not presented because later tests prove that the radome is the major contributor to standing waves within the chamber and therefore does not truly represent the capability of a vertical test range.

To verify the fact that the radome was the major contributor to the problem (i.e., larger reflection coefficients than anticipated and polarization sensitivity) measurements were performed with the source antenna inside the chamber just under the radome and compared to the pattern taken through the radome (Figure 13(a)). Thus, an interference pattern (see Figure 13(b)) between incident and reflected energy was produced by essentially removing the radome from between the source and receive antenna. In both cases a horn antenna was moved across the chamber at a 45° angle (from the vertical). The +45° refers to the horn antenna being tilted toward the chamber wall containing the wall mount. And -45° refers to an angle toward the wall without the wall mount.

Plots in Figure 14 compare the standing waves of the A-sandwich radome for the two polarizations E-perpendicular and E-parallel as defined below:



Except at angles far off normal incidence reflected energy from E-perpendicular is nearly always greater than from E-parallel because of the zero at the Brewster angle occurring only with parallel polarization. From Figure 14 it can be

TABLE 1.—125Mc Chamber Reflectivity Levels

Track Orientation θ_T	Dipole Orientation θ_D	Dipole Height Above Floor				
		3 ft.	4 ft.	5 ft.	6 ft.	8 ft.
0°-----	0°	24 db	31 db	29 db	29 db	29 db
0°-----	90°	25 db	24 db	27 db	26 db	25 db
90°-----	0°	*27 db	*26 db	*28 db	*24 db	*20 db
90°-----	90°	26 db	27 db	28 db	29 db	25 db
45°-----	45°	29 db	22 db	24 db	23 db	*25 db
45°-----	135°	29 db	31 db	31 db	32 db	27 db
135°-----	45°	26 db	21 db	21 db	21 db	22 db
135°-----	135°	*22 db	*21 db	*24 db	*24 db	*21 db

*Measurements performed at 123Mc using NASA battery operated signal generator.

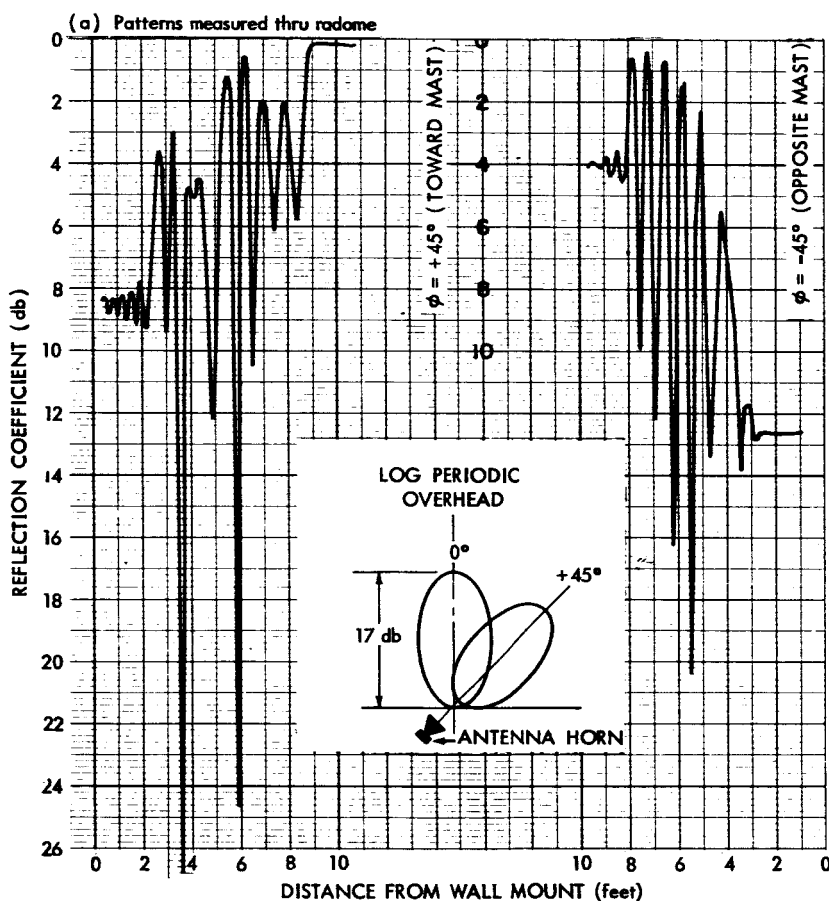


FIGURE 13a.—Radiation plots showing effect of radome on reflection coefficient at 1200 Mc.

seen that the difference in reflection coefficients is 14 db. The Brewster angle (where the reflection value goes to zero) is defined as follows:

$$\theta_B = \tan^{-1} \sqrt{\frac{\epsilon_2}{\epsilon_1}} \text{ where } \epsilon_2 > \epsilon_1$$

TABLE 2.—400Mc Chamber Reflectivity Levels

Track Orientation θ_T	Dipole Orientation θ_D	Dipole Heights Above Floor				
		2 ft.	4 ft.	5 ft.	6 ft.	8 ft.
0°-----	0°	31 db	39 db	31 db	35 db	32 db
0°-----	90°	39 db	39 db	31 db	32 db	31 db
90°-----	0°	31 db	32 db	32 db	31 db	30 db
90°-----	90°	31 db	32 db	32 db	32 db	31 db
45°-----	45°	31 db	39 db	39 db	34 db	32 db
45°-----	135°	32 db	31 db	30 db	31 db	32 db
135°-----	45°	35 db	32 db	31 db	35 db	30 db
135°-----	135°	32 db	32 db	30 db	31 db	39 db

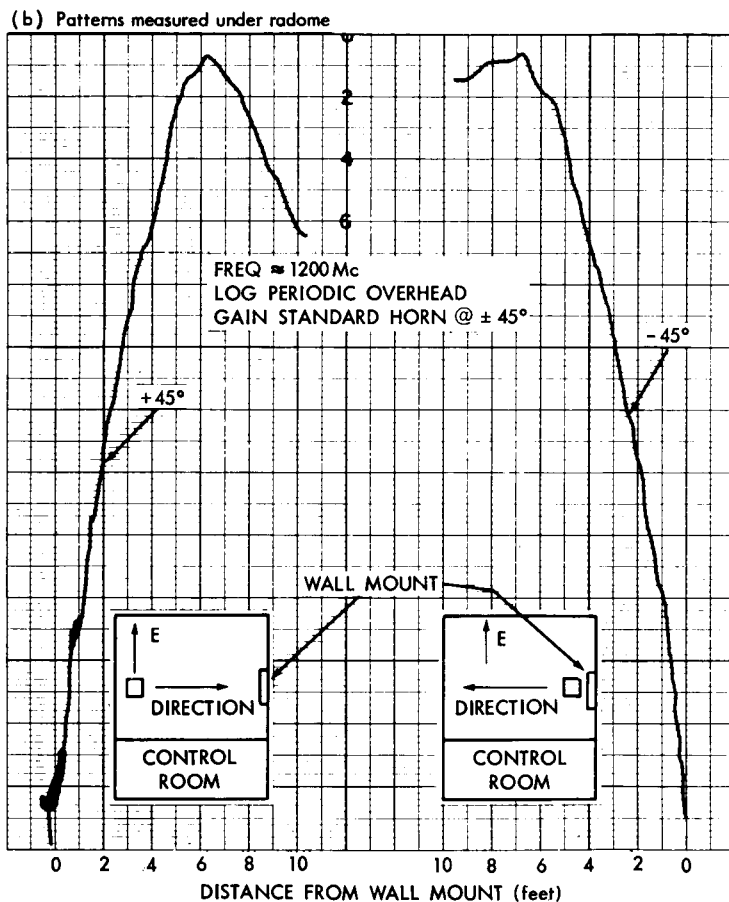


FIGURE 13b.—Radiation plots showing effect of radome on reflection coefficient at 1200 Mc.

An important point which is illustrated in Figure 15(a), (b) is the degree of accuracy with which the antenna radiation patterns can be measured through a radome which exhibits sig-

nificant reflection characteristics. In this case the radiation patterns of an antenna (gain standard at 1,200Mc) was measured first in a free space pattern range method and then the measurement

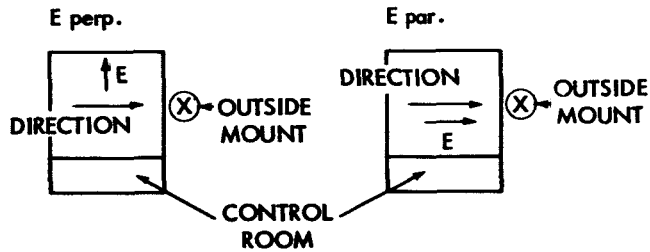
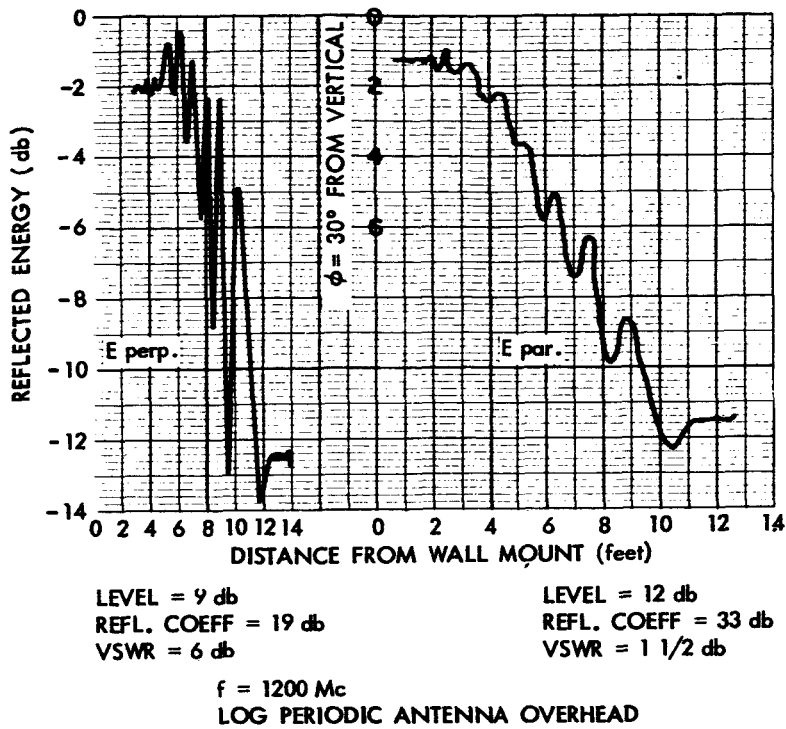


FIGURE 14.—Effect of radome on parallel and perpendicular polarization.

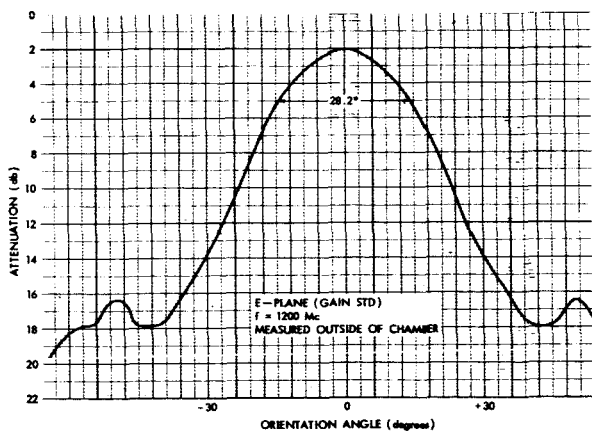


FIGURE 15a.—Pattern of horn antenna (E-plane) measured outside and inside of chamber.

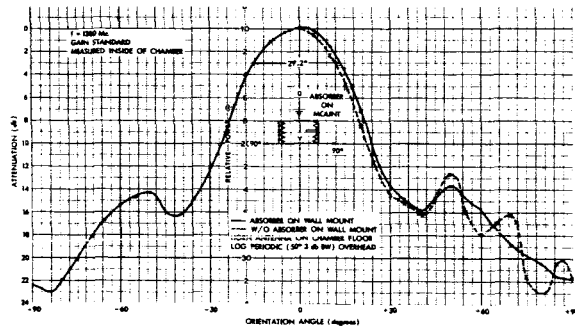


FIGURE 15b.—Pattern of horn antenna (E-plane) measured outside and inside of chamber.

was repeated inside the vertical chamber. To measure the pattern inside the chamber, the antenna was rotated about a fixed axis and not moved across the chamber. Therefore, the "available" reflections which exist across the chamber are not probed and therefore, a rather respectable pattern can be achieved. As a matter of interest Figure 15(b) shows the effect the wall mount can produce in terms of reflected energy.

IV. CONCLUSIONS

Measurements taken from the 1/9th-scale model of the test range are in reasonable agreement with the measurements of the full scale range. The structural S-sandwich radome definitely reduces the performance of the facility at higher frequencies, but does not affect operation in the frequency range of primary interest (125-400Mc).

The concept of a vertical test range composed of an electrically small termination chamber with a r-f transparent radome has been found feasible and provides at moderate cost a convenient, quasi-all-weather, facility for accurate measurement of antenna characteristics.

V. RECOMMENDATIONS

Considerable confidence may be placed on the results of measuring scale models of anechoic chambers. Since the instrumentation is not difficult and the cost is small it is recommended that more extensive use be made of scale models to check chamber and anechoic material performance.

A multipanel sandwich radome has definite frequency limitations when used as part of an antenna test range. Therefore it is recommended that a thick, low-dielectric-constant foam radome be used.

REFERENCES

1. BUCKLEY, E. F., "Outline of Evaluation Procedures for Microwave Anechoic Chamber," *Microwave Journal*, August, 1963.
2. EMERSON, W., "Chamber Information," unpublished report of B. F. Goodrich Company.
3. JASIK, "Antenna Engineering Handbook", New York: McGraw-Hill Book Company, 1961.
4. HARVEY, "Microwave Engineering", New York, Academic Press, 1963.
5. Electronic Space Structures Corporation, "Ground Radomes", 1964.

VI. ENVIRONMENTAL TESTING

SOLAR SIMULATION TESTING OF AN EARTH SATELLITE AT GODDARD SPACE FLIGHT CENTER*

R. E. BERNIER, R. H. HOFFMAN, A. R. TIMMINS, AND E. I. POWERS

INTRODUCTION

The use of solar simulation as a test for evaluating the thermal performance of a spacecraft is still relatively new and controversial. Reference (1) reported successful use of carbon arcs in testing the Telstar spacecraft. Additional information on the use of carbon arcs as a solar source should be useful in evaluating its effectiveness as a thermal design evaluation technique. At Goddard Space Flight Center carbon arcs have been used for achieving the solar simulation testing of spacecraft sized for the Delta and Scout boosters. This report presents data and experience from such testing, using the results from the UK-2/S-52 International Satellite† as an illustrative example.

OBJECTIVES OF THE TEST

The primary purpose of exposing a spacecraft to solar simulation testing is to verify the thermal design of the spacecraft while it is in full operation.

An additional objective is to check experiment operation in the vacuum chamber with simulated solar energy. Some experiments are directly stimulated by the sun, as in the case of the ozone measurement experiments on the UK-2/S-52. Others, such as the micrometeoroid detector, use sunlight as a secondary medium by measuring the amount of sunlight passing through the punctures in a foil. Spacecraft subsystems also use sunlight as an event marker, switching operational modes as a sunrise or sunset condition is encountered.

A benefit derived from a solar test of a spacecraft is the exposure of the spacecraft coatings and exposed surfaces to the thermal radiation environment encountered in orbit. In this way, possible physical incompatibilities may be discovered.

For these reasons, and the basic desire of environmental testing groups to demonstrate spacecraft performance under the simulated environment, a solar test of the UK-2/S-52 International Satellite was conducted.

THERMAL DESIGN AND PREDICTIONS

As previously stated, the primary purpose of conducting a solar test is to verify the thermal design of the spacecraft. Therefore, a brief discussion of the radiation inputs and the thermal model is presented so that a better appreciation of the test results may be possible.

For the UK-2/S-52, internal power dissipation is relatively small compared to the total radiation input, and does not have a significant effect on the satellite mean temperature. In general, the magnitude of this effect depends upon the emittance (ϵ) and the surface area (e.g. with a surface of low absolute ϵ , internal power may raise the temperature significantly because the skin has a limited capacity for reradiation.)

Direct solar heating, earth-reflected solar heating (albedo), and earth-emitted radiation (earthshine) represent the significant inputs to the satellite. It is apparent that an adequate thermal design is predicated upon a reasonably accurate knowledge of these thermal radiation inputs. The major source of energy, direct solar radiation, is, fortunately, the most accurately obtained. Since the sun's rays impinging upon the satellite are virtually parallel, the problem is simply one of determining the instantaneous

*Published as *NASA Technical Note D-2614*. This work was also presented at the Technical Meeting on Equipment Exposition at Philadelphia, Pennsylvania, April 13-15, 1964, and published in the *Proceedings of the Institute of Environmental Sciences*, April 1964, pages 209-216.

†This satellite is scheduled for launch during the first quarter of 1964.

orientation of each external face with respect to the solar vector.

Determination of the Thermal Model

Of prime importance in the thermal analysis of a spacecraft is the determination of the thermal model. The model, an approximate mathematical representation of the satellite, is composed of a number of isothermal nodes or areas. The selection of these nodes is governed partly by convenience in working around interfaces, accuracy requirements, and a desire to minimize engineering and computer time.

At first inspection of the UK-2/S-52, Figure 1, it was recognized that the broad-band ozone detector was essentially independent of the spacecraft itself. Therefore, separate thermal models were developed for the main spacecraft and broad-band ozone detector and thermally coupled by radiation and conduction interchange. The spacecraft was divided into 20 nodes and the broad-band ozone detector into 19 nodes.

The thermal model of the ozone detector is shown in Figure 2. The most critical elements of this experiment are the monitor cell and the thorium-coated glass-enclosed tube at the top, each of which must be maintained below 60°C.

The thermal model of the spacecraft is shown in Figure 3. Since the satellite is spinning about its longitudinal axis, skin temperatures tend to be uniform about the axis. This, together with the symmetrical design of the spacecraft, greatly simplifies the thermal considerations. As can be seen from Figure 1, there are a few components exposed to space. These are the ferrite rod antennae mounted in two fiberglass containers, the foil of the four micrometeorite experiments, and the four apertures of the ozone scanners. There are also several openings in the bottom dome around the boom and paddle arm mounts which expose certain internal elements to space.

Since elements exposed to space generally undergo significant fluctuations in temperature, three of the ten external nodes were assigned to the ferrite rods, micrometeoroid foil and the ozone scanner apertures. Of the ten internal nodes, five are structural elements and five relate to the experiments. The experiments on the upper shelf were considered as one node, since each is

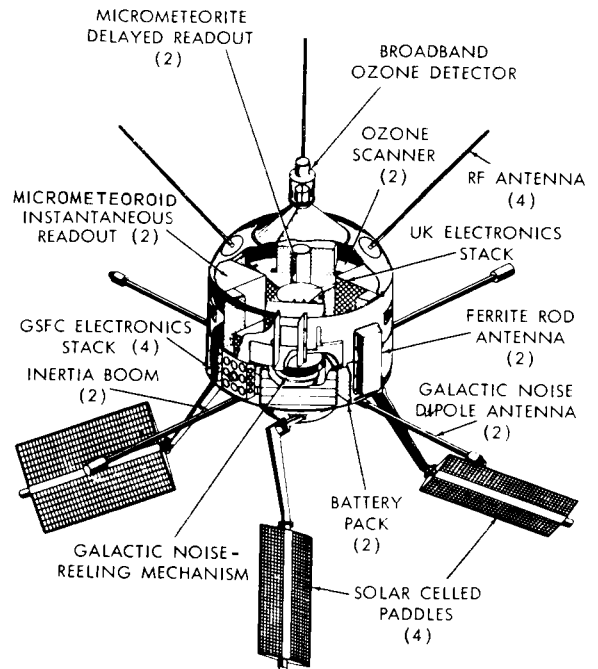


FIGURE 1.—International satellite UK-2/S-52.

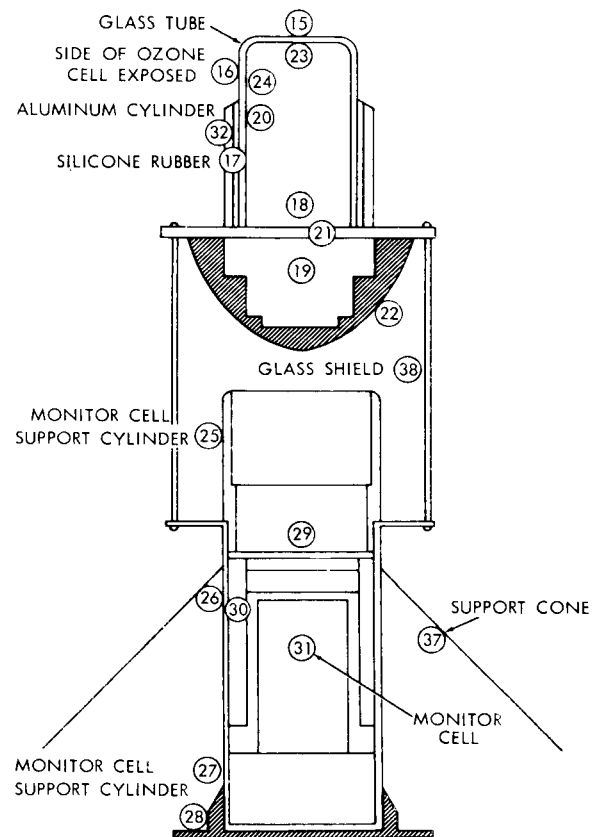


FIGURE 2.—Broad band ozone detector node locations.

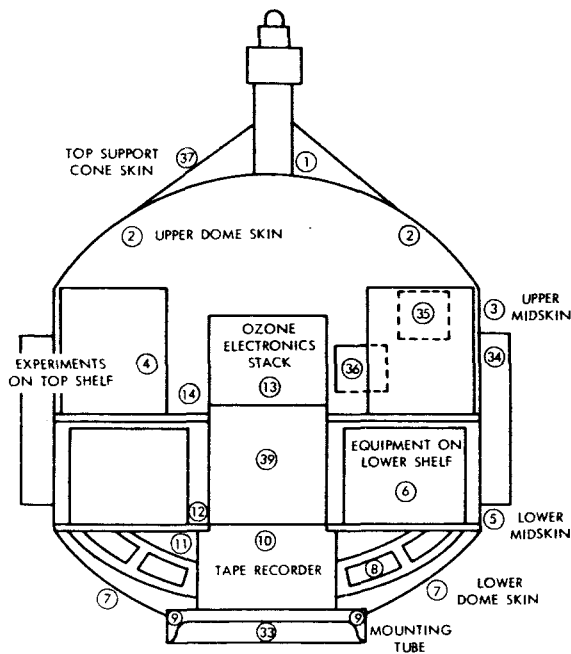


FIGURE 3.—Thermal model for UK-2/S-52, main body node locations.

similarly influenced by the temperatures of the upper dome and midskin while the power dissipated is negligible. The equipment on the lower shelf was also considered one node. The batteries, however, were investigated in further detail because of local hotspots while being charged. The three remaining internal nodes represent the ozone electronics stack, the galactic noise-reeling mechanism and the tape recorder.

Conduction and Radiation Interchange

Every node is thermally coupled to one or more nodes by conduction and/or radiation interchange. External nodes also radiate to space. To determine the radiative coupling, the shape factor area product and effective emittance between nodes were determined. The shape factor calculations were simplified by reducing the internal nodes to simple geometric forms (flat plates, cylinders, spheres, etc.) and employing sources such as Reference (2). Almost the entire interior of the spacecraft was painted black to minimize thermal gradients. The effective emittance values were, therefore, approximately equal to 1.

The conduction interchange presented a problem in some cases since there was no way of

accurately determining the conductance across joints. In these cases, the extremes were considered, assuming both nominal contact and virtual isolation of the two nodes to determine how large a gradient might exist. Ten percent of perfect contact area was usually used for a nominal value.

Solar Input

As mentioned earlier, the major source of heat input to an orbiting satellite is direct solar radiation. The solar radiation absorbed by an external surface is $SA_p\alpha$, where S is the solar constant, A_p is the projected area to sunlight, and α is the solar absorptance of the surface.

Determining the projected areas of each surface element or node for different positions around the spin axis and for various aspect angles was accomplished by taking pictures of a one-fifth scale model of the spacecraft. The satellite was designed to operate within solar aspect angles (angle between the solar vector and spin axis) of 45° to 135° . However, for a complete analytical study the projected areas were determined for all aspect angles, at 15° increments.

Albedo and Earthshine

A computer program (Reference (3)) was employed to determine the values of albedo and earthshine incident upon the rotating surfaces throughout the orbit. Average orbital values were used for the two orbits considered. Albedo and earthshine account for approximately 15 to 30 percent of the total external heat input, albedo being greater in the minimum sunlight orbit.

Modifications in Thermal Analysis

In altering the thermal analysis for use in the solar environment test chamber, the energy inputs of albedo and earthshine were equated to zero since no attempt was made in the test to simulate them. The effects of paddle shading at high aspect angles was also removed from the computer analysis since the test was conducted without the solar paddles fixed to the spacecraft.

Calibration tests were performed on samples of the spacecraft coatings to determine their absorptance properties when illuminated by the solar simulator in the test chamber. These properties were then inserted into the thermal model in

place of the values used for orbital predictions. It is noted here that the difference between the orbital absorptance properties and the carbon arc chamber values for the UK-2/S-52 coatings was negligible. However, the practice of using test-condition absorptance properties to predict test temperatures can be extremely important when a source with a poor spectral match is used on a coating whose absorptance response is not flat in the visible wavelength region.

The thermal model was used to predict spacecraft temperatures from carbon arc radiation intensity values which were introduced as input fluxes to the external nodes of the spacecraft. Agreement between predicted and actual test temperatures would then corroborate the thermal design of the spacecraft. Differences in predicted versus test temperatures would indicate areas requiring more study, either in the design or in the test technique.

DESCRIPTION OF THE SOLAR SIMULATION TEST

The test was performed in a seven-foot diameter, eight-foot long, cylindrical thermal-vacuum chamber. Located at one end of the chamber is a one-foot diameter quartz port, through which the carbon arc beam was introduced (Figure 4).

The spacecraft was mounted on a rotator-gimbal mount which provided two axis motion;

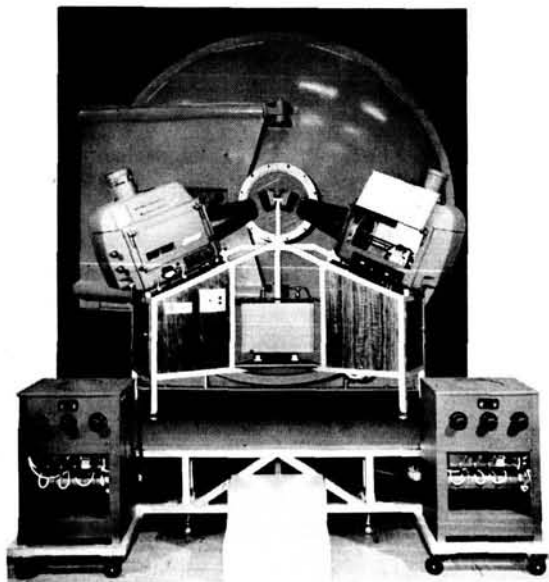


FIGURE 4.

spin about the centerline of the spacecraft and inclination relative to the incident simulated solar radiation. Because of the physical limitations of the facility size, the inertia booms, galactic noise experiment dipole booms and solar paddles were not included in the test configuration. Also, shortened telemetry antennas were substituted for the full-length antennas during this test (Figure 5). Two modes of information

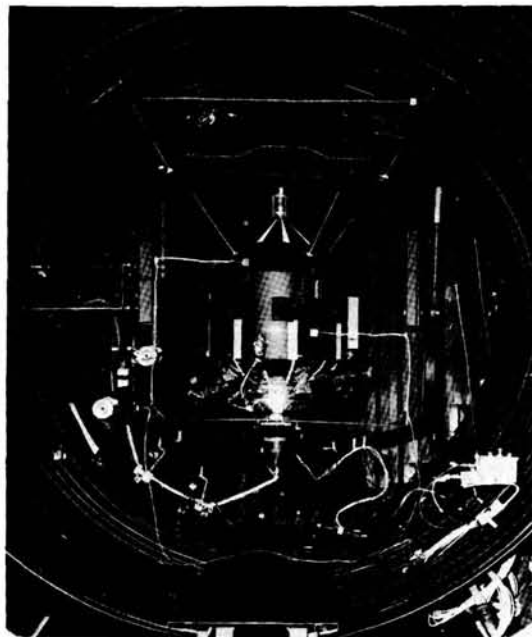


FIGURE 5.—UK-2/S-52 mounted on rotator-gimbal in the thermal-vacuum chamber.

were available from the spacecraft; normal telemetry transmission and command receiver, and a test hardline-slip ring combination for power and supplementary data. The former system provided information in the same format as its orbital operation while the latter provided the capability of recharging the on-board batteries and transmitting data from temperature sensors mounted for test purposes. These sensor outputs were scanned during the test to minimize the number of wires penetrating the chamber.

The pressure environment was in the range of 1×10^{-7} mm of Hg while the chamber walls were at approximately liquid nitrogen temperatures, -190°C . The simulated solar radiation was set at an equivalent one solar constant as described

in the problems section of this report. The spacecraft was fully operational from an experiment and subsystem basis while rotating on its spin axis.

The spacecraft was separated from the rotator-gimbal mount by a nylon insulator. Heat was supplied to the mount to such a degree that a small temperature gradient existed at all times. This minimized the energy flow between the spacecraft and mount. Since the temperatures were monitored throughout the test, approximate values of heat gain or loss were calculated. These values were introduced into the thermal model, to be reflected in the spacecraft energy balance and test temperature predictions.

Three spacecraft-incident solar radiation aspects were tested. They consisted of the 90° aspect or broad-side solar exposure, the 45° aspect or maximum top-to-bottom temperature gradient expected in flight, and the 135° aspect or maximum bottom-to-top temperature gradient expected.

The variation of sun exposure in orbit is from 63 percent to 100 percent. The 63 percent exposure consisted of a series of cycles of 55-minute sunlight—32-minute shade periods. The 90° aspect position was tested at both of the above exposures to determine the effect on mean spacecraft temperature. The remaining aspects of 45° and 135° were tested at the 100 percent sunlight condition only, because the primary objective was to study the internal temperature gradients resulting from these aspects.

PROBLEMS ASSOCIATED WITH SOLAR SIMULATION

Some of the difficulties encountered in attempting to simulate solar radiation are related below. The efforts to compensate for these problems are discussed in detail.

The choice of radiation source in large part is dictated by the desire to match as closely as possible the spectrum of the sun at orbital altitude. In addition, the ease of handling the source during the test must be considered.

Spectrum

Any significant departure in the spectral distribution of the simulator source from the sun's spectrum, as defined in the Johnson Curve, may

cause a change in the absorbed energy of the exposed surface.

Carbon arc was chosen as the solar source because of its close spectral match with the Johnson Curve. Open-arc sources, however, present several operational problems, and these are discussed in the following paragraphs.

Uniformity and Degradation of the Carbon Arc Solar Simulator

The carbon arc system consisted of two modified reflector arc lamps. The system contained no optics except the quartz port in the vacuum chamber. Two arc lamps were used for two reasons. First, the carbon rod feed-mechanism created a shadow on the reflector and a discontinuity in the intensity of the beam from a single lamp. Second, the necessity for replacing the consumable rod in each lamp every hour warranted the use of multiple lamps to minimize the interruption of radiation energy input to the spacecraft.

The use of a reflector system is beneficial from a power and efficiency standpoint, but creates uniformity and degradation problems. The exposure of the reflector to the open arc permits the deposit of vaporized carbon on the reflector surface, resulting in loss of reflectance efficiency and an ever-changing intensity distribution in the projected beam. To compensate for this condition during the test, intensity mappings of the projected beam were made at selected intervals in the test cycle. A rolled-ribbon thermocouple radiation detector mounted on an X-Y plotting board was used for this mapping procedure. This provided an interest calibration of the radiation beam and also a reference point for total intensity adjustment to compensate for reflector degradation.

Determination of Carbon Arc Intensity

The first step in determining the intensity or solar flux on the spacecraft is to determine the effective (α/ϵ) of the total spacecraft when illuminated by the natural sun. This is then combined with the carbon arc intensity and the natural sun intensity as follows:

$$I_s(\alpha/\epsilon)_s = I_c(\alpha/\epsilon)_c \quad (1)$$

The effective (α/ϵ) is determined by the following equation:

$$\alpha \text{ eff} = \bar{A}_1(\alpha)_1 + \bar{A}_2(\alpha)_2 + \dots + \bar{A}_n(\alpha)_n \quad (2)$$

$$\epsilon \text{ eff} = \bar{A}_1(\epsilon)_1 + \bar{A}_2(\epsilon)_2 + \dots + \bar{A}_n(\epsilon)_n \quad (3)$$

$$(\alpha/\epsilon) \text{ eff} = \frac{\alpha \text{ eff}}{\epsilon \text{ eff}} \quad (4)$$

where: \bar{A} = ratio of coating area to total exposed area as given in Figure 7.

α = absorptance.

ϵ = emittance.

Referring to Tables 1 and 2, which give the coating thermal properties and their areas, respectively, the following effective (α/ϵ) 's are derived:

$$(\alpha/\epsilon)_{\text{eff-solar}} = 1.04 \quad (5)$$

$$(\alpha/\epsilon)_{\text{eff-chamber}} = 1.03. \quad (6)$$

The second step in determining the intensity is to compensate for the time required to change a carbon rod. The arc system used employs two separate lamps, each illuminating one-half of the

spacecraft. Each lamp has one carbon-rod change every hour; 1.5 minutes is allowed for this change. Therefore, the following ratio is used to adjust total energy for this down-time:

$$I_{\text{total}}(t)_{\text{total}} = I_{\text{eff}}(t)_{\text{eff}} \quad (7)$$

$$I_{\text{eff}} = 1.03 I_{\text{total}}. \quad (8)$$

Combining equations (1), (5) and (6), we get, assuming I_s = solar constant:

$$I_c = 1.01 \text{ s.c.} \quad (9)$$

Combining equations (8) and (9):

$$I_{\text{eff}} = 1.04 \text{ s.c.} \quad (10)$$

Intensity Calibration and Monitoring of the Carbon Arc Solar Simulator

The lack of an absolute solar standard detector made it necessary to use a secondary method of intensity calibration. The method chosen was the integrating black-ball technique where a thin-shell aluminum ball, the approximate size of the spacecraft, is placed in the test chamber in the exact location the spacecraft will occupy during the test. The integrating ball has a small thin-shell ball suspended in its center. The inner ball temperature is measured by a thermocouple. Both balls are coated with a black paint of known absorptance and emittance. A calculation is made to determine the stabilization temperature of the ball for a flux of 1.04 solar constant. The chamber is then evacuated and the shrouds flooded with liquid nitrogen to create the same simulated environment that the spacecraft will see. The arc output is then adjusted so that the

TABLE 1.—Spacecraft Coating Thermal Properties

Coating	Solar			Chamber		
	α	ϵ	α/ϵ	α	ϵ	α/ϵ
Black.....	.96	.86	1.12	.96	.86	1.12
White.....	.27	.86	0.31	.26	.86	0.30
Evaporated Al.....	.12	.04	3.0	.11	.04	2.77

TABLE 2.—Spacecraft Coating Area Evaluation

Spacecraft Surface	Area Ft ²	% of Total A.	White %	White % of T.A.	Black %	Black % of T.A.	Eval. Al. %	Evap. Al. % of T.A.	Al. Foil %	Al. Foil % of T.A.
Top Dome.....	2.82	23.5	0	0	100	23.5	0	0	0	0
B.B. Support Cone...	1.16	10.0	25	2.5	75	7.5	0	0	0	0
Upper Mid Skin.....	2.50	21.0	0	0	100	21.0	0	0	0	0
Lower Mid Skin.....	2.58	21.5	20	4.3	80	17.2	0	0	0	0
Dottom Dome.....	2.53	21.5	10	2.1	0	0	90	19.4	0	0
Mounting Ring.....	0.30	2.5	0	0	0	0	0	0	100	2.5
Total Area.....	11.89			8.9		69.2		19.4		2.5

ball system stabilizes to the predetermined temperature. Once this focus adjustment is made, the arc beam is then mapped for uniformity and intensity with the radiation detector on the X-Y plotting board located external to the chamber. The readings of the detector are then benchmarked to the intensity incident on the black ball and used throughout the test of the spacecraft as a relative monitoring point.

The use of the black-ball technique for calibrating the arcs serves a dual purpose; the first, arc calibration as described above and, second, inclusion of inherent chamber energy (reflected or radiative) in the calibration of total energy absorbed by the black ball. The black coating of the ball is spectrally flat in its absorption characteristics and, therefore, arrives at the desired stabilization temperature by summing the energies from the carbon arcs and the inherent chamber sources.

This technique is applicable if the spacecraft to be tested is also essentially spectrally flat in its absorption characteristics as in the case of the UK-2/S-52.

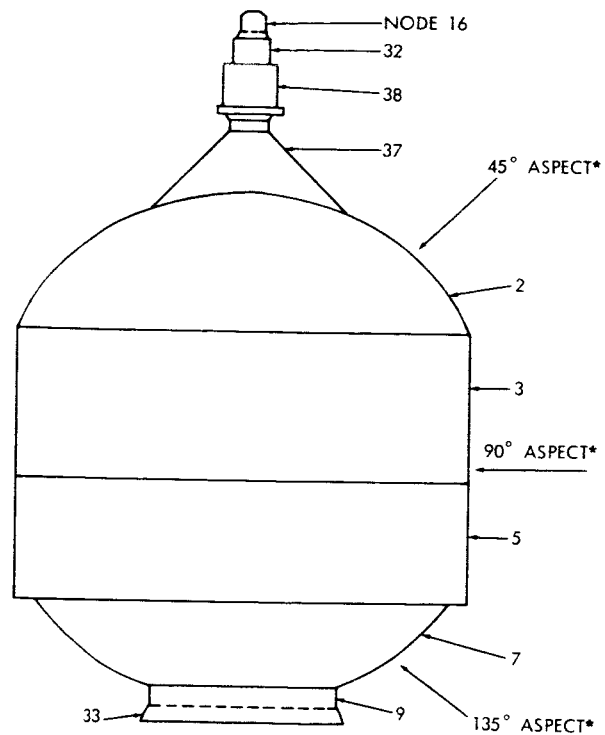
Beam Collimation

The radiation beam projected, by the solar simulator is divergent with a $7\frac{1}{2}^\circ$ half-angle. This permits the system to project a 36-inch diameter beam at the centerline plane of the spacecraft. This divergence rate is not considered severe enough to cause unnatural shadowing, however a measured energy change per unit area of approximately one percent per inch of depth variation is produced in the test volume. Therefore, incident intensity compensation was made to the thermal model of the spacecraft to take into account this change of intensity with depth.

Determination of Local Solar Intensity Input to The Spacecraft

The thermal model segments the spacecraft into thermal nodes; those used for this test are illustrated in Figure 6 along with the definition of aspect angles used in the test series.

The intensity uniformity of the projected beam was mapped with an X-Y plotting board, as previously described. Figure 7 shows a typical uniformity distribution plot taken with the GE



*Angle represents spacecraft-sun aspect, and arrow represents sunlight.

FIGURE 6.—Computer thermal nodes and spacecraft aspect angles.

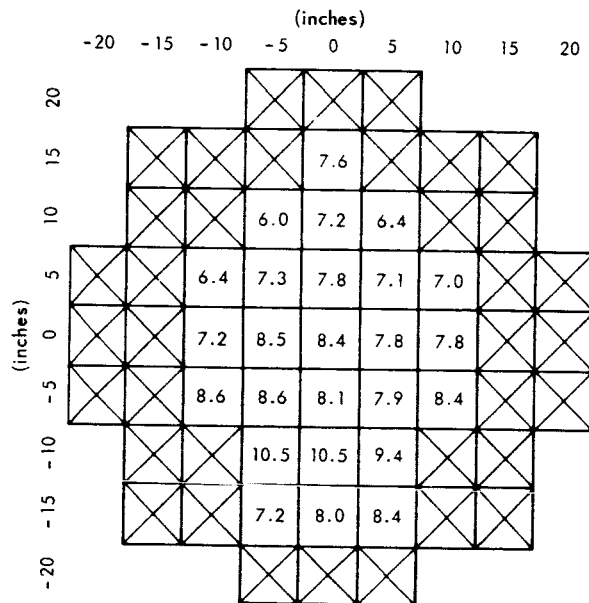


FIGURE 7.—Carbon arc uniformity distribution survey.

detector; the numbers are millivolt output readings which, when used in conjunction with a calibration curve, yield relative intensity readings. The uniformity plots were used to determine the local intensity at the individual thermal nodes given in Figure 6. Three planes of intensity perpendicular to the incident beam were plotted; at the spacecraft centerline, six inches nearer the beam source and 12 inches nearer the beam source. This was done to determine the change of intensity with distance from the source, since the beam is diverging at the rate of $7\frac{1}{2}^\circ$ half-angle. This calibration indicated a change of six percent intensity for each position or a range of 12 percent over the full 12 inches of the beam depth covered. The above information was then used in refining the local intensity values to be used in the thermal analysis. Table 3 gives these values.

It should be noted here that in a reflector-focus arc system, the reflector degrades in performance as it becomes coated with carbon deposit from the open arc. This degradation, decreases the total intensity for a given focal length, and randomly changes the uniformity pattern as some areas of the reflector receive carbon deposit. Because of this condition, in-test monitoring is necessary to change the arc focal length, compensating for the reflector degradation. Since

no adequate real-time in-chamber monitoring device was available, the detector and X-Y plotting board were used whenever the test schedule called for a simulated shade period in an orbit cycle. These external in-test uniformity plots were then used to indicate the status of the solar flux just prior to the shade period. In addition, they were used to determine what re-focusing was necessary to restore the total intensity back to the desired 1.04 solar constant. Therefore, a test prediction is made from intensity plots obtained during the test.

SOLAR SIMULATION TEST RESULTS

Figures 8, 9, and 10 are thermal profiles that have been included to show the comparison between predicted and test temperatures. Included are the thermal gradients existing in the spacecraft at aspect angles of 135° and 45° .

From the data below, the following observations are made:

The effect of aspect angle on the mean temperature of the spacecraft, as represented by the battery, was 23°C .

The effect of sunlight-shade exposure time on the mean temperature of the spacecraft was 33°C . Figure 11 presents a sample of data from this test. It is valuable as an indicator of the interactions of spacecraft structure and components.

TABLE 3.—Test Chamber Total Intensity Inputs To The Predict Program

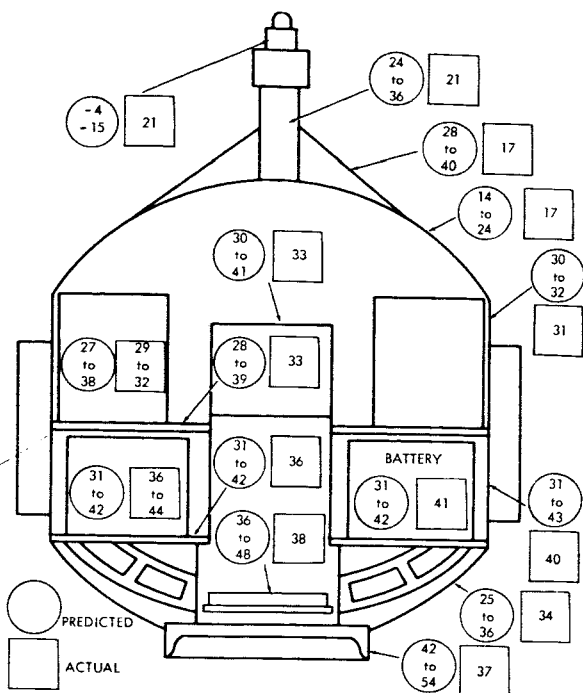
Spacecraft Node (See Figure 6)	90° Aspect 100% Sun I_T	90° Aspect 63% - 37% Sun I_T	45° Aspect 100% Sun I_T	135° Aspect 100% Sun I_T
2.....	550	597	529	0
3.....	470	505	492	529
5.....	468	500	454	484
7.....	581	496	0	495
9.....	522	478	0	518
9.....	-27	-18	-17	+1.0
16.....	330	365	728	490
32.....	330	365	715	504
37.....	583	729	592	0
38.....	390	400	689	525

Notes:

1. Units $\frac{\text{BTU}}{\text{HR Ft}^2}$

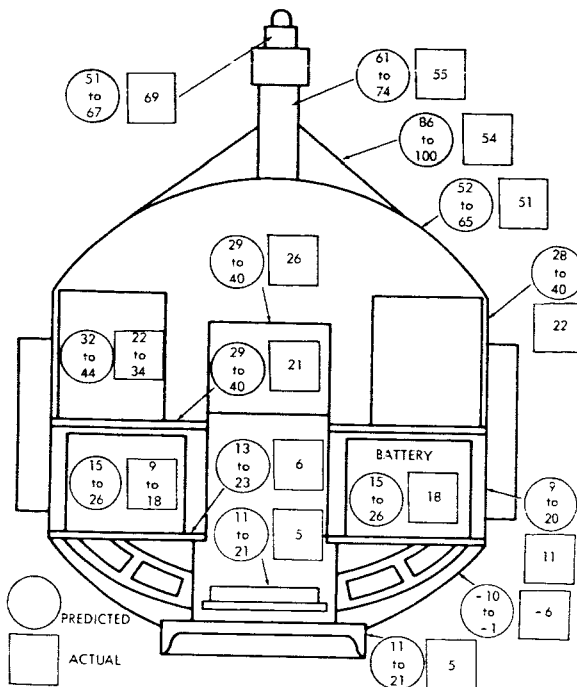
2. The intensities (I_T) Represents Available Energy To The Nodes

3. Node 9 • Input Is Conductive Energy Across The Spacecraft Rotator Interface



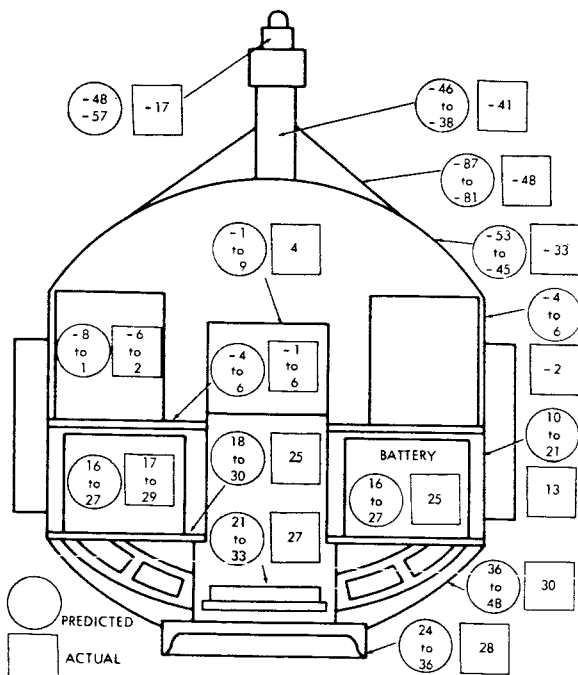
Sun Cycle: 100%
Aspect: 90°

FIGURE 8.—Temperature predictions versus actual test °C.



Sun Cycle: 100%
Aspect: 45°

FIGURE 10.—Temperature predictions versus actual test °C.



Sun Cycle: 100%
Aspect: 135°

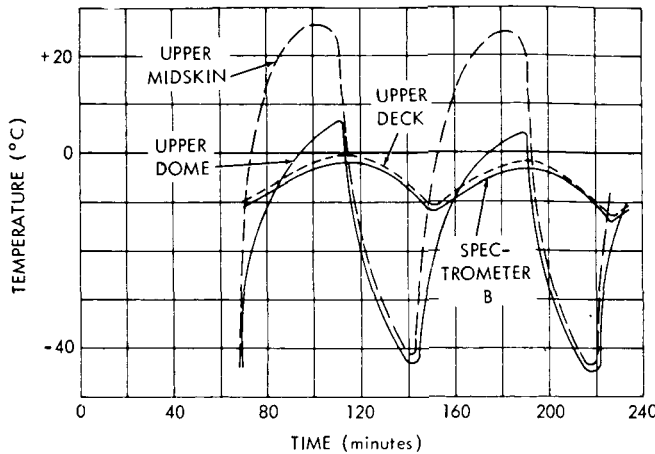
FIGURE 9.—Temperature predictions versus actual test °C.

The temperature excursion of the broad-band ozone detector due to sun aspect angle was about 85°C (69°C to -17°C) from test data, whereas the predicted excursion was approximately 110°C (59°C to -52°C). The test results indicate that the external portion of the detector does not reach the extreme temperatures that are predicted, but more closely follow the temperature of its enclosed base, the support cone, and the top dome.

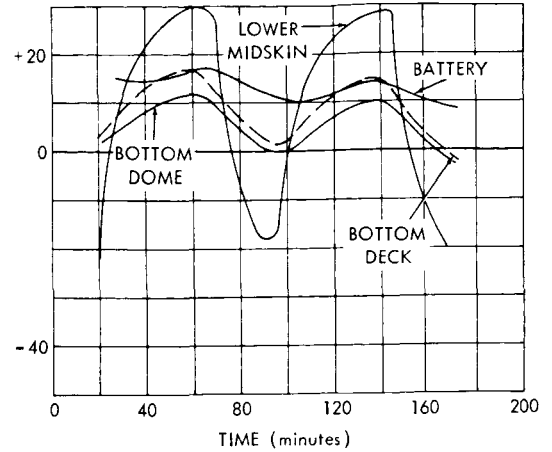
Two of the experiments required solar energy for activation, thus complete performance was obtained under simulated space conditions.

A coating failure occurred which directed attention to a review of the adhesion characteristics under stress and the preparation of surfaces to be painted.

The spacecraft was fully operational throughout the test and no problems were experienced with components or ground station. One minor exception was an occasional noise interference between the carbon arcs and the spacecraft programmer which controls operating modes of the spacecraft.



Typical Upper Deck and Skins Interaction During Cycles



Typical Lower Deck and Skins Interaction During Cycles

FIGURE 11.—Typical cyclic data.

CONCLUSIONS

1. The solar simulation test verified the assumptions made for the thermal model excluding the broad-band experiment.

2. The spacecraft will operate satisfactorily under space conditions.

3. The experiments which were stimulated by the simulated solar energy performed properly, and indicate successful operation will be obtained in space.

4. Carbon arc is a useful solar simulation source although a simulator utilizing optics, rather than reflector focus, would provide better uniformity.

5. The non-uniformity of the arc beam was adequately compensated by the rotating spacecraft and the nodal energy input to the computer program.

REFERENCES

1. HRYCAK, UNGER, and WITTENBERG, "Thermal Testing of the Telstar Satellite," *Institute of Environmental Sciences 1963 Proceedings*, p. 377.
2. HAMILTON and MORGAN, "Radiant-Interchange Configuration Factors," *NACA TN-2836*, December, 1952.
3. POWERS, E. I., "Thermal Radiation to a Flat Surface Rotating About an Arbitrary Axis in An Elliptical Earth Orbit: Application to Spin Stabilized Satellites," *NASA TN D-2147*, Goddard Space Flight Center, Greenbelt, Maryland, January, 1964.

CORIOLIS EFFECTS IN COMBINED ENVIRONMENT TESTING*

JOSEPH C. BOYLE

When the movement of a body is described with respect to a set of coordinate axes that are in motion, it is necessary to distinguish between absolute and relative motions. The relations between forces, motions, and inertial properties of the body are derived herein for the special case in which the set of coordinate axes is rotating at constant angular velocity about a fixed axis. Illustrative examples are included for a body undergoing linear vibration while being rotated about a distant axis and for a body spinning on its own axis while being rotated.

INTRODUCTION

During the launch of a spacecraft, sustained accelerations that are essentially rectilinear are produced. In ground testing, rectilinear accelerators, such as rocket sleds, are of limited utility because of the short period of time during which sustained accelerations may be maintained without involving many miles of track and because of the high level vibration inadvertently introduced by track impact and rocket combustion. As a consequence, curvilinear accelerators, such as centrifuges, are used; with these, sustained accelerations may be prolonged indefinitely.

Except for the fact that a radial acceleration gradient will exist when a curvilinear accelerator is used, it makes little difference which type of accelerator is used if the only requirement of the test is to produce sustained acceleration. However, if additional input motions such as vibration or spin of the test body are included, the resultant accelerations will differ depending on whether a rectilinear or curvilinear accelerator is employed.

In describing the vibration or spin of a test body mounted on a centrifuge, the most convenient references are a set of coordinate axes that are stationary with respect to the centrifuge, as illustrated by Figure 1. Here the x axis coincides with a horizontal radius of rotation, the y axis is horizontal and tangent to the circular path followed by the point of origin, and the z axis is vertical and parallel to the axis of rotation. Motion recorded with respect to these coordinate axes represents

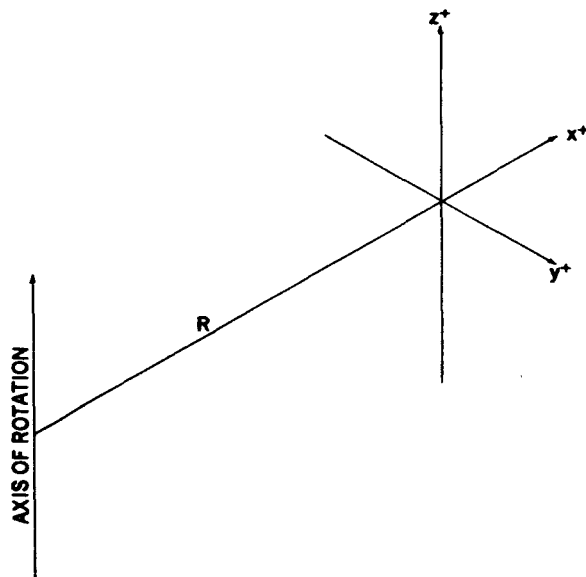


FIGURE 1.—An orthogonal set of coordinate axes, x , y , and z , that rotate with the centrifuge.

the *apparent* motion that would be seen by an observer riding the centrifuge. It does not, however, represent the *true* motion that would be seen by a stationary observer at a fixed point remote from the centrifuge (e.g., in prolongation of its axis of rotation). Suppose, for example, that a test specimen is simultaneously rotated on a centrifuge and vibrated in a radial direction. To the observer on the centrifuge, the vibratory motion appears to be purely rectilinear. Its actual path, however, combines radial and rotary motion, as shown in Figure 2.

*Published as NASA Technical Note D-1914, June 1964.

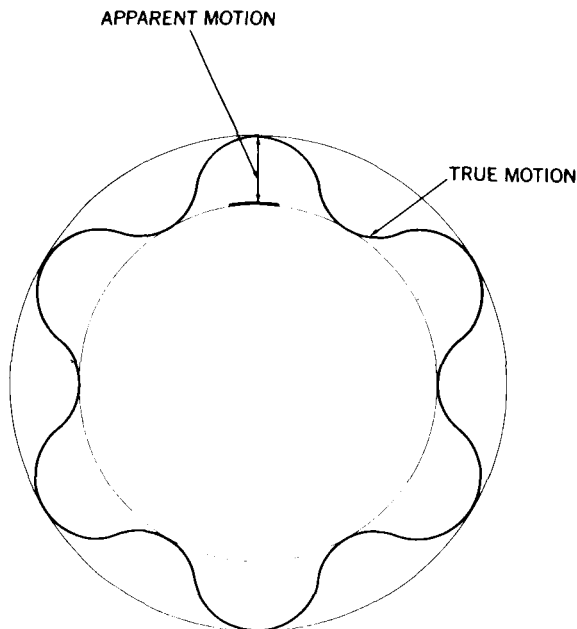


FIGURE 2.—Comparison of true and apparent motion of a point that is simultaneously vibrating and rotating at constant angular velocity.

This difference between true and apparent motion must be kept in mind when determining the acceleration to which a test body will be subjected. Newton's second law of motion, relating the change of momentum of a body to the applied force, applies to absolute or true motion rather than to relative or apparent motion. (For purposes here, the true motion of a body may be defined as its motion with respect to a set of coordinate axes fixed on the earth). Thus, although the motion of a body may be described with respect to a rotating set of axes, accelerations must be related to forces on the basis of the absolute motion. In the general solution of this problem, the methods of vector analysis are used to express the absolute acceleration of a particle as the sum of three component accelerations. The resulting expression, known as Coriolis' law, may be stated as follows:*

If a particle moves along a path and the path itself is not stationary but rotates, the instantaneous acceleration of the particle is the vector sum of three component accelerations. These are:

1. The acceleration the particle would have if the path were fixed and the particle moved along the path with velocity \mathbf{V}_R and acceleration \mathbf{a}_r .†
2. The acceleration of the point of the path coincident with the particle at a particular instant of time, that is, the acceleration the particle would have if it were fixed on the path and the path rotated.
3. The vector product $2\boldsymbol{\Omega} \times \mathbf{V}_R$ called the Coriolis component of acceleration.

The work encompassed by this paper and its appendixes consists of the determination of the equations of motion for a number of specific applications of Coriolis' law as applied to combined motion testing on a centrifuge. Some interpretation of these equations is included, and numerical results are derived for specific cases. The work is entirely analytical, no experimentation having been performed.

VIBRATION COMBINED WITH CENTRIFUGING

Consider a test body being simultaneously vibrated and centrifuged,‡ as shown in Figure 3.

†All symbols are defined in Appendix A. Boldface letters denote vector quantities.

‡For a detailed analysis, see Appendix B.

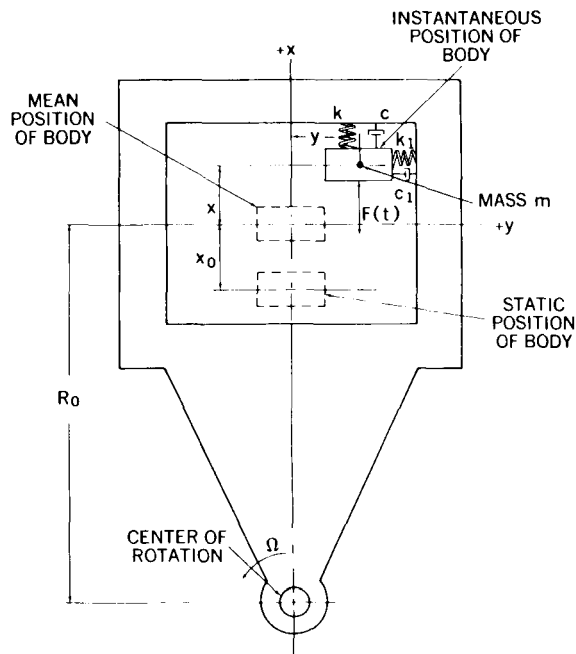


FIGURE 3.—Combined vibration and centrifuging of a body on compliant supports.

*An excellent treatment of this concept appears in: Seely, F. B., Ensign, E. N., Jones, P. G., "Analytic Mechanics for Engineers," 5th edition, New York: John Wiley and Sons, Inc., 1958.

It is assumed that the test body and the centrifuge structure are rigid and that the only compliance is in the springs k and k_1 . Damping is indicated by the coefficients c and c_1 .

Under static conditions the body is presumed to be in the position indicated. When the centrifuge is brought up to speed, the x spring deflects an amount x_0 , so that $k x_0 = mR_0\Omega^2$.

Subsequent vibration induced by a periodic force $F(t)$ causes an oscillation about the mean position. It is this oscillation and the associated acceleration that are of interest. The angular rotational rate of the centrifuge is assumed constant.

As shown in Appendix B, the x and y components of absolute acceleration may be derived by using Coriolis' law. They are

$$a_x = \ddot{x} - \Omega^2 (R_0 + x) + 2\Omega\dot{y} \quad (1)$$

and

$$a_y = \ddot{y} - \Omega^2 y - 2\Omega\dot{x} \quad (2)$$

From Newton's second law, the equations of motion are

$$m[\ddot{x} - \Omega^2(R_0 + x) + 2\Omega\dot{y}] = F(t) - k(x + x_0) - c\dot{x} \quad (3)$$

and

$$m(\ddot{y} - \Omega^2 y - 2\Omega\dot{x}) = -k_1 y - c_1 \dot{y} \quad (4)$$

Equations 3 and 4 describe the motion of the body with respect to the rotating x and y axes, and may be solved for particular forms of the input function $F(t)$. In the event that $F(t)$ is a sinusoid [$F(t) = F_0 \sin \omega t$], Equations 3 and 4 may be solved simultaneously to yield, for the steady-state vibratory motion,

$$\frac{x}{F_0} = \frac{k_1 - m\Omega^2 - m\omega^2 + j\omega c_1}{(k - m\Omega^2 - m\omega^2 + j\omega c)(k_1 - m\Omega^2 - m\omega^2 + j\omega c_1) - 4m^2 \omega^2 \Omega^2} \quad (5)$$

and

$$\frac{y}{F_0} = \frac{j 2m\Omega\omega}{(k - m\Omega^2 - m\omega^2 + j\omega c)(k_1 - m\Omega^2 - m\omega^2 + j\omega c_1) - 4m^2 \omega^2 \Omega^2} \quad (6)$$

Equations 5 and 6 represent the motion relative to the rotating set of axes. For the absolute acceleration, the values of x/F_0 and y/F_0 from Equations 5 and 6 have been substituted into Equations 1 and 2 to yield

$$\frac{a_x}{F_0} = \frac{(k_1 - m\Omega^2 - m\omega^2 + j\omega c_1)(-\omega^2 - \Omega^2) - 4m\Omega^2 \omega^2}{(k - m\Omega^2 - m\omega^2 + j\omega c)(k_1 - m\Omega^2 - m\omega^2 + j\omega c_1) - 4m^2 \Omega^2 \omega^2} \quad (7)$$

and

$$\frac{a_y}{F_0} = \frac{-2\Omega\omega(jk_1 - \omega c_1)}{(k - m\Omega^2 - m\omega^2 + j\omega c)(k_1 - m\Omega^2 - m\omega^2 + j\omega c_1) - 4m^2 \Omega^2 \omega^2} \quad (8)$$

Another quantity of interest is the ratio a_y/a_x , which may be defined as the "Coriolis crosstalk." This takes the form

$$\frac{a_y}{a_x} = \frac{-2\Omega\omega(jk_1 - \omega c_1)}{(k_1 - m\Omega^2 - m\omega^2 + j\omega c_1)(-\omega^2 - \Omega^2) - 4m\Omega^2 \omega^2} \quad (9)$$

Equations 5 through 8 may be described as transfer functions in that they relate output displacements and accelerations to the input force F_0 . Similarly, Equation 9 relates output ac-

celeration in the y direction to input acceleration in the x direction. These ratios also may be considered as compliances in that they relate motion to input force. Ratios of the peak values of the

sinusoidally varying inputs and outputs are represented; and, as the expressions are in general complex, there will be both a magnitude and phase relation between the peak values of input and output quantities. In the derivation of these expressions, the mean values of displacement and acceleration have been suppressed so that the resulting equations relate only the time-varying components of motion.

Equations 5 through 9 provide predictions of the input force required to produce a given acceleration to the test body at a given frequency (Equations 7 and 8), the resultant "Coriolis crosstalk" (Equation 9), and the extent of the displacements that will occur relative to the rotating reference system (Equations 5 and 6).

It is obvious that the rotational rate of the centrifuge affects all five of these quantities, since terms in Ω appear in each equation. It is similarly apparent that these five quantities are functions of shake frequency ω . Although quantitative evaluations depend on the specific parameters k , k_1 , c , c_1 , and m and on the frequencies Ω and ω , some general conclusions can be drawn from an examination of these expressions.

As has been stated, the quantities are frequency-dependent. If damping can be neglected, certain critical frequencies are found to exist. These frequencies, derived in Appendix B, are the following:

$$\omega_{10} = \sqrt{\frac{k_1 - m\Omega^2}{m}} \quad (10)$$

Equation 10 specifies the frequency at which $x/F_0=0$. At this frequency no compliance exists in the x direction and no x motion will occur, regardless of the magnitude of F_0 . However, acceleration will be present.

To gain a clearer insight, refer to the equation of motion in the x direction, Equation 3. Since no relative motion occurs, the equation reduces to $-m\Omega^2 R_0 + 2m\Omega \dot{y} = (F_0 \sin \omega t) - kx_0$. Equating the fixed and alternating portions separately to zero, we have $-m\Omega^2 R_0 = -kx_0$ and $2m\Omega \dot{y} = F_0 \sin \omega t$. These equations indicate that the input force will not produce x motion, since it is balanced by the Coriolis force resulting from the velocity in the y direction. Despite the absence of x motion, an alternating acceleration component of magnitude $2\Omega \dot{y} = F_0 \sin \omega t / m$ will exist in the x direction.

$$\omega_{11}^2 = \frac{k + k_1 + 2m\Omega^2 \pm \sqrt{(k_1 - k)^2 + 8m\Omega^2(k + k_1)}}{2m} \quad (11)$$

At these frequencies all x and y compliances, in terms of relative displacements and absolute acceleration, will be infinite. These are effectively resonant frequencies for the system.

$$\omega_{12}^2 = \frac{k_1 + 2m\Omega^2 \pm \sqrt{k_1^2 + 8mk_1\Omega^2}}{2m} \quad (12)$$

At these frequencies the accelerational compliance in the x direction a_x/F_0 will be zero: this indicates that no alternating component of a_x will occur at these frequencies regardless of input. Since a finite, nonzero compliance exists for a_y at these frequencies, the crosstalk ratio will be infinite.

It is interesting to note that, at the frequency defined by Equation 12, there will be relative motion in the x direction despite the fact that the alternating component of absolute acceleration is zero.

Obviously the extreme values discussed cannot

occur in practice, since there is always some damping present. However, resonant or anti-resonant effects can be expected at the associated frequencies.

To study the effect of variation in parameters on frequency response, it is desirable to represent the transfer functions in a normalized form. In doing this, the natural frequencies in the x and y directions are defined as:

$$\omega_x^2 = \frac{k - m\Omega^2}{m}$$

and

$$\omega_y^2 = \frac{k_1 - m\Omega^2}{m}$$

Damping ratios are defined as:

$$\xi = \frac{c}{c_c}$$

where

$$c_c = 2m\omega_x$$

$$\xi_1 = \frac{c_1}{c_{c_1}}$$

and

$$y_{ST} = \frac{F_0}{k_1 - m\Omega^2}$$

and

$$c_{c_1} = 2m\omega_y$$

The acceleration which would be produced by application of F_0 to the mass m is defined as

$$a_f = \frac{F_0}{m}$$

The static deflections in the x and y directions are defined as

$$x_{ST} = \frac{F_0}{k - m\Omega^2}$$

Substitution of these quantities into Equations 5 through 12 yields the following normalized expressions:

$$\frac{x}{x_{ST}} = \frac{-1 + \left(\frac{\omega}{\omega_y}\right)^2 - j2\xi_1 \frac{\omega}{\omega_y}}{4 \left(\frac{\Omega}{\omega_x}\right)^2 \left(\frac{\omega}{\omega_y}\right)^2 - \left[1 - \left(\frac{\omega}{\omega_y}\right)^2 + j2\xi_1 \frac{\omega}{\omega_y}\right] \left[1 - \left(\frac{\omega}{\omega_x}\right)^2 + j2\xi \frac{\omega}{\omega_x}\right]} \quad (5N)$$

$$\frac{y}{y_{ST}} = \frac{-j2 \left(\frac{\Omega}{\omega_x}\right) \left(\frac{\omega}{\omega_x}\right)}{4 \left(\frac{\Omega}{\omega_x}\right)^2 \left(\frac{\omega}{\omega_y}\right)^2 - \left[1 - \left(\frac{\omega}{\omega_y}\right)^2 + j2\xi_1 \frac{\omega}{\omega_y}\right] \left[1 - \left(\frac{\omega}{\omega_x}\right)^2 + j2\xi \frac{\omega}{\omega_x}\right]} \quad (6N)$$

$$\frac{a_x}{a_f} = \frac{\left[1 - \left(\frac{\omega}{\omega_y}\right)^2 + j2\xi_1 \frac{\omega}{\omega_y}\right] \left[\left(\frac{\omega}{\omega_x}\right)^2 + \left(\frac{\Omega}{\omega_x}\right)^2\right] + 4 \left(\frac{\Omega}{\omega_x}\right)^2 \left(\frac{\omega}{\omega_y}\right)^2}{4 \left(\frac{\Omega}{\omega_x}\right)^2 \left(\frac{\omega}{\omega_y}\right)^2 - \left[1 - \left(\frac{\omega}{\omega_y}\right)^2 + j2\xi_1 \frac{\omega}{\omega_y}\right] \left[1 - \left(\frac{\omega}{\omega_x}\right)^2 + j2\xi \frac{\omega}{\omega_x}\right]} \quad (7N)$$

$$\frac{a_y}{a_f} = \frac{j2 \frac{\Omega}{\omega_x} \frac{\omega}{\omega_x} + j2 \left(\frac{\Omega}{\omega_y}\right) \left(\frac{\Omega}{\omega_x}\right)^2 \left(\frac{\omega}{\omega_y}\right) - 4\xi_1 \left(\frac{\Omega}{\omega_x}\right) \left(\frac{\omega}{\omega_x}\right) \left(\frac{\omega}{\omega_y}\right)}{4 \left(\frac{\Omega}{\omega_x}\right)^2 \left(\frac{\omega}{\omega_y}\right)^2 - \left[1 - \left(\frac{\omega}{\omega_y}\right)^2 + j2\xi_1 \frac{\omega}{\omega_y}\right] \left[1 - \left(\frac{\omega}{\omega_x}\right)^2 + j2\xi \frac{\omega}{\omega_x}\right]} \quad (8N)$$

$$\frac{a_y}{a_x} = \frac{j2 \left(\frac{\Omega}{\omega_x}\right) \left(\frac{\omega}{\omega_x}\right) + j2 \left(\frac{\Omega}{\omega_y}\right) \left(\frac{\Omega}{\omega_x}\right)^2 \left(\frac{\omega}{\omega_y}\right) - 4\xi_1 \left(\frac{\Omega}{\omega_x}\right) \left(\frac{\omega}{\omega_x}\right) \left(\frac{\omega}{\omega_y}\right)}{\left[1 - \left(\frac{\omega}{\omega_y}\right)^2 + j2\xi_1 \frac{\omega}{\omega_y}\right] \left[\left(\frac{\omega}{\omega_x}\right)^2 + \left(\frac{\Omega}{\omega_x}\right)^2\right] + 4 \left(\frac{\Omega}{\omega_x}\right)^2 \left(\frac{\omega}{\omega_y}\right)^2} \quad (9N)$$

$$\frac{\omega_{10}}{\omega_x} = \frac{\omega_y}{\omega_x} \quad (10N)$$

$$\frac{\omega_{11}^2}{\omega_x^2} = \frac{1 + 4 \frac{\Omega^2}{\omega_x^2} + \frac{\omega_y^2}{\omega_x^2} \pm \sqrt{\left(1 + 4 \frac{\Omega^2}{\omega_x^2} + \frac{\omega_y^2}{\omega_x^2}\right)^2 - 4 \frac{\omega_y^2}{\omega_x^2}}}{2} \quad (11N)$$

$$\frac{\omega_{12}^2}{\omega_x^2} = \frac{\omega_y^2}{2\omega_x^2} \left[1 + 3 \frac{\Omega^2}{\omega_y^2} \pm \sqrt{\left(1 + 9 \frac{\Omega^2}{\omega_y^2} \right) \left(1 + \frac{\Omega^2}{\omega_y^2} \right)} \right]. \tag{12N}$$

The basic parameters governing the responses of the test body are ω_x , ω_y , Ω , ξ , and ξ_1 . The curves of Figures 4 through 18 reveal the nature of the frequency responses at steady state conditions and clearly demonstrate the effect of the critical frequencies on the shape of the response curves. In addition, the effect of variation in the parameters is shown by plotting families of curves in which ω_y/ω_x , Ω/ω_x , and ξ are individually

varied. For the sake of simplicity, ξ and ξ_1 have been assumed equal. ω_x has been selected as the basic parameter for normalizing the input frequency ω as well as for normalizing ω_y and Ω .

Of particular interest is the variation of Coriolis crosstalk with frequency (Figures 16, 17, and 18). In addition to the peak response at ω_{12} , differentiation of the undamped Coriolis crosstalk with respect to ω reveals the existence of stationary

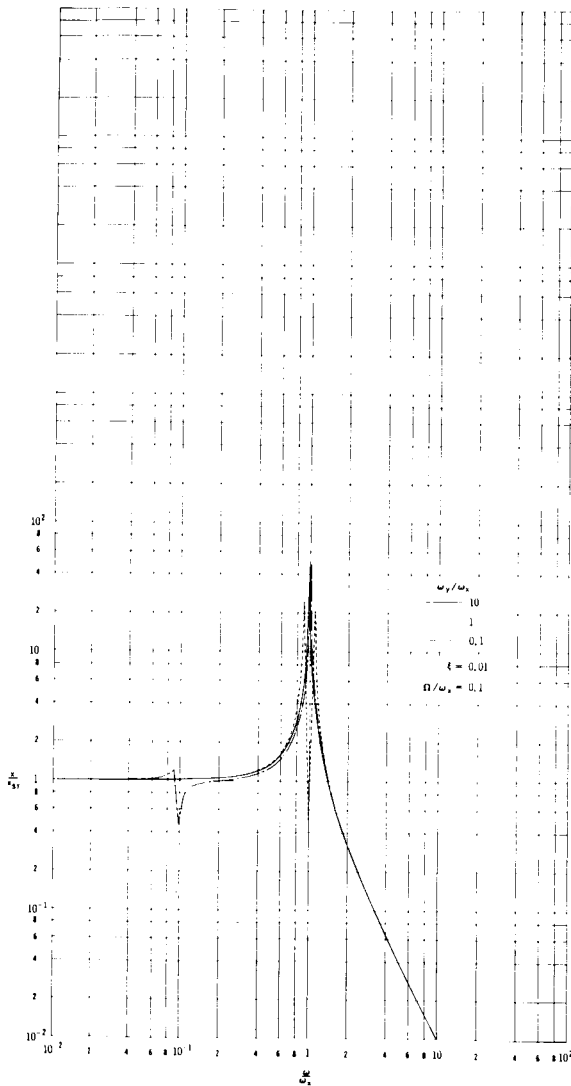


FIGURE 4.—Variation of x-displacement frequency response with ω_y/ω_x .

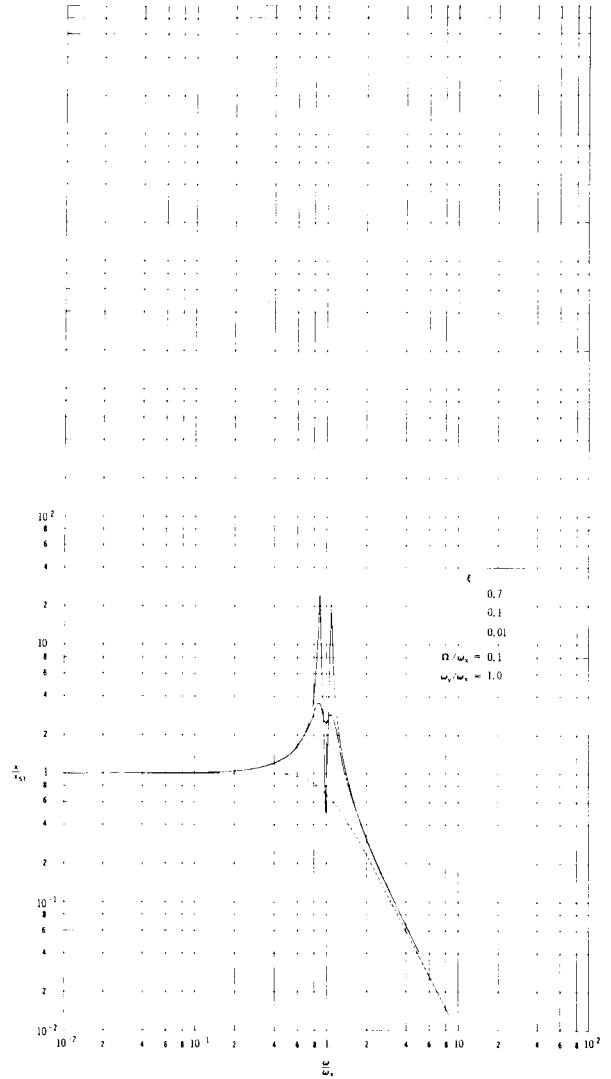


FIGURE 5.—Variation of x-displacement frequency response with ξ .

values at $\omega = \Omega$ and at $\omega = \omega_y/\sqrt{3}$. For zero damping, these values are

$$a_y/a_x = j \text{ at } \omega = \Omega \text{ and}$$

$$a_y/a_x = \frac{j3\sqrt{3}\Omega(\omega_y^2 + \Omega^2)}{\omega_y(\omega_y^2 + 9\Omega^2)} \text{ at } \omega = \frac{\omega_y}{\sqrt{3}}$$

The equations developed in this section and the plots derived from these equations show that the response to a test body to a sinusoidal forcing function may be markedly affected by rotation of the centrifuge and the resultant Coriolis accelera-

tions. Whereas with the centrifuge at standstill there would exist only one critical frequency at $\omega/\omega_x = 1$, Figure 6 shows that the presence of centrifuge rotation causes a pair of criticals to be present in the x/x_{ST} response. These criticals spread farther apart as the centrifuge speeds up, with a notch or anti-node appearing between them.

In the transverse direction, y/y_{ST} response (Figures 7, 8, and 9) exhibits criticals at the same frequency. No transverse motion would be present at standstill.

The accelerational responses a_x/a_f and a_y/a_f also show peaks at the same critical frequency as

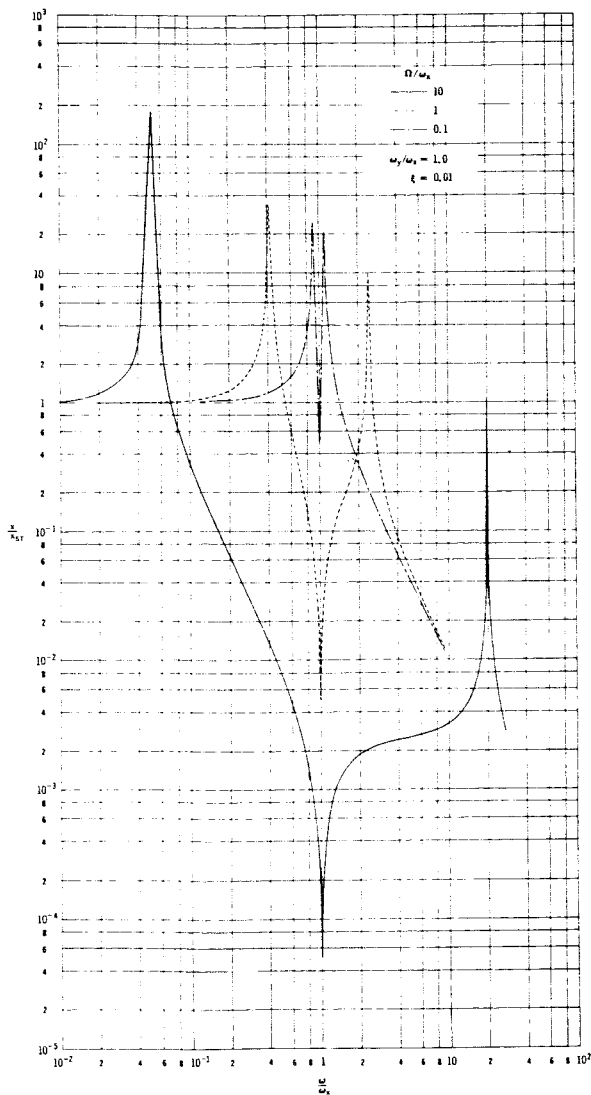


FIGURE 6.—Variation of x-displacement frequency response with Ω/ω_x .

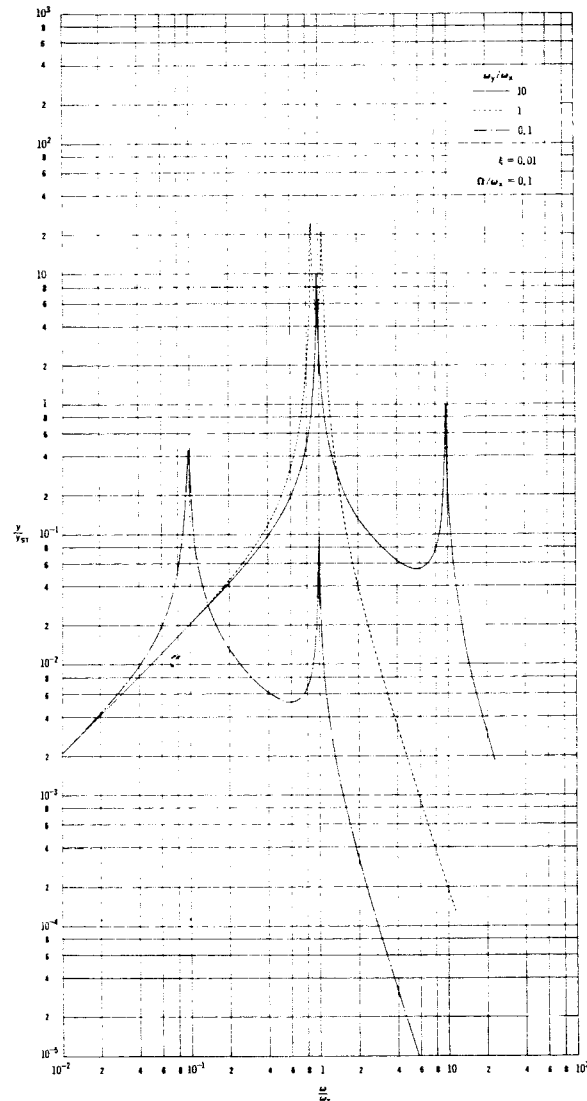


FIGURE 7.—Variation of y-displacement frequency response with ω_y/ω_x .

indicated by Figures 10 through 15. In addition to the peaks, the a_x/a_f curves indicate the presence of an intermediate notch. This notch is significant in that it indicates that little compliance exists to the forcing function, requiring a larger input. It is also significant that there is no corresponding dip in the a_y/a_f curves, resulting in Coriolis crosstalk (Figures 16, 17, and 18) peaks at these frequencies.

Increasing the amount of damping reduces the severity of all resonant peaks and notches (Figures 5, 8, 11, 14, and 17). It is significant to note that the stationary values occurring at $\omega/\omega_x = \Omega/\omega_x$

are not affected by damping. The Coriolis crosstalk has an invariant value of unity at this frequency ratio.

SPIN COMBINED WITH CENTRIFUGING

Suppose that a test body is to be simultaneously spun and centrifuged, as shown in Figure 19.

Assume the following:

1. The test body is rigid.
2. The rotational axes are perpendicular to each other.
3. ω is constant.
4. Ω is constant.

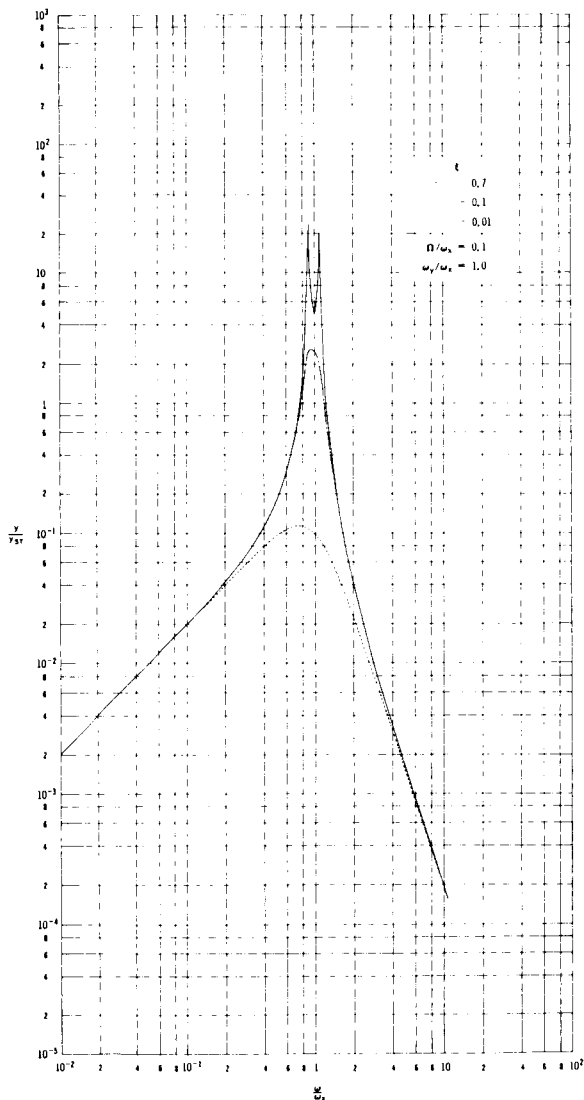


FIGURE 8.—Variation of y-displacement frequency response with ξ .

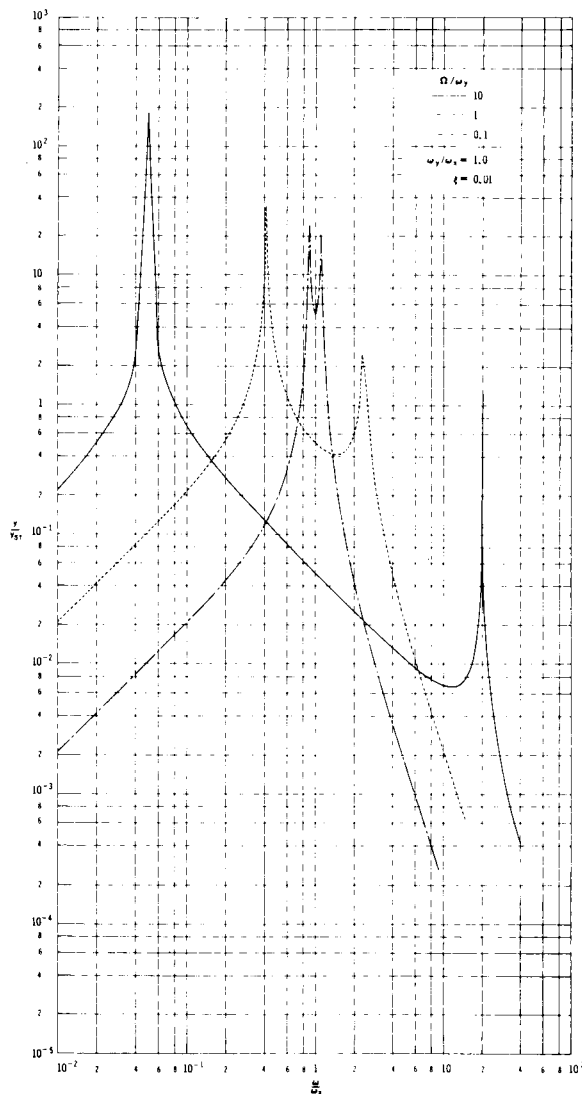


FIGURE 9.—Variation of y-displacement frequency response with Ω/ω_x .

5. The centrifuge structure is rigid.

The three components of acceleration needed to define the absolute acceleration of a particle of the body in Figure 19, are

1. Acceleration relative to rotating axes, $\omega \times (\omega \times \mathbf{r})$;
2. Acceleration if particle were fixed on the rotating axis at a given instant, while the axes rotated, $\Omega \times (\Omega \times \mathbf{R})$;
3. The Coriolis component, $2\Omega \times \mathbf{V}_R$.

The first of these quantities is the centripetal acceleration due to the spin rate ω , and the second is the centripetal acceleration due to the

spin rate Ω . Both of these components are constant in magnitude for any given particle of the test body. The third, the Coriolis component, varies in magnitude and direction as the test body spins; this component will be examined in some detail.

Certain general conclusions may be drawn relative to the inertial properties of the test body and the resultant Coriolis accelerations. These are derived in Appendix C and may be stated as follows:

1. There will be no resultant linear Coriolis acceleration of the test body provided the

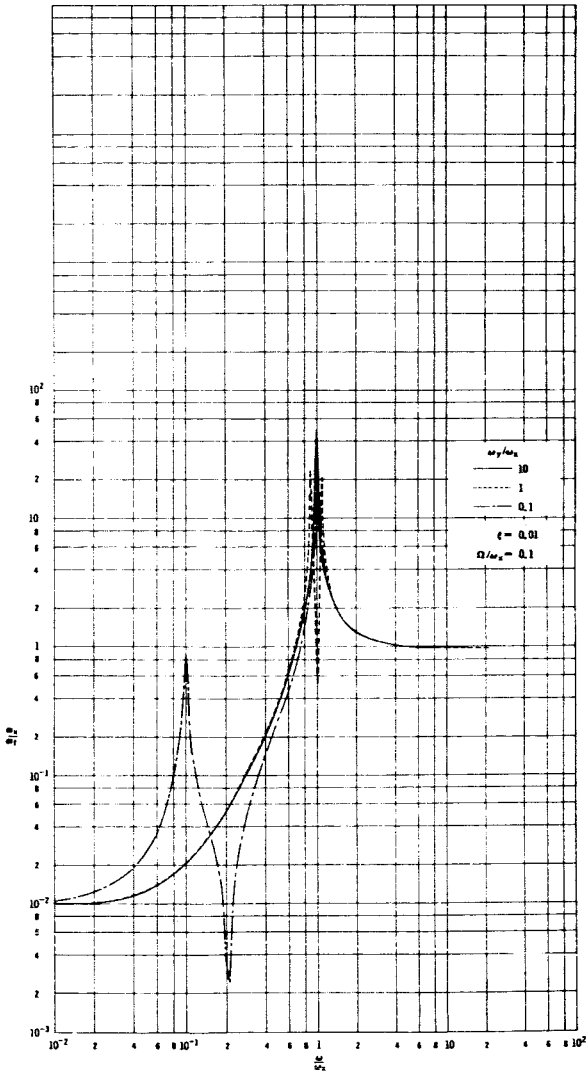


FIGURE 10.—Variation of a_x/a_f frequency response with ω_y/ω_x .

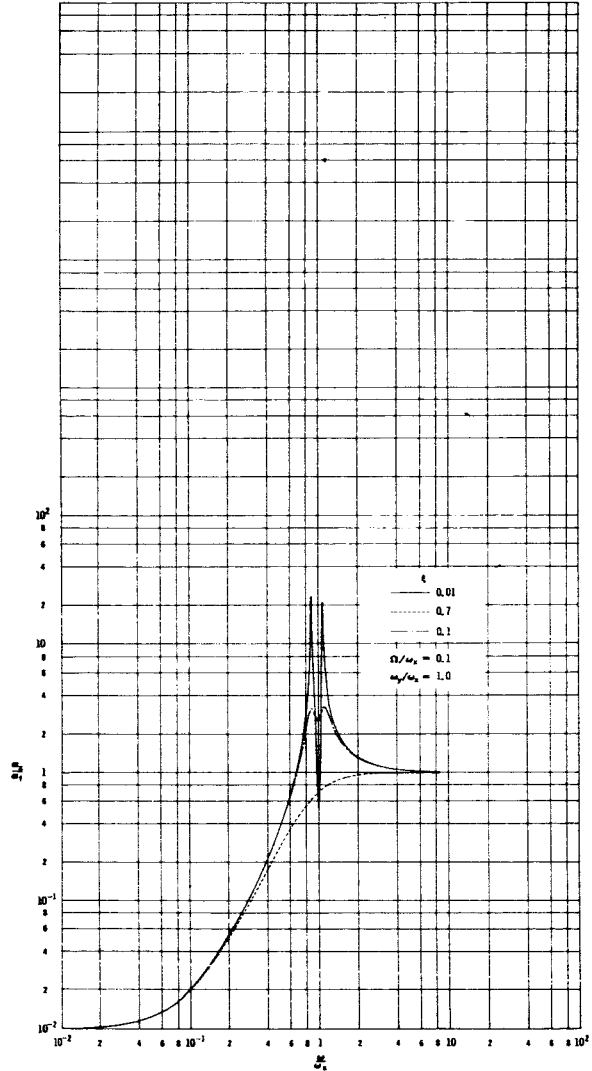


FIGURE 11.—Variation of a_x/a_f frequency response with ξ .

axis of spin passes through the center of gravity.

2. If the mass moments of inertia of the test

$$M_y = 2\Omega\omega I_{x'y'} \sin 2\omega t - 2\Omega\omega I_{y'} \cos^2 \omega t - 2\Omega\omega I_{z'} \sin^2 \omega t, \tag{13}$$

$$M_z = 2\Omega\omega I_{y'z'} (\cos^2 \omega t - \sin^2 \omega t) + \Omega\omega \sin 2\omega t (I_{y'} - I_{z'}). \tag{14}$$

3. If the mass moments of inertia about the principal transverse axes are equal ($I_{y'} = I_{z'}$), the resultant Coriolis torque about the y and z axes will be constant as expressed by the

equations

$$\left. \begin{aligned} M_y &= -I_{x'}\Omega\omega, \\ M_z &= 0. \end{aligned} \right\} \tag{15}$$

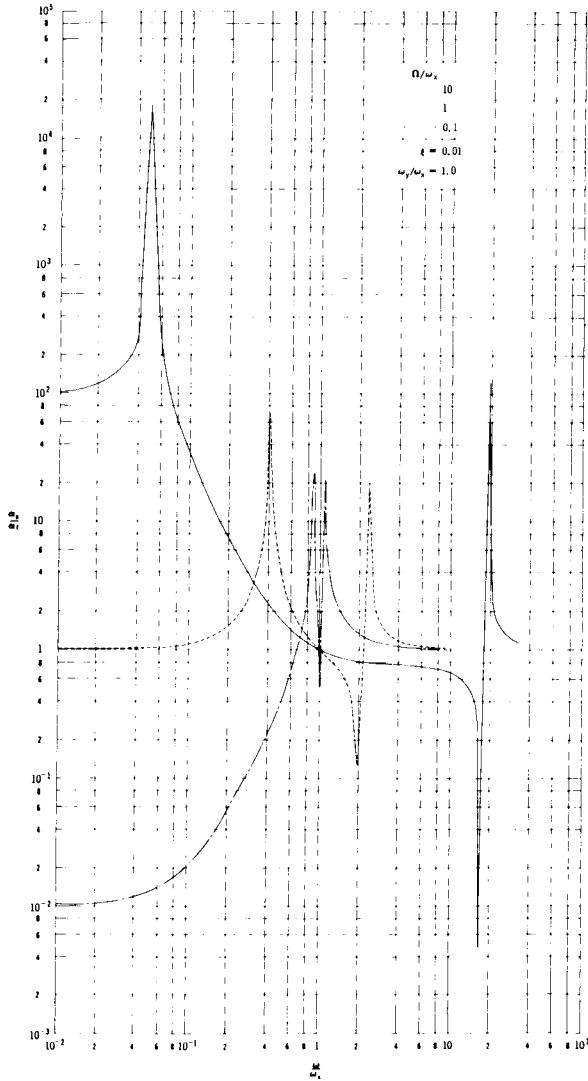


FIGURE 12.—Variation of a_x/a_f frequency response with Ω/ω_x .

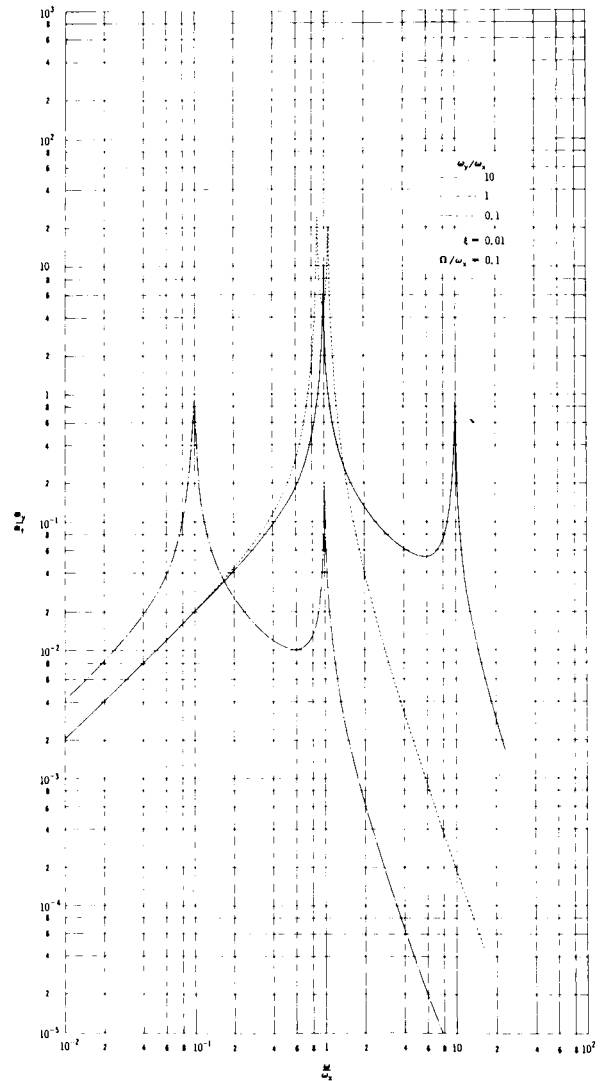


FIGURE 13.—Variation of a_y/a_f frequency response with ω_y/ω_x .

4. It is of interest to consider the situation where the axes of rotation are not perpendicular to one another, such as when the spin axis is tilted through an angle θ . In this case the reference axes are rotated about the y axis so that the x and z axes make an angle θ with the radial and vertical directions, respectively. Under these circumstances the total Coriolis moment is obtained by substituting $\Omega \cos \theta$ for Ω in Equation 13 through 15. In particular, for the case where $I'_y = I'_z$

$$M_y = - I_x \Omega \cos \theta ;$$

and, in vector notation,

$$M_y = I_x \Omega \times \omega ,$$

$$M_z = 0 .$$

The in-line component of Ω produces a radially directed Coriolis force that is constant in magnitude. This force adds to or subtracts from the centrifugal force, depending on whether the $\Omega \sin \theta$ and ω vectors are in the same or in opposite directions.

While the resultant moment of the symmetrical body is constant with respect to the y and z axes

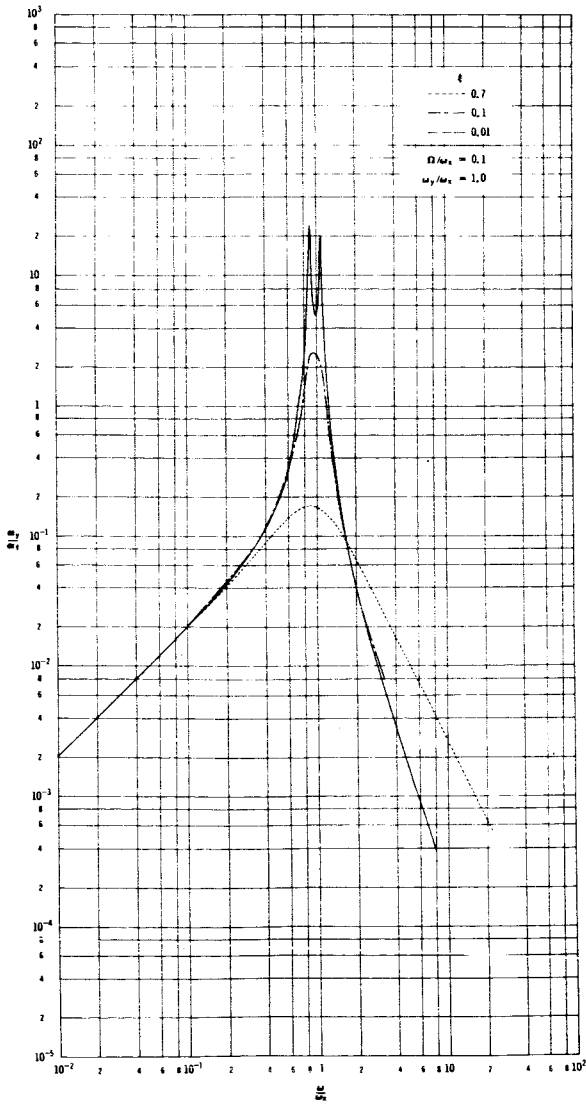


FIGURE 14.—Variation of a_y/a_f frequency response with ξ .

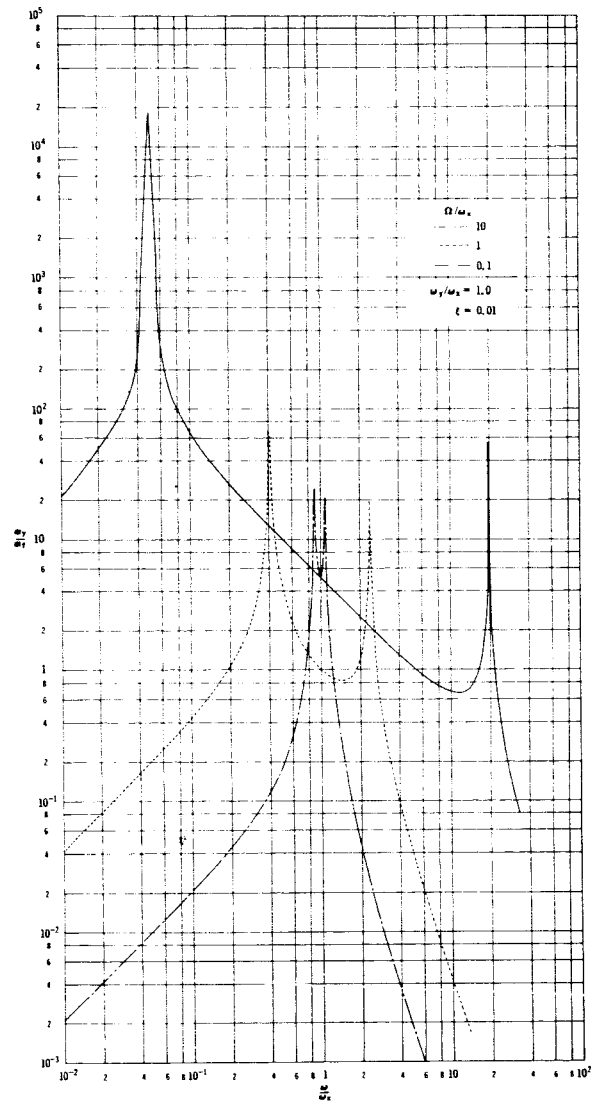


FIGURE 15.—Variation of a_y/a_f frequency response with Ω/ω_x .

and while the resultant cross-sectional force is zero, the axially directed force felt by a single particle will be an alternating one, equal to

$$dF_c = dm2\Omega \cos \theta \omega r \sin (\omega t + \alpha).$$

A logical question is "How serious is the Coriolis effect on a spinning test body?" There is no general answer to this question, but a specific illustration may be of interest. Consider a solid, homogeneous, cylindrical test body of length l and radius r_0 , spinning while being centrifuged. The

resultant Coriolis moment on such a body is

$$M_y = I_x \omega \times \Omega = \frac{mr_0^2}{2} \omega \Omega \sin \left(\frac{\pi}{2} + \theta \right).$$

Assuming that the body is supported at its base, we may compare the above moment with the cantilever gravity moment of the body about the base:

$$M_g = \frac{mg l}{2}.$$

Also of interest is the maximum alternating com-

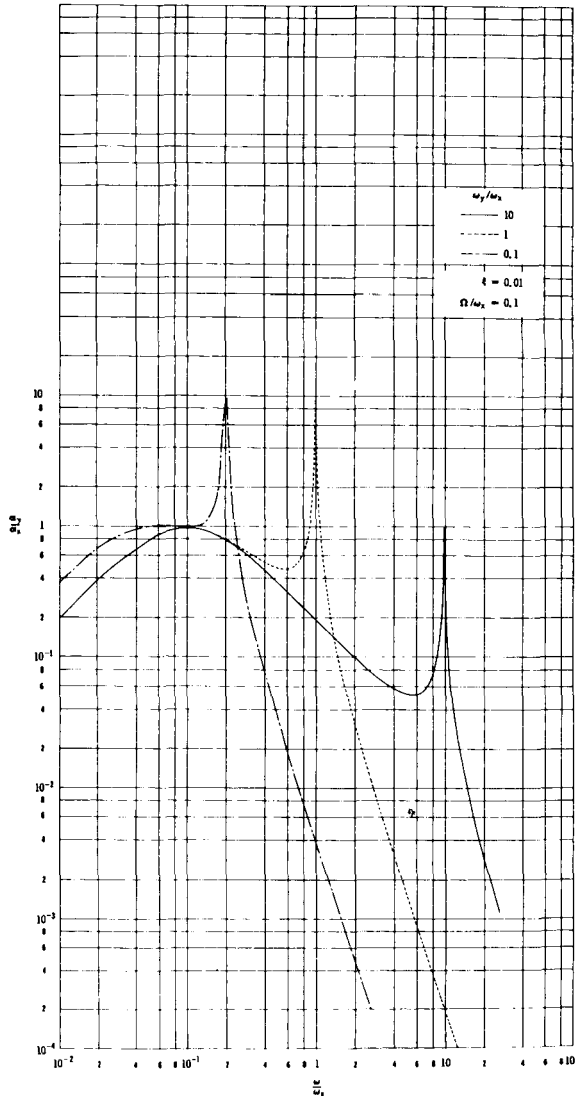


FIGURE 16.—Variation of Coriolis crosstalk frequency response with ω_y/ω_x .

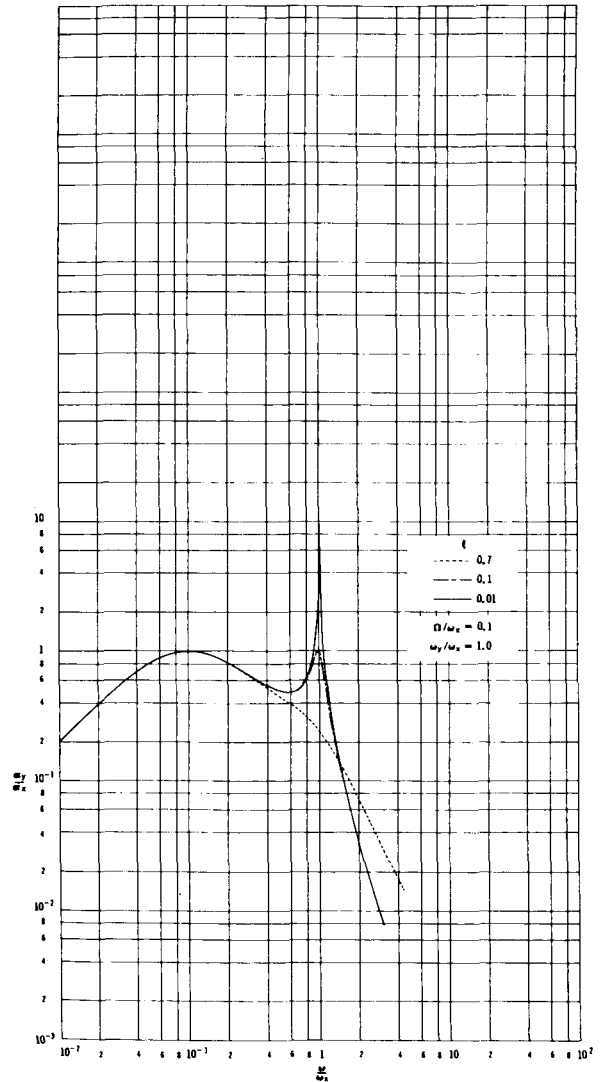


FIGURE 17.—Variation of Coriolis crosstalk frequency response with ξ .

ponent of Coriolis acceleration of a particle at the outer radius, $2\Omega\omega r_0 \cos \theta$.

If the body is to be tested in a horizontal position, the resultant of the gravity and Coriolis moments may be reduced if the directions of ω and Ω are such as to make them of opposite sign. In addition, an angle of tilt may be imposed in the vertical, so that a component of the sustained centripetal acceleration acts transversely to the spin axis. This permits a further adjustment of the resultant bending moment. Analysis (Appendix D) indicates that it is possible to tilt the test body to an angle such that the net moment will be zero on any selected transverse section but

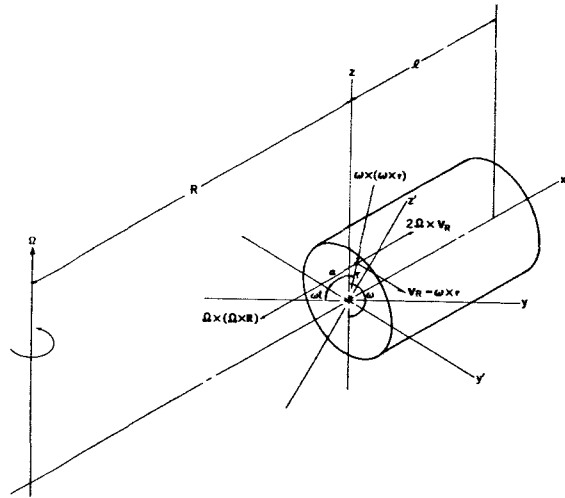


FIGURE 19.—Test body subjected to simultaneous spin and centrifuging.

cannot be rendered zero over the whole length of the test body. In addition, transverse shears will be generated. The following normalized relations (see Appendix D) have been derived to express the forces and moments existing on the cross section of a homogeneous cylindrical body:

$$\frac{C_{x'}}{qlg} = \frac{\Omega^2 R_i x'}{lg} \cos \theta - \frac{x'}{l} \sin \theta$$

$$\frac{S_{x'}}{qlg} = \frac{\Omega^2 R_i x'}{gl} \sin \theta + \frac{x'}{l} \cos \theta$$

$$\frac{M_{x'}}{ql^2 g/2} = \frac{\Omega^2 R_i x'^2}{l^2 g} \sin \theta + \frac{x'^2}{l^2} \cos \theta$$

$$- \frac{x' r_0^2 \Omega \omega \cos \theta}{l^2 g}$$

In these equations the shear and compressive forces have been normalized with respect to the body weight. The bending moment has been normalized with respect to the moment of the body weight at the base when acting as a cantilever.

A desirable type of test would be one in which no transverse shears or bendings were applied to the body as it was spun and centrifuged. As pointed out in Appendix D, these conditions can be approximated by proper choice of tilt angle. In the specific example treated therein, desirable values for forces and moments and the corresponding tilt

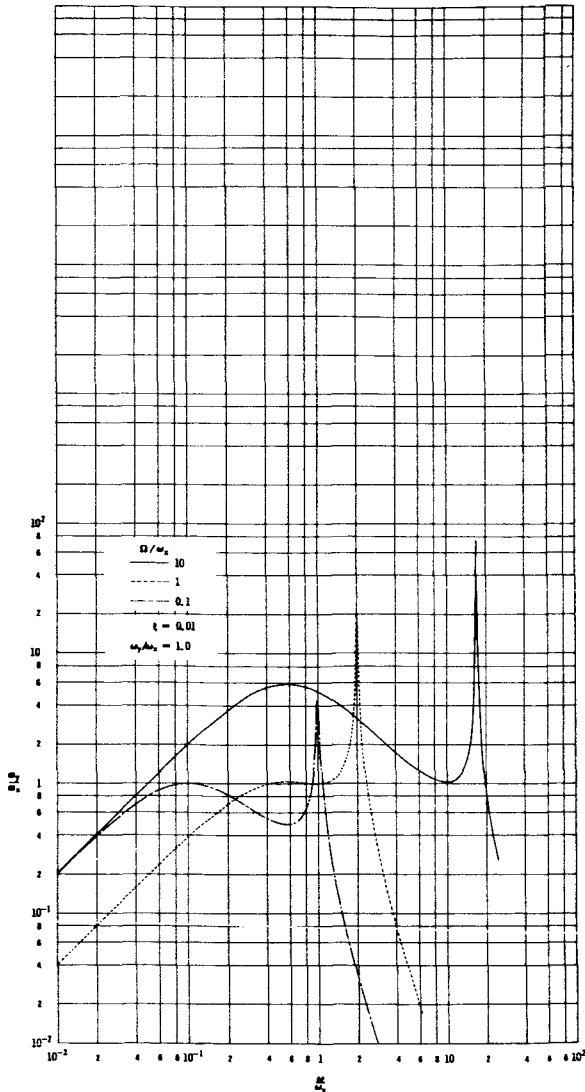


FIGURE 18.—Variation of Coriolis crosstalk frequency response with Ω/ω_x .

angles are:

For

$$\frac{C_{x'}}{q\ell g} = \frac{x'}{\ell} \frac{\Omega^2 R}{g}; \quad \theta = 0^\circ \text{ and } -6.8^\circ.$$

For

$$\frac{S_{x'}}{q\ell g} = 0; \quad \theta = -3.4^\circ.$$

For minimum moment: $\theta = -0.8^\circ$

For this particular example numerical parameters were assumed as:

$$r_0 = 12 \text{ in.},$$

$$l = 30 \text{ in.}, \quad \omega = 16.7 \text{ rad/sec},$$

$$\Omega = 3 \text{ rad/sec}, \quad R_i = 720 \text{ in.}$$

Figures 20, 21, and 22 show the normalized bending moment, compressive force, and shear force present at the tilt angles of interest as a function of position along the length of the test body.

It can be seen that compromise is necessary, since the ideal conditions cannot be simultaneously realized. If the body is tilted 0.8 degree, the bending moment will be minimized to 14.1 percent of the gravity cantilever moment. The maximum

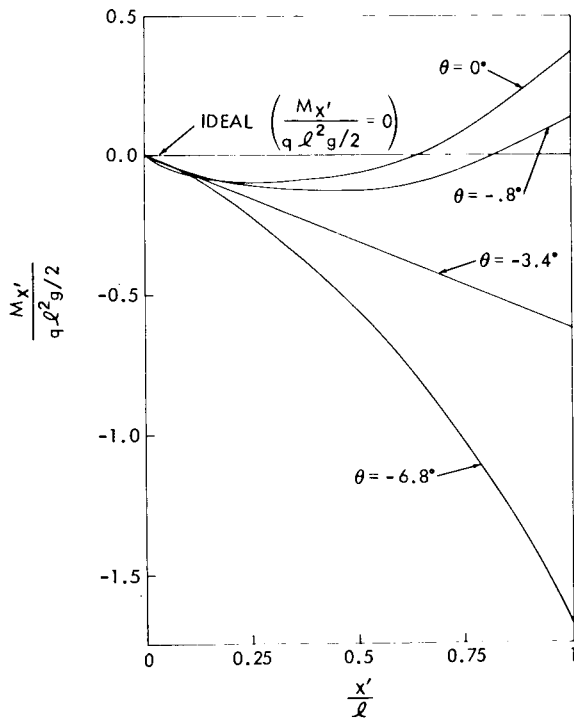


FIGURE 20.—Bending moment curves for several angles of tilt of test body.

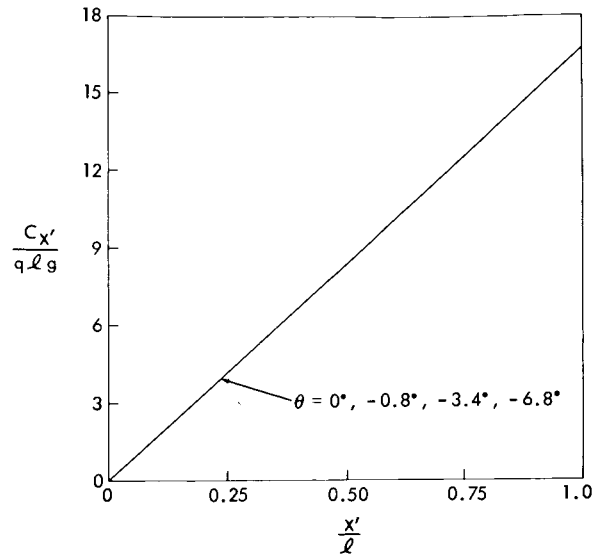


FIGURE 21.—Curves of resultant compression for several angles of tilt of test body.

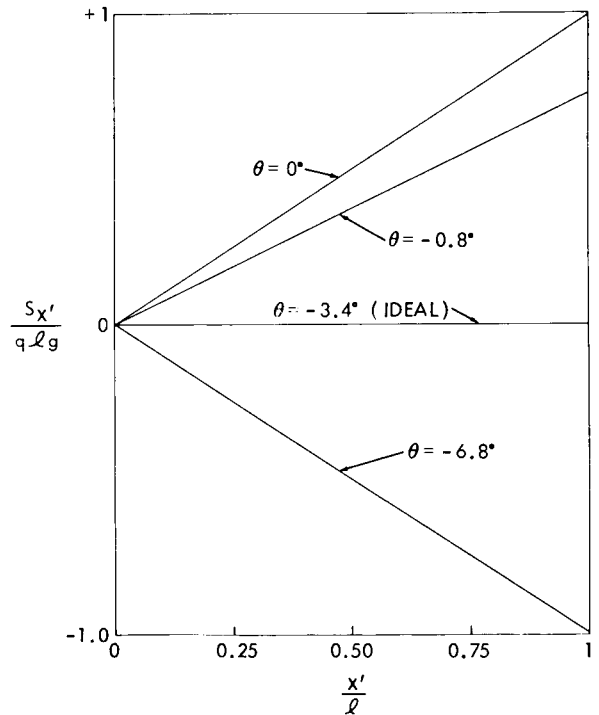


FIGURE 22.—Curves of resultant shear for several angles of tilt of test body.

shear will be 76.5 percent of the body weight. If, on the other hand, a tilt angle of -3.4 degrees is chosen, rendering shear zero over the body length, the maximum bending moment will be increased to 62.3 percent of the gravity cantilever moment.

If no tilt is applied, the shear will equal the gravity loading and the maximum bending moment will be 37.6 percent of the gravity cantilever moment. A tilt angle of -6.8 degrees is undesirable since it produces a maximum moment of 161 percent of gravity cantilever moment.

The decision as to the extent of tilt to be applied will depend on the relative sensitivities of individual test bodies to undesired shears and moments. The particular example chosen illustrates also that the desired loadings can be very sensitive to small changes in angle of tilt.

With regard to the alternating Coriolis acceleration felt by a single particle or localized mass, this will amount to $(2\Omega\omega r_0 \cos \theta)/g$ when normalized with respect to gravity. For the particular problem illustrated, this amounts to 3.11 at $\theta=0^\circ$.

VIBRATION AND SPIN COMBINED WITH CENTRIFUGING

In the case of vibration and spin combined with centrifuging, the relative velocity of any point on the test body has two components such that

$$\mathbf{V}_R = \mathbf{V}_S + \mathbf{V}_V$$

The resulting Coriolis acceleration

$$\begin{aligned} \mathbf{A}_c &= 2\boldsymbol{\Omega} \times \mathbf{V}_R \\ &= 2\boldsymbol{\Omega} \times (\mathbf{V}_S + \mathbf{V}_V) \\ &= (2\boldsymbol{\Omega} \times \mathbf{V}_S) + (2\boldsymbol{\Omega} \times \mathbf{V}_V) \end{aligned} \quad (16)$$

The two quantities on the right side of Equation 16 already have been considered individually in the preceding sections. If vibration and spin are combined with centrifuging, the resultant Coriolis acceleration may be computed as the vector sum of the accelerations produced by the spin and vibration separately.

CONCLUSIONS

Among the considerations discussed in the foregoing analysis, the following are significant.

With respect to *combined vibration and centrifuging*, the preceding analysis makes it evident that Coriolis effects ought not to be disregarded either during the design or operational phase of a combined environment facility. Since these effects are frequency-dependent, it would appear

prudent to design the mounting for the test body and shaker so that the criticals fall outside the range of operating frequencies if this is possible. If operation must be conducted in the critical range, the damping introduced will play a significant part.

In regard to *combined spin and centrifuging*, the analysis indicates the importance of the test body's inertial properties. In particular:

1. There will be no resultant linear Coriolis acceleration of a test body provided the axis of spin passes through the center of gravity of the section.

2. If the mass moments of inertia of the test body about its principal transverse axes are equal, there will be no alternating components of the resultant torque about the y and z axes. The resultant torque will have the value $I_x \Omega \times \omega$ about the y axis. Conversely, if these mass moments of inertia are unequal, there will be alternating components of torque about the y and z axes.

3. Any individual element of mass of the test body will be subjected to an alternating Coriolis force.

4. The bending moment on the test body due to the Coriolis forces may be lessened by tilting the spin axis with respect to the rotational axis of the centrifuge; however, transverse shears are generated in the process. Any decision to use this technique should be based on the practical considerations of actual structural configuration of the test body, the masses, locations, and attachment methods of internal components. The assumption of test body rigidity was made to permit a simplified analysis; however, it is quite possible that compliances may drastically affect the response to the alternating Coriolis forces. Thus, it appears that no generalization regarding the seriousness of Coriolis effects is warranted. Rather, the decision to test and the extent of testing spinning bodies on a centrifuge should be made on an individual case basis.

It should be emphasized that the conclusions drawn in this study are based on the following major assumptions: (1) Both the centrifuge and the test body are rigid, (2) the input vibratory force is sinusoidal and directed along the longitudinal axis of the test body, and (3) rotational and spin speeds are constant.

APPENDIX A

LIST OF SYMBOLS

a_f acceleration produced by application of constant force F_0 to mass m	k x-directed spring constant
a_r acceleration relative to rotating reference axes	k_1 y-directed spring constant
a_x x-directed component of absolute acceleration	l length of test body
$a_x(s)$ La Place transform of a_x	M bending moment
a_y y-directed component of absolute acceleration	M_g bending moment due to gravity
$a_y(s)$ La Place transform of a_y	$M_{x'}$ bending moment at position x'
C compressive force	m mass
$C_{x'}$ compressive force at position x'	q mass per unit length
c x-directed damping coefficient	R distance from center of rotation of centrifuge to c.g. of an element of a spinning body
c_c x-directed coefficient of critical damping	R_i distance from center of rotation to innermost end of spinning test body
c_{c1} y-directed coefficient of critical damping	R_0 distance from center of rotation of centrifuge to steady state or mean position of c.g. of test body
c_1 y-directed damping coefficient	r radial distance from axis of spin to elementary particle
F force	r_0 radial distance from axis of spin to circumference of cylindrical test body
F_c Coriolis force	S shear force
$F(s)$ La place transform of $F(t)$	$S_{x'}$ shear force at position x'
$F(t)$ time-varying external force	s La Place operator
F_0 peak value of sinusoidally varying force	t time
f frequency	V linear velocity
g acceleration of gravity	V_R particle linear velocity relative to rotating reference axes
I_x, I_y, I_z mass moment of inertia with subscript indicating axis of reference	V_s, V_v spin and vibration components, respectively, of particle velocity relative to rotating reference axes
$I_{y'z'}$ product of inertia with subscripts indicating axes of reference	x, y, z displacements with respect to a set of
i, j unit vectors along the x and y axes, respectively	
j $\sqrt{-1}$	

	axes that are fixed relative to the centrifuge and rotate with it	ρ mass density
x_{ST}	static deflection in the x direction resulting from application of force F_0	ϕ angle between Ω and V_R vectors
y_{ST}	static deflection in the y direction resulting from application of force F_0	Ω angular velocity of centrifuge
x_0	initial spring deflection	ω angular spin rate of test body; also angular frequency of sinusoidal force or motion
x', y', z'	displacements relative to axes that are fixed in the test body and move with it	ω_x rotating natural frequency in the x direction
α	angular position of element of mass	ω_y rotating natural frequency in the y direction
β	a projection of ϕ	\times vector cross-product
θ	angle of tilt of spin axis	Rotation—angular motion relative to the centrifuge axis
ξ	x-directed ratio of actual to critical damping	Spin—angular motion relative to the test body spin axis
ξ_1	y-directed ratio of actual to critical damping	

APPENDIX B

EQUATIONS OF MOTION FOR COMBINED VIBRATION AND CENTRIFUGING

In Figure 3 of the text, it is assumed that the test body and the centrifuge structure are rigid and that the only compliance is in the springs k and k_1 ; damping is indicated by the coefficients c and c_1 .

Under static conditions the body is presumed to be in the position indicated. When the centrifuge is brought up to speed, the x spring deflects an amount x_0 , so that $kx_0 = mR_0\Omega^2$. Subsequent vibration induced by a periodic force $F(t)$ causes an oscillation about the mean position. This oscillation and the associated acceleration are of interest. The angular rotational rate of the centrifuge is assumed constant.

To define the absolute acceleration of the body, the three component accelerations must be determined in accordance with Coriolis' law:

1. The acceleration the body would have if the rotating axes were to be fixed and only the relative motion of the body with respect to these axes existed.
2. The acceleration the body would have if it were fixed on the rotating axes at a given instant while the axes rotated.
3. The vector product $2\Omega \times \mathbf{V}_R$, called the Coriolis component of acceleration.

The first acceleration component is simply

$$\mathbf{a}_r = \ddot{x} \mathbf{i} + \ddot{y} \mathbf{j} .$$

The second acceleration component is the centripetal acceleration

$$\boldsymbol{\omega} \times [\boldsymbol{\omega} \times (\mathbf{R}_0 + \mathbf{x})].$$

This is directed radially and has a magnitude of $-\Omega^2 (R_0 + x)$. If we assume y to be small compared with R_0 , the x component of centripetal acceleration may be specified as $-\Omega^2 (R_0 + x)$ and the y component as $-\Omega^2 y$.

The third acceleration component, $2\Omega \times \mathbf{V}_R$, may be written as

$$2\Omega \times (\dot{x} \mathbf{i} + \dot{y} \mathbf{j}) = 2\Omega \times \dot{x} \mathbf{i} + 2\Omega \times \dot{y} \mathbf{j} .$$

The Coriolis component in the x direction has, therefore, a magnitude of $2\Omega \dot{y}$, and the component in the y direction has a magnitude of $-2\Omega \dot{x}$.

Summing the above, the component of absolute acceleration in the x direction is

$$a_x = \ddot{x} - \Omega^2 (R_0 + x) + 2\Omega \dot{y}$$

and, in the y direction,

$$a_y = \ddot{y} - \Omega^2 y - 2\Omega \dot{x} .$$

Applying Newton's second law, we obtain

$$\left. \begin{aligned} m[\ddot{x} - \Omega^2 (R_0 + x) + 2\Omega \dot{y}] &= F(t) - k(x + x_0) - c\dot{x} \\ \text{and} \\ m(\ddot{y} - \Omega^2 y - 2\Omega \dot{x}) &= -k_1 y - c_1 \dot{y} . \end{aligned} \right\} \quad (\text{B1})$$

The simultaneous solution of Equations B1 is required. Expanding the first equation gives

$$m\ddot{x} - m\Omega^2 R_0 - m\Omega^2 x + 2m\Omega \dot{y} = F(t) - k_x - kx_0 - c\dot{x} .$$

Since the average spring force in the x direction will provide the average centripetal acceleration,

$$-m\Omega^2 R_0 = -kx_0,$$

so that

$$m\ddot{x} - m\Omega^2 x + 2m\Omega\dot{y} = F(t) - kx - c\dot{x}. \quad (B2)$$

Taking the La Place transforms of Equations B1 and B2, we get

$$ms^2 y - 2m\Omega sx - m\Omega^2 y = -k_1 y - c_1 sy \quad (B3)$$

and

$$ms^2 x - m\Omega^2 x + 2m\Omega sy = F(s) - kx - csx. \quad (B4)$$

In the transformed Equations B3 and B4 and in subsequent transformed equations, the variables x and y are understood to be functions of s, not t.

Solving for y from Equation B3 gives

$$y = \frac{2m\Omega sx}{ms^2 + k_1 - m\Omega^2 + c_1 s}. \quad (B5)$$

Substituting Equation B5 into B4 yields

$$ms^2 x - m\Omega^2 x + kx + csx + \frac{4m^2 \Omega^2 s^2 x}{ms^2 + k_1 - m\Omega^2 + c_1 s} = F(s),$$

$$x \left(ms^2 + cs + k - m\Omega^2 + \frac{4m^2 \Omega^2 s^2}{ms^2 + k_1 - m\Omega^2 + c_1 s} \right) = F(s),$$

$$x \frac{[(ms^2 + cs + k - m\Omega^2)(ms^2 + k_1 - m\Omega^2 + c_1 s) + 4m^2 \Omega^2 s^2]}{ms^2 + k_1 - m\Omega^2 + c_1 s} = F(s),$$

$$\frac{x(s)}{F(s)} = \frac{ms^2 + k_1 - m\Omega^2 + c_1 s}{(ms^2 + cs + k - m\Omega^2)(ms^2 + k_1 - m\Omega^2 + c_1 s) + 4m^2 \Omega^2 s^2}, \quad (B6)$$

$$\frac{y(s)}{F(s)} = \frac{2m\Omega s}{(ms^2 + cs + k - m\Omega^2)(ms^2 + k_1 - m\Omega^2 + c_1 s) + 4m^2 \Omega^2 s^2}. \quad (B7)$$

If the input function $F(t)$ is a sinusoid $= F_0 \sin \omega t$, the steady-state inverse transformation of Equations B6 and B7 may be obtained by setting s equal to $j\omega$.

We then have

$$\frac{x(t)}{F_0} = \frac{k_1 - m\Omega^2 - m\omega^2 + j\omega c_1}{(k - m\Omega^2 - m\omega^2 + j\omega c)(k_1 - m\Omega^2 - m\omega^2 + j\omega c_1) - 4m^2 \Omega^2 \omega^2} \quad (B8)$$

and

$$\frac{y(t)}{F_0} = \frac{j 2m\Omega \omega}{(k - m\Omega^2 - m\omega^2 + j\omega c)(k_1 - m\Omega^2 - m\omega^2 + j\omega c_1) - 4m^2 \Omega^2 \omega^2}. \quad (B9)$$

In addition to the relative displacements x and y , the absolute acceleration components in the x and y directions are of interest. Considering just the alternating component of acceleration in the x direction, we have

$$a_x = \ddot{x} - \Omega^2 x + 2\Omega \dot{y}$$

and, in the y direction,

$$a_y = \ddot{y} - 2\Omega \dot{x} - \Omega^2 y$$

Their transforms are

$$\begin{aligned} a_x(s) &= s^2 x - \Omega^2 x + 2\Omega sy \\ &= x(s^2 - \Omega^2) + 2\Omega sy \end{aligned} \quad (B10)$$

and

$$\begin{aligned} a_y(s) &= s^2 y - \Omega^2 y - 2\Omega sx \\ &= y(s^2 - \Omega^2) - 2\Omega sx \end{aligned} \quad (B11)$$

Substituting Equations B6 and B7 into B10 and B11 gives

$$\frac{a_x(s)}{F(s)} = \frac{(ms^2 + k_1 - m\Omega^2 + c_1 s)(s^2 - \Omega^2) + 4m\Omega^2 s^2}{(ms^2 + cs + k - m\Omega^2)(ms^2 + k_1 - m\Omega^2 + c_1 s) + 4m^2 \Omega^2 s^2} \quad (B12)$$

$$\frac{a_y(s)}{F(s)} = \frac{2m\Omega s(s^2 - \Omega^2) - 2\Omega s(ms^2 + k_1 - m\Omega^2 + c_1 s)}{(ms^2 + cs + k - m\Omega^2)(ms^2 + k_1 - m\Omega^2 + c_1 s) + 4m^2 \Omega^2 s^2} \quad (B13)$$

$$\frac{a_y(s)}{a_x(s)} = \frac{2m\Omega s(s^2 - \Omega^2) - 2\Omega s(ms^2 + k_1 - m\Omega^2 + c_1 s)}{(ms^2 + k_1 - m\Omega^2 + c_1 s)(s^2 - \Omega^2) + 4m\Omega^2 s^2} \quad (B14)$$

Substituting $s=j\omega$ in Equations B12, B13, and B14 results in

$$\frac{a_x(t)}{F_0} = \frac{(k_1 - m\Omega^2 - m\omega^2 + j\omega c_1)(-\omega^2 - \Omega^2) - 4m\Omega^2 \omega^2}{(k - m\Omega^2 - m\omega^2 + j\omega c)(k_1 - m\Omega^2 - m\omega^2 + j\omega c_1) - 4m^2 \Omega^2 \omega^2} \quad (B15)$$

$$\frac{a_y(t)}{F_0} = \frac{-2\Omega \omega (jk_1 - \omega c_1)}{(k - m\Omega^2 - m\omega^2 + j\omega c)(k_1 - m\Omega^2 - m\omega^2 + j\omega c_1) - 4m^2 \Omega^2 \omega^2} \quad (B16)$$

$$\frac{a_y(t)}{a_x(t)} = \frac{-2\Omega \omega (jk_1 - \omega c_1)}{(k_1 - m\Omega^2 - m\omega^2 + j\omega c_1)(-\omega^2 - \Omega^2) - 4m\Omega^2 \omega^2} \quad (B17)$$

The conditions that will produce extreme values of the quantities defined by Equations B6, B7, B15, B16, and B17 are of interest.

The equations for $x(t)/F_0$, $y(t)/F_0$, $a_x(t)/F_0$, and $a_y(t)/F_0$ all have the same denominator. Equating this denominator to zero gives

$$(k - m\Omega^2 - m\omega^2 + j\omega c)(k_1 - m\Omega^2 - m\omega^2 + j\omega c_1) - 4m^2 \Omega^2 \omega^2 = 0$$

The real and imaginary parts must separately equal 0. For the real part,

$$(k - m\Omega^2 - m\omega^2)(k_1 - m\Omega^2 - m\omega^2) - \omega^2 cc_1 - 4m^2 \Omega^2 \omega^2 = 0$$

or

$$m^2 \omega^4 - [m(k + k_1) + 2m^2 \Omega^2 + cc_1] \omega^2 + (k - m\Omega^2)(k_1 - m\Omega^2) = 0 .$$

Solving the quadratic in ω^2 yields

$$\begin{aligned} \omega^2 &= \frac{m(k + k_1) + 2m^2 \Omega^2 + cc_1 \pm \sqrt{[m(k + k_1) + 2m^2 \Omega^2 + cc_1]^2 - 4m^2 (k - m\Omega^2)(k_1 - m\Omega^2)}}{2m^2} \\ &= \frac{k + k_1 + 2m\Omega^2 \pm \sqrt{(k_1 - k)^2 + 8m\Omega^2 (k + k_1)}}{2m} . \end{aligned} \quad (\text{B18})$$

For the imaginary part,

$$\omega c (k_1 - m\Omega^2 - m\omega^2) + \omega c_1 (k - m\Omega^2 - m\omega^2) = 0 ,$$

$$c = c_1 = 0 .$$

The frequencies that produce zeros of the numerators are also of interest. For x/F_0 , we require that

$$\left. \begin{aligned} k_1 - m\Omega^2 - m\omega^2 + j\omega c_1 &= 0 , \\ k_1 - m\Omega^2 - m\omega^2 &= 0 , \\ \omega &= \sqrt{\frac{k_1 - m\Omega^2}{m}} , \\ j\omega c_1 &= 0 , \\ c_1 &= 0 ; \end{aligned} \right\} \quad (\text{B19})$$

for $y(t)/F_0$,

$$j2m\Omega\omega = 0 ,$$

$$\omega = 0 ;$$

and, for $a_x(t)/F_0$,

$$(k_1 - m\Omega^2 - m\omega^2 + j\omega c_1)(-\omega^2 - \Omega^2) - 4m\Omega^2 \omega^2 = 0 .$$

For the real part,

$$m\omega^4 - (k_1 + 2m\Omega^2)\omega^2 - \Omega^2(k_1 - m\Omega^2) = 0 ,$$

$$\omega^2 = \frac{k_1 + 2m\Omega^2 \pm \sqrt{(k_1 + 2m\Omega^2)^2 + 4m\Omega^2(k_1 - m\Omega^2)}}{2m} ; \quad (\text{B20})$$

and, for the imaginary part,

$$\omega c_1 (-\omega^2 - \Omega^2) = 0 ,$$

$$c_1 = 0 .$$

For $a_y(t)/F_0$, we require that

$$-2\Omega\omega(jk_1 - \omega c_1) = 0 .$$

For the imaginary part,

$$-j2\Omega\omega k_1 = 0 ,$$

$$\omega = 0 ;$$

and, for the real part,

$$2\Omega\omega^2 c_1 = 0 ,$$

$$\omega = 0 .$$

Since the zeros of numerators and denominators do not occur at the same frequency, we may conclude the following:

1. At the frequencies defined by Equation B18, $x(t)/F_0$, $y(t)/F_0$, $a_x(t)/F_0$, and $a_y(t)/F_0$ all will have infinite values provided no damping is present.
2. At the frequency defined by Equation B19, $x(t)/F_0$ will be zero provided no damping is present.
3. At the frequency defined by Equation B20, $a_x(t)/F_0$ will be zero provided no damping is present.
4. The crosstalk ratio, $a_y(t)/a_x(t)$, will be infinite at the frequencies defined by Equation B20 provided no damping is present.

APPENDIX C

DERIVATION OF EXPRESSIONS FOR CORIOLIS FORCES AND MOMENTS DUE TO COMBINATION OF SPIN AND CENTRIFUGING

Consider the combined centrifugal and spin of a body of arbitrary shape as shown by Figure C1. The y and z axes are fixed system references, while the y' and z' axes are fixed in the body and spin with it. The spin axis x' is perpendicular to both sets of axes and passes through their common intersection point. In addition, the spin axis is assumed to be perpendicular to the centrifuge rotational axis. The origin of axes is located at some arbitrary point within the test body, not necessarily coincident with the center of gravity.

The Coriolis force on an elementary particle is

$$\begin{aligned} dF_c &= dm \, 2\Omega \times V_R = dm \, 2\Omega \omega r \sin(\omega t + \alpha) \\ &= dm \, 2\Omega \omega r [\sin \omega t \cos \alpha + \cos \omega t \sin \alpha] \end{aligned}$$

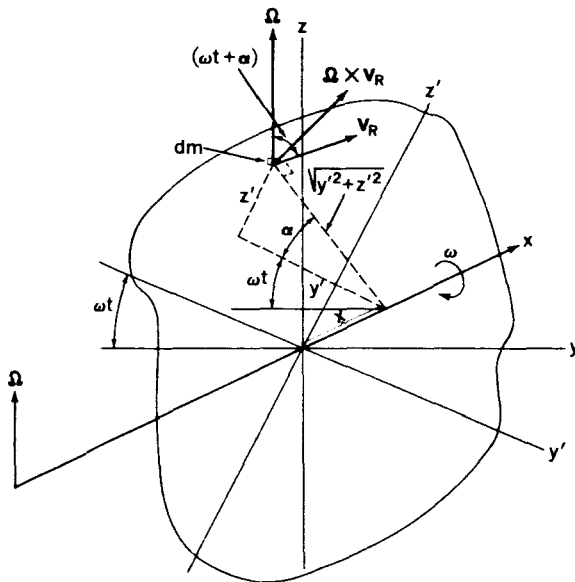


FIGURE C1.—Body of arbitrary shape subjected to combined spin and centrifuging with Ω and ω axes orthogonal.

$$dm = \rho \, dy' \, dz' \, dx' ,$$

$$r \cos \alpha = -y' ,$$

$$r \sin \alpha = z' ,$$

$$dF_c = -2\Omega\omega \sin \omega t \, \rho y' \, dy' \, dz' \, dx' + 2\Omega\omega \cos \omega t \, \rho z' \, dz' \, dy' \, dx' ,$$

$$F_c = -2\Omega\omega \sin \omega t \iiint \rho y' \, dy' \, dz' \, dx' + 2\Omega\omega \cos \omega t \iiint \rho z' \, dz' \, dy' \, dx' .$$

The triple integrals in the equation above represent the first moment of gravity about each of the axes y' and z' . If the integral of each of these

moments is zero, the spin axis must pass through the center of gravity of the body and there will be no resultant linear Coriolis force on the body.

Each elementary Coriolis force will produce a moment about the coordinate axes y' and z' :

$$\begin{aligned} dM_{y'} &= -dF_c z' \\ &= +2\Omega\omega \sin \omega t \iiint \rho y' z' dy' dz' dx' - 2\Omega\omega \cos \omega t \iiint \rho z'^2 dz' dy' dx'. \end{aligned}$$

By axial transformation,

$$dM_y = dM_{y'} \cos \omega t + dM_{z'} \sin \omega t,$$

$$dM_{z'} = -dM_{y'} \sin \omega t + dM_z \cos \omega t.$$

Since

$$dM_{y'} = -dF_c z'$$

and

$$dM_{z'} = dF_c y',$$

then

$$\begin{aligned} dM_y &= -dF_c z' \cos \omega t + dF_c y' \sin \omega t \\ &= 2\Omega\omega \sin \omega t \rho y' dy' z' dz' dx' \cos \omega t - 2\Omega\omega \cos^2 \omega t \rho z'^2 dz' dy' dx' \\ &\quad - 2\Omega\omega \sin^2 \omega t \rho y'^2 dy' dz' dx' + 2\Omega\omega \sin \omega t \cos \omega t \rho z' dz' y' dy' dx', \\ M_y &= 2\Omega\omega \sin 2\omega t \iiint y' z' \rho dy' dz' dx' - 2\Omega\omega \cos^2 \omega t \iiint z'^2 \rho dy' dz' dx' \\ &\quad - 2\Omega\omega \sin^2 \omega t \iiint y'^2 \rho dy' dz' dx'. \end{aligned} \tag{C1}$$

This is, in general, a variable.

By definition,

$$I_{x'} = \iiint (y'^2 + z'^2) \rho dy' dz' dx',$$

$$I_{y'} = \iiint (x'^2 + z'^2) \rho dy' dz' dx',$$

$$I_{z'} = \iiint (x'^2 + y'^2) \rho dy' dz' dx'.$$

and

$$I_{y'z'} = \iiint y' z' \rho dy' dz' dx'.$$

If $I_{x'}$, $I_{y'}$, and $I_{z'}$ are about the principal axes of the body, then $I_{y'z'} = 0$. If, furthermore, $I_{y'} = I_{z'}$, then

$$\iiint y'^2 \rho dy' dz' dx' = \iiint z'^2 \rho dy' dz' dx'.$$

Making these substitutions into Equation C1, we have

$$M_y = -2\Omega\omega \iiint \frac{y'^2 + z'^2}{2} \rho dy' dz' dx' = -I_{x'} \Omega\omega.$$

where

$$\sin \beta = \frac{\sin \theta}{\sqrt{\cos^2 \theta \sin^2 (\omega t + a) + \sin^2 \theta}} .$$

Thus, the coplanar component is

$$dm \, 2\Omega r \sin \theta .$$

In determining the resultant moment due to the Coriolis forces acting on the elementary section, we may make use of the previous analysis to obtain the following results. The perpendicular component of F_c is

$$- 2\Omega \omega \cos \theta \sin \omega t \iiint \rho y' \, dy' \, dz' \, dx' + 2\Omega \omega \cos \theta \cos \omega t \iiint \rho z' \, dz' \, dy' \, dx' .$$

By the same reasoning as in the previous analysis, there will be no resultant Coriolis force perpendicular to the cross section provided the body is

spun about an axis through the center of gravity.

The moments due to the perpendicular component of Coriolis force are

$$\begin{aligned} M_y = 2\Omega \omega \cos \theta & \left[\sin 2\omega t \iiint y' z' \rho \, dy' \, dz' \, dx' - \cos^2 \omega t \iiint z'^2 \rho \, dy' \, dz' \, dx' \right. \\ & \left. - \sin^2 \omega t \iiint y'^2 \rho \, dy' \, dz' \, dx' \right] \end{aligned}$$

and

$$\begin{aligned} M_z = 2\Omega \omega \cos \theta I_{y'z'} & (\cos^2 \omega t - \sin^2 \omega t) \\ & + \Omega \omega \cos \theta \sin 2\omega t (I_{y'} - I_{z'}) . \end{aligned}$$

Under the conditions that the body is spun about an axis through its center of gravity and that it has symmetry such that $I_{y'} = I_{z'}$,

$$I_{y'z'} = 0 .$$

Then,

$$\begin{aligned} M_y & = - I_x \Omega \omega \cos \theta \\ & = I_x \Omega \times \omega \end{aligned}$$

and

$$M_z = 0 .$$

The coplanar component of the elementary Coriolis force ($dm 2\Omega \times V_R \sin \beta$) is radially

directed. This results in $M_x = M_y = M_z = 0$ as far as this component of the Coriolis force is concerned. The y and z force components may be evaluated as follows:

Elementary coplanar y component

$$\begin{aligned} & = dm \, 2\Omega r \sin \theta \cos(\omega t + a) \\ & = dm \, 2\Omega \omega \sin \theta [r \cos a \cos \omega t - r \sin a \sin \omega t] \end{aligned}$$

but

$$r \sin a = z'$$

and

$$r \cos a = y' .$$

$$\text{Total y component} = 2\Omega \omega \sin \theta \left[\cos \omega t \iiint \rho y' \, dy' \, dz' \, dx' - \sin \omega t \iiint \rho z' \, dy' \, dz' \, dx' \right] .$$

In similar fashion:

$$\text{Total z component} = -2\Omega\omega \sin\theta \left[\sin\omega t \iiint \rho y' dy' dz' dx' + \cos\omega t \iiint \rho z' dy' dz' dx' \right] .$$

If the spin axis passes thru the center of gravity, both these components will vanish.

With the directions of rotation of Ω and ω as shown in Figure C2, the force will be directed radially inward. If either rotation is reversed, the force will be directed radially outward.

APPENDIX D

LOADS PRODUCED BY COMBINATION OF SPIN AND CENTRIFUGING

Assume a uniform homogeneous test body being spun with constant angular velocity ω while being centrifuged at constant angular velocity Ω , as shown in Figure D1.

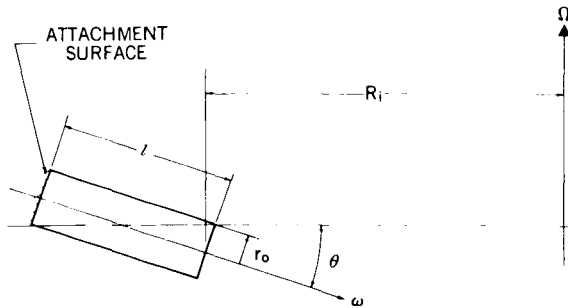


FIGURE D1.—Combined spin and centrifuging of uniform homogeneous body.

To determine the forces acting on this body, we may consider an elementary transverse slice of width dx' located at a distance x' from the in-board end, as shown in Figure D2.

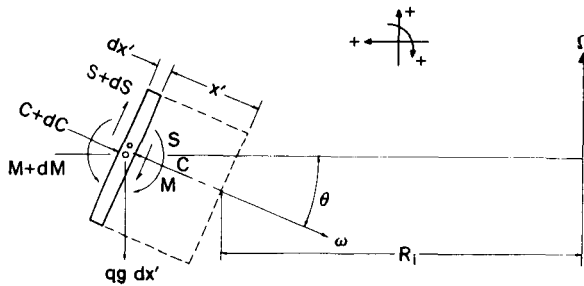


FIGURE D2.—Elementary transverse slice of spinning test body.

In the horizontal direction,

$$q dx' \Omega^2 (R_1 + x' \cos \theta) = dC \cos \theta + dS \sin \theta. \quad (D1)$$

In the vertical direction,

$$- qg dx' - dC \sin \theta + dS \cos \theta = 0. \quad (D2)$$

In rotation about a transverse axis through the center of gravity,

$dI \Omega \omega \cos \theta$ + centrifugal component of moment

$$= (S + dS) \frac{dx'}{2} + \frac{S dx'}{2} - dM.$$

The centrifugal component of moment may be derived as follows (see Figure D3).

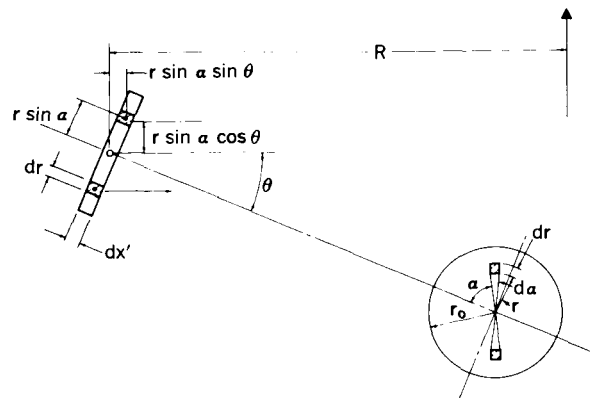


FIGURE D3.—Centripetal acceleration on symmetrically located masses of spinning body.

The moment about center due to the element of mass at radius less than R is

$$dM_1 = \Omega^2 \frac{qr}{\pi r_0^2} da dr dx' (R -$$

$$r \sin \alpha \sin \theta) r \sin \alpha \cos \theta.$$

The moment about center due to the element of mass at radius greater than R is

$$dM_2 = \Omega^2 \frac{qr}{\pi r_0^2} da dr dx'$$

$$(R + r \sin \alpha \sin \theta) r \sin \alpha \cos \theta.$$

The moment due to the pair of elementary masses is

$$\begin{aligned} dM_2 - dM_1 &= 2\Omega^2 \frac{qr^2}{\pi r_0^2} dr \sin \alpha da dx' \sin \theta r \sin \alpha \cos \theta \\ &= 2\Omega^2 \frac{q}{\pi r_0^2} \sin \theta \cos \theta r^3 dr \sin^2 \alpha da dx' . \end{aligned}$$

Integrating,

$$\begin{aligned} \text{Moment} &= 2\Omega^2 \frac{q}{\pi r_0^2} \sin \theta \cos \theta dx' \int_{r=0}^{r=r_0} \int_{\alpha=0}^{\alpha=\pi} r^3 dr \sin^2 \alpha da \\ &= 2\Omega^2 \frac{q}{\pi r_0^2} \sin \theta \cos \theta dx' \left[\frac{r^4}{4} \cdot \left(\frac{\alpha}{2} - \frac{\sin 2\alpha}{4} \right) \right]_{r=0, \alpha=0}^{r=r_0, \alpha=\pi} \\ &= 2\Omega^2 \frac{q}{\pi r_0^2} \sin \theta \cos \theta dx' \left[\frac{r_0^4}{4} \left(\frac{\pi}{2} \right) \right] \\ &= \Omega^2 \frac{q}{\pi r_0^2} \frac{dx' \pi r_0^4}{4} \cdot \sin \theta \cos \theta \\ &= \frac{\Omega^2 q dx' r_0^2}{4} \sin \theta \cos \theta . \end{aligned}$$

The moment equation now may be written as

$$dI \Omega \cos \theta - \Omega^2 \frac{q dx' r_0^2}{4} \sin \theta \cos \theta = S dx' - dM . \quad (D3)$$

Integrating Equations D1, D2, and D3,

$$I_x \Omega \cos \theta - \frac{\Omega^2 q x' r_0^2}{4} \sin \theta \cos \theta = \int S dx' - M_x , \quad (D4)$$

$$q \Omega^2 R_i x' \cos \theta + q \frac{\Omega^2 \cos^2 \theta x'^2}{2} = C_x \cos^2 \theta + S_x \sin \theta \cos \theta , \quad (D5)$$

$$- q g x' \sin \theta = C_x \sin^2 \theta - S_x \cos \theta \sin \theta . \quad (D6)$$

Unknowns are C_x , S_x , and M_x .

From Equations D5 and D6,

$$C_{x'} = q\Omega^2 R_i x' \cos \theta + \frac{q\Omega^2 \cos^2 \theta x'^2}{2} - qgx' \sin \theta \quad (D7)$$

$$S_{x'} = C_{x'} \frac{\sin \theta}{\cos \theta} + \frac{qgx'}{\cos \theta}$$

$$\begin{aligned} S_{x'} &= q\Omega^2 R_i x' \sin \theta + \frac{q\Omega^2 x'^2}{2} \sin \theta \cos \theta - \frac{qgx' \sin^2 \theta}{\cos \theta} + \frac{qgx'}{\cos \theta} \\ &= q\Omega^2 R_i x' \sin \theta + \frac{q\Omega^2 x'^2 \sin \theta \cos \theta}{2} + qgx' \cos \theta \end{aligned} \quad (D8)$$

$$\begin{aligned} M_{x'} &= \frac{q\Omega^2 R_i x'^2 \sin \theta}{2} + \frac{q\Omega^2 x'^3 \sin \theta \cos \theta}{6} + \frac{qgx'^2 \cos \theta}{2} - I_{x'} \Omega^2 \omega \cos \theta \\ &\quad + \frac{\Omega^2 qx' r_0^2}{4} \sin \theta \cos \theta \end{aligned} \quad (D9)$$

In many centrifuge operations,

$$R_i \gg r_0,$$

$$R_i \gg x',$$

and

$$\Omega < \omega.$$

If these conditions are satisfied, Equations D7, D8, and D9 become

$$C_{x'} = q\Omega^2 R_i x' \cos \theta - qgx' \sin \theta \quad (D10)$$

$$S_{x'} = q\Omega^2 R_i x' \sin \theta + qgx' \cos \theta \quad (D11)$$

$$M_{x'} = \frac{q\Omega^2 R_i x'^2 \sin \theta}{2} + \frac{qgx'^2 \cos \theta}{2} - \frac{qx' r_0^2}{2} \Omega \omega \cos \theta \quad (D12)$$

In this case, R_i may be taken as the radius to the center of gravity of the test body.

To provide a basis for comparison, the forces

and moments have been normalized with respect to gravity loadings, resulting in the following relations:

$$\frac{C_{x'}}{qlg} = \frac{\Omega^2 R_i x' \cos \theta}{lg} - \frac{x' \sin \theta}{l} \quad (D13)$$

$$\frac{S_{x'}}{qlg} = \frac{\Omega^2 R_i x' \sin \theta}{lg} + \frac{x' \cos \theta}{l} \quad (D14)$$

$$\frac{M_{x'}}{ql^2 g/2} = \frac{\Omega^2 R_i x'^2 \sin \theta}{l^2 g} + \frac{x'^2 \cos \theta}{l^2} - \frac{x' r_0^2 \Omega \omega \cos \theta}{l^2 g} \quad (D15)$$

For most realistic testing, it is desirable to achieve the conditions:

$$\frac{C_{x'}}{q'g} = \frac{qx' \Omega^2 R_i}{q'lg} = \frac{x'}{l} \frac{\Omega^2 R_i}{g} \quad (D16)$$

$$\frac{S_{x'}}{q'lg} = 0 \quad (D17)$$

$$\frac{M_{x'}}{ql^2 g/2} = 0 \quad (D18)$$

These conditions are to be satisfied for $0 \leq x \leq l$. It is of interest to determine the extent to which these conditions can be realized by varying θ . To meet the condition of Equation D16 requires

$$\frac{\Omega^2 R_i x' \cos \theta}{lg} - \frac{x' \sin \theta}{l} = \frac{x'}{l} \frac{\Omega^2 R_i}{g}$$

Solutions of this equation are

$$\theta = 0$$

$$\theta = \arcsin \frac{-2}{(g/\Omega^2 R_i) + (\Omega^2 R_i/g)}$$

To meet the condition of Equation D17 requires

$$\frac{\Omega^2 R_i x' \sin \theta}{lg} + \frac{x' \cos \theta}{l} = 0$$

maximum and minimum values over the length of the body. Stationary values are found where

$$\frac{\partial}{\partial x} \frac{M_{x'}}{ql^2 g/2} = 0$$

$$\frac{\partial}{\partial x} \frac{M_{x'}}{ql^2 g/2} = \frac{2\Omega^2 R_i x' \sin \theta}{l^2 g} + \frac{2x' \cos \theta}{l^2} - \frac{r_0^2 \Omega \omega \cos \theta}{l^2 g} = 0$$

from which

$$x' = \frac{r_0^2 \Omega \omega \cos \theta}{2\Omega^2 R_i \sin \theta + 2g \cos \theta}$$

If this value of x' is substituted into Equation D15, we obtain a stationary value of $M_{x'}/\frac{ql^2 g}{2}$:

$$\frac{M_{x'}}{gl^2 g/2} = \frac{-(r_0^2 \Omega \omega \cos \theta)^2}{4l^2 g (\Omega^2 R_i \sin \theta + g \cos \theta)}$$

or

$$\tan \theta = \frac{-g}{\Omega^2 R_i}$$

To meet the conditions of Equation D18 requires

$$\frac{\Omega^2 R_i x'^2 \sin \theta}{l^2 g} + \frac{x'^2 \cos \theta}{l^2} - \frac{x' r_0^2 \Omega \omega \cos \theta}{l^2 g} = 0$$

Here the solution is

$$\tan \theta = \frac{r_0^2 \Omega \omega - gx'}{\Omega^2 R_i x'}$$

In this particular case, $\tan \theta$ is a function of x' , indicating that no unique angle exists which will make the bending moment vanish over the whole length of the test body. We may, however, minimize the moment by adjusting θ to equal the

This may be equated to

$$\frac{M_e}{ql^2 g/2} = \frac{\Omega^2 R_i \sin \theta}{g} + \cos \theta - \frac{r_0^2 \Omega \omega \cos \theta}{lg}$$

and the resultant expression solved for θ .

We obtain from this a quadratic expression in $\tan \theta$ as follows:

$$\tan^2 \theta + \left(\frac{2g}{\Omega^2 R_i} - \frac{r_0^2 \Omega \omega}{l \Omega^2 R_i} \right) \tan \theta - \left(\frac{r_0^2 \Omega \omega}{2l \Omega^2 R_i} \right)^2 + \left(\frac{g}{\Omega^2 R_i} \right)^2 - \frac{r_0^2 \Omega \omega}{l \Omega^2 R_i} \cdot \frac{g}{\Omega^2 R_i} = 0$$

As a specific example, assume a test body being spun while being rotated on a centrifuge under the following conditions:

$$\omega = 16.7 \text{ rad/sec} ,$$

$$\Omega = 3 \text{ rad/sec} ,$$

$$l = 30 \text{ in.} ,$$

$$r_0 = 12 \text{ in.} ,$$

$$R_i = 720 \text{ in.}$$

Under these circumstances, the angles of tilt

of interest are:

$$\text{For } \frac{C_{x'}}{q/g} - \frac{x'}{l} \frac{\Omega^2 R_i}{g} , \quad \theta = 0^\circ \text{ and } -6.8^\circ .$$

$$\text{For } \frac{S_{x'}}{q/g} = 0 , \quad \theta = -3.4^\circ .$$

$$\text{For minimum moment, } \theta = -0.8^\circ .$$

Figures 20, 21, and 22 (on page 590) in the test show the normalized bending moment, compressive force, and shear force resulting from various values of axial tilt of the test body, and compare these with the ideal.

PROBLEMS IN THE CONSTRUCTION OF A SPACE ENVIRONMENT SIMULATOR*

DANA S. COPE

Goddard Space Flight Center is constructing a space environment simulator which has advanced the state of the art of environment simulation. Problems concerning the material and construction of reflectors are presented in some detail. At present this facility provides oil-diffusion, cryogenic pumping, and controlled wall temperatures. Solar hardware fabrication is in the process of completion.

INTRODUCTION

When it is determined to build a new facility extending the state-of-the-art, many decisions must be made on the basis of the best information available tempered with human judgment. As the design and construction develop, many problems arise which are more difficult to solve than was first anticipated, as well as some thought to be important which solve themselves. When this occurs the facility has often progressed to the point where it is difficult or impossible to back up or start over. Therefore compromises must be made as best seen from the new position.

This presentation is a discussion of how some of these problems were met in the construction of a space environment simulator recently put in operation at Goddard Space Flight Center (Figures 1 and 2). At present the facility provides oil diffusion and cryogenic pumping and controlled wall temperatures but not solar simulation. Solar hardware fabrication has not yet been completed.

REQUIREMENTS

1. Size—a work space $27\frac{1}{2}$ ft. in diameter \times 40 ft. in height
2. Pressure—an ultimate pressure with full solar simulation of 1×10^{-8} torr in 24 hr
3. Shroud temperatures— 100°K and -65° to $+100^{\circ}\text{C}$
4. Solar simulation—an intensity of 50 to 275 watts/ft.² and a uniformity of ± 10 percent over any square foot
5. Collimation—a ± 4 degree half angle

6. Spectral Distribution—equivalent to zero air mass solar radiation from 0.3 to 4.0 microns
7. Area covered—a 20 ft. diameter circle.

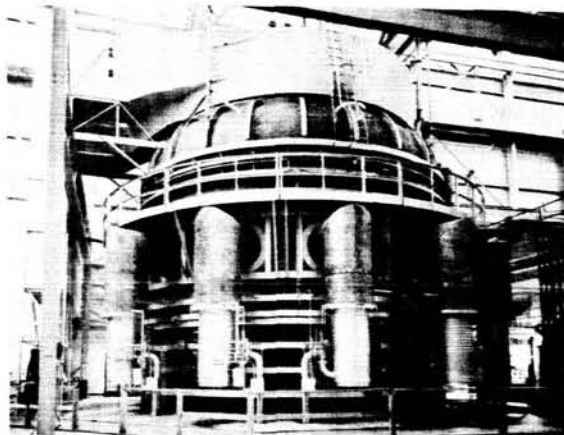


FIGURE 1.—Space environment simulator.

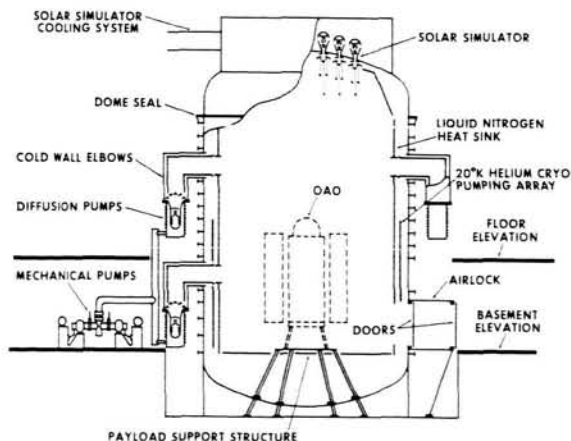


FIGURE 2.—Cross-section of the simulator.

*Published as NASA Technical Note D-1917, March 1964.

THE CHAMBER

The chamber walls are constructed of 304 stainless steel with No. 4 mill finish to keep outgassing to a minimum. Stainless steel was chosen rather than stainless clad steel to avoid both the possible welding problems and the effect of LN_2 spillage and still maintain the desired interior finish. The chamber is externally reinforced with carbon steel. All welds except penetrations are full penetration welds and have been radiographed to insure against porosity within the weld. We now believe that all welds should be full penetration multi-pass welds.

The head is removable to provide access to the tank for placement of the test loads (Figure 3).



FIGURE 3.—Head handler and overhead crane.

This is accomplished by means of a specially designed head handler (Figure 4) that lifts the head 12 in. and then moves it horizontally on its own rails a distance of 40 ft. The spacecraft to be tested is placed in the vessel by means of a 15 ton overhead bridge crane. An air lock (Figure 5) installed at the payload table level provides access to the normal working level so that a man in a space suit can enter to make repairs to a satellite and save an expensive test. The chamber has not as yet been man-rated.

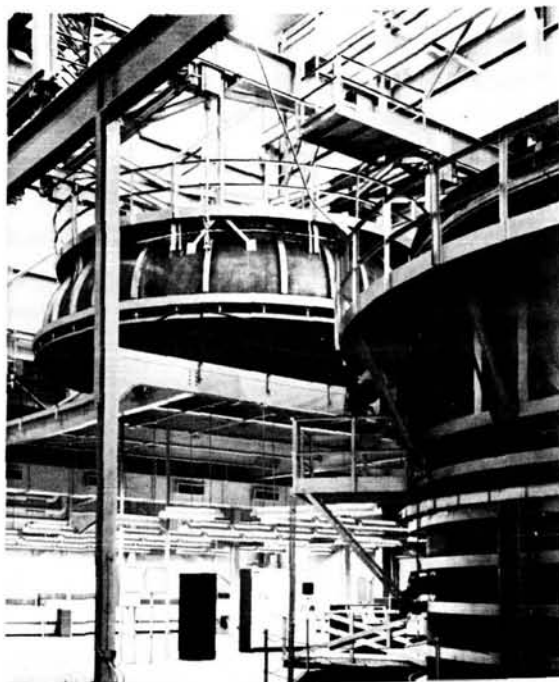


FIGURE 4.—Head handler removed from the chamber.

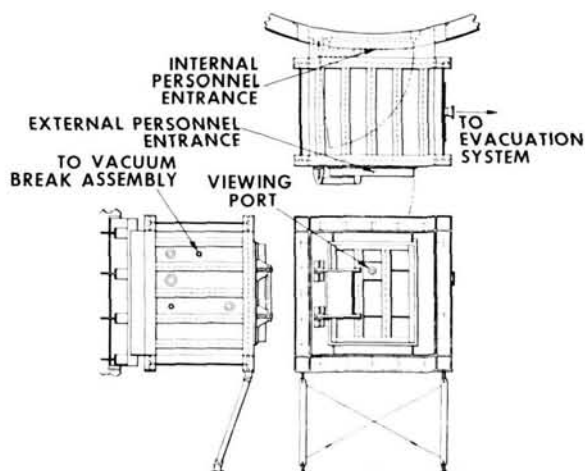


FIGURE 5.—Details of the air lock.

The head is supported on flanges welded to the shell and head (Figure 6). Thin plates of stainless steel are welded to these flanges and to the seal flanges to give flexibility for sealing. This avoids all the problems of main vessel distortion with temperature or time. The lower seal flange is rigidly supported at 36 equally spaced points and the upper flange is adjusted

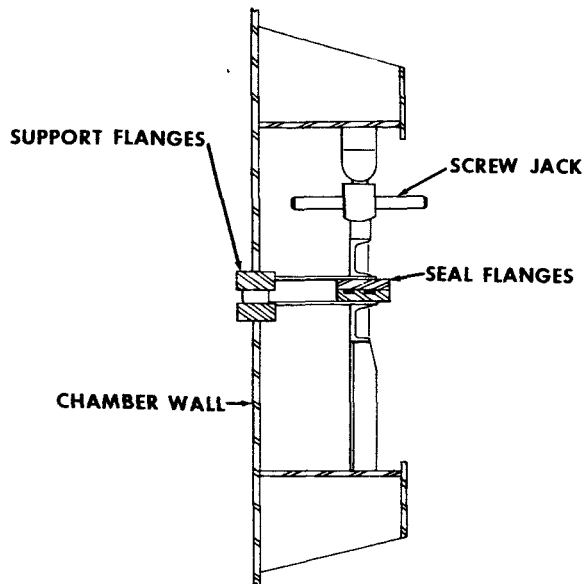


FIGURE 6.—Head seal section.

in relation to it by means of jacks placed at the corresponding 36 points. The seal between these flanges consists of pumped double 3/8 in. Buna N "o" rings. The head seal has worked out very well and has caused no leak problem. All penetrations over 12 in. also have pumped double "o" rings. The seals for the solar penetrations will be discussed later.

A removable floor is used for protection of the bottom heat sink when personnel are working inside the chamber. This consists of pie-shaped lightweight aluminum sections covered with checkered plate. Hard points have been located around the walls and in the dome to support any expected test load.

Leak testing of the chamber took approximately 6 weeks rather than the 2 weeks originally scheduled. The most important lesson learned was that on a tank this large it is impractical to try on the first attempt to pump to a low enough pressure to use the mass spectrometer. A complete and thorough subsystem and component leak check should be made. The final leak rate was 2.5×10^{-5} atm.-cc/sec.

Only one major change in design would be suggested for another chamber of comparable size. The inner door seal is not dependable when the air lock is evacuated (which leaves essentially the same pressure on both sides). This is true

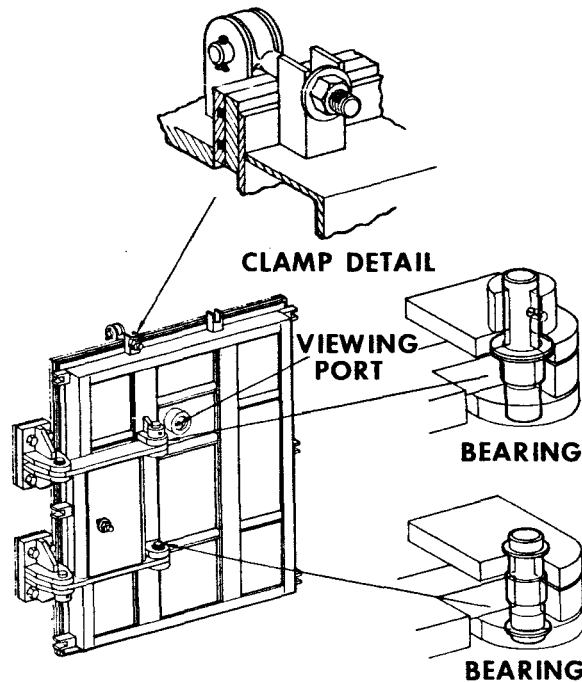


FIGURE 7.—Air lock door.

even though this door (Figure 7) is secured with ten clamping bolts. Since this clamping arrangement also makes the operation of the door slow, a better design should be found.

It is interesting to note that there were very few leaks in the field welds but most of the shop welds had to be redone.

The Vacuum System

The design of the vacuum system as originally designed is based on the following equipment, installed as shown schematically in Figure 8:

1. Twenty-five LN₂ cooled elbows. These will eliminate any possibility of a heat load on the test object because of their optical density.
2. Twenty-five 32 in. diffusion pumps rated at 50,000 liters/sec each. These pumps are so designed that heating elements and pump oil can be changed while the system is under vacuum.
3. Twenty-five concentric disc water-cooled baffles between the LN₂ cooled elbows and the diffusion pumps to prevent gross backstreaming.

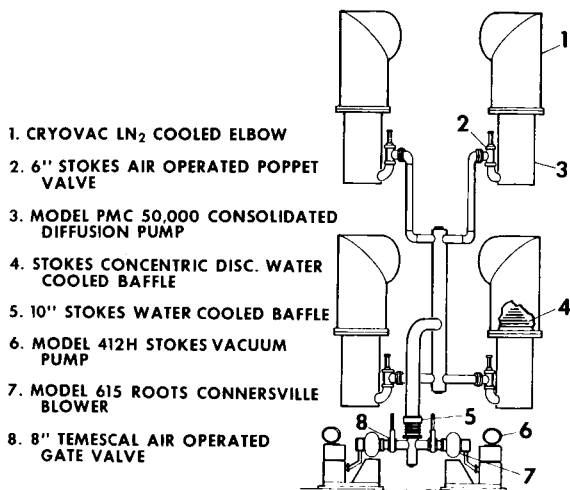


FIGURE 8.—Vacuum pumping system schematic: (1) LN_2 cooled elbow; (2) 6 in. air operated poppet valve; (3) Model PMC 50,000 diffusion pump; (4) concentric disc water cooled baffle; (5) 10 in. water cooled baffle; (6) Model 412H vacuum pump; (7) Model 615 roots blower; (8) 8 in. air-operated gate valve.

4. Twenty-five 6 in. pneumatically operated poppet valves in the forearms of the diffusion pumps.
5. Six 10 in. water-cooled optically dense baffles installed between the diffusion and roughing pumps.
6. Twelve roughing vacuum pump units each consisting of a blower and mechanical high-vacuum pump.
7. Twelve 8 in. pneumatically operated gate valves to isolate each of the twelve roughing vacuum pump units.

In the early design phase it became necessary to economize and this was done by eliminating eight diffusion pumps and four vacuum pump units and the associated LN_2 elbows, water baffles, and valves. Provision was made for adding this equipment at a later date.

The original design was based on reaching 1×10^{-8} torr in 24 hr. when used with the LN_2 heat sinks and the GHe cold wall. With the elimination of these pumps, the design specification was changed to 5×10^{-8} in 48 hours.

This system has been installed with no troubles other than those normally found on jobs of this size. However, a complete leak check of all components is considered extremely valuable in

the performance of the overall leak check of the chamber.

Since placing the chamber in operation several pump down tests have been run and much lower pressures have been reached in much less time than specified. As the pump down curve (Figure 9) shows a pressure of 9.6×10^{-10} torr was reached

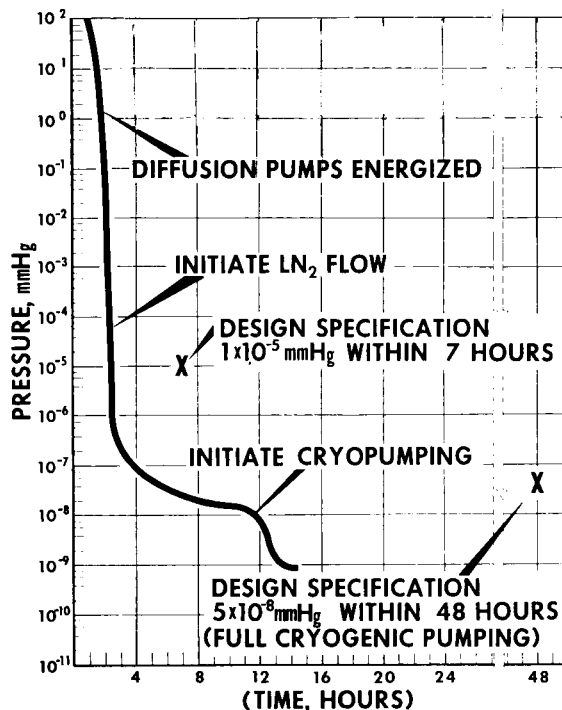


FIGURE 9.—Pump down curve.

in 14 hr. This compares with the requirements of 5×10^{-8} torr in 48 hr. as called for in the specifications.

Wall Temperature Control

To simulate outer space it is necessary to cover all walls, floor, and top with a cold wall at $< 100^\circ\text{K}$. This temperature will give the equivalent of a 1 percent return of energy radiated by the satellite under test. The cold wall was constructed using tube-in-sheet aluminum panels assembled to cover as completely as possible side walls, bottom, and top, except for the area used for the solar simulator which has a separate LN_2 cooling system. The LN_2 panels were mounted 30 in. inside the chamber, at the sacrifice of working

space, to make it possible for a man to get behind them for repair of leaks. This has proved very beneficial. The aluminum panels are painted with black epoxy paint on the inner surface for maximum absorptivity and are bright on the back side for maximum reflectivity to minimize heat absorption from the external vessel wall.

Cooling to $<100^{\circ}\text{K}$ is accomplished by means of LN_2 pumped under sufficient pressure to remain liquid at all times. It is cooled outside the chamber in a subcooler using evaporating LN_2 at atmospheric pressure for refrigeration. The panels are manifolded in a way which assures a uniform pressure drop in each circuit.

When the temperature is to be controlled between -65° and $+100^{\circ}\text{C}$, gaseous nitrogen is circulated through the system and either cooled with a 3 stage Freon 22 system or heated with steam, as conditions require.

LN_2 for the operation of the cold wall is stored adjacent to the building in a vacuum insulated ball holding 64,000 gallons.

These cold walls have been installed with very few leaks and the entire system commenced operation without excessive trouble.

Cryogenics

The cryopumping is accomplished with a 1 kilowatt capacity gaseous helium refrigeration unit circulating through helium panels. These are placed in the lower half of the chamber behind baffles projecting from the LN_2 cold wall at 45 degrees (frequently called Santeller array). This protects the helium panels from any heat load while utilizing their capacity to condense those that will not condense the LN_2 cold wall but will at 20°K . The helium panels are supplied with gaseous helium at 15°K to maintain a maximum temperature of 20°K .

The upper half of the LN_2 cold wall is provided with shields so that helium panels can be installed in this portion of the chamber at a later date.

The Solar Simulator

The problems in building the solar simulator were more severe than those in any other part of this project. This is because the state-of-the-art was extended further than in the other portions of the job and also because of production problems.

After much consideration of the state of the art of solar simulation, it was decided to set the following specifications:

1. Intensity—50 watts/ft.² to 275 watts/ft.²
2. Uniformity— ± 10 percent over any square foot.
3. Collimation— ± 4 degree half angle.
4. Spectral Distribution—Equivalent to zero air mass solar radiation from 0.3 to 4.0 microns.
5. Area Covered—20 ft. diameter circle.

After a review of all available information, an on-axis system proposed by Minneapolis-Honeywell consisting of multiple-modules, each self-contained, was selected. The design selected is based on using 127 modules placed in a hexagon of 20 in. centers (Figure 10). Together these

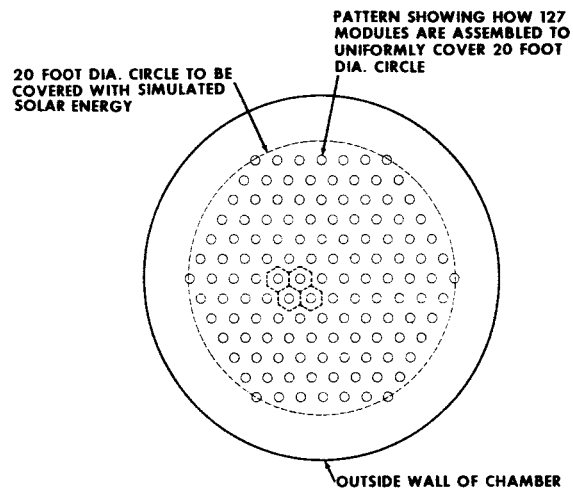


FIGURE 10.—Solar module arrangement.

will cover a circle 20 ft. in diameter. Because of the collimation angle there is some overlap between modules. Therefore, the intensity will drop somewhat at the edges, starting to fall off outside a 17 ft. diameter circle. This decision was based on these considerations:

Advantages

1. Lower initial cost.
2. Smaller optics.
3. Smaller penetrations.
4. Uses $2\frac{1}{2}$ kw lamps for earth orbit intensity.
5. Possible to use partial system for limited area.

6. Can be proven by building and testing one module.
7. Better efficiency.

Disadvantages

1. Burned out lamp causes dark area.
2. Many penetrations.
3. Does not preclude re-reflection (re-reflection will not exceed 2 to 3 percent with the multiple modules).

Each module is identical to all the others and consists of (Figure 11):

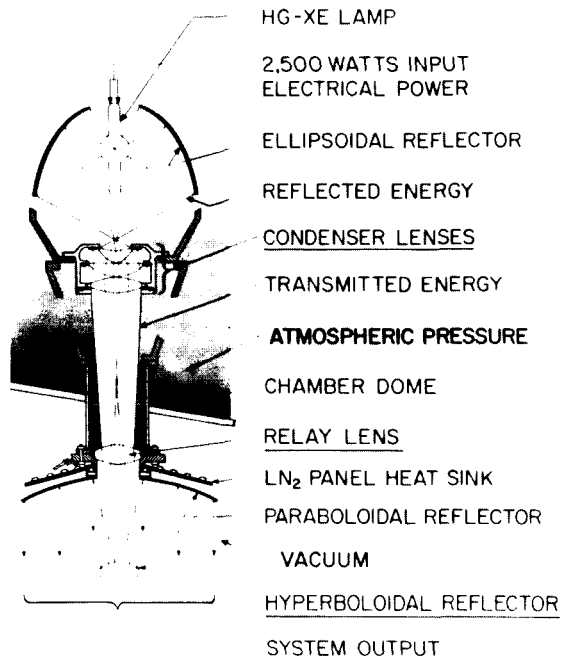


FIGURE 11.—Solar simulator module schematic.

1. The light source (presently planned as a 2500 watt mercury-xenon short arc lamp, with power supplies and holding fixtures designed so that later a 5000 watt lamp can be used).
2. Elliptical reflector to collect the energy.
3. Three condenser lenses.
4. Relay lens.
5. Hyperboloidal reflector.
6. Paraboloidal reflector.

An overall efficiency from lamp input to energy delivered on the test floor of approximately

12½ percent is required to make the system produce the desired intensity of 130 w/ft.² using the 2½ kw lamp. The requirement for this degree of energy transfer efficiency in the system is the basic cause of the extreme fabrication accuracies required, as discussed below. The basic design of the system has changed very little since the original conception, but there has been much detail refinement and many concurrent troubles.

It was necessary to settle on detailed design of the lenses early because of the urgency to complete the facility. Once this lens grinding was started, it limited design conditions on other elements of the system.

Mechanical Tolerances

While the lenses were being ground, it became apparent that a detailed tolerance analysis was essential for firm mechanical design and alignment requirements. Therefore a study was instituted and 150 hrs of IBM 7090 computer time were utilized to determine the effect on the system of movements of each of the elements relative to others and of incremental inaccuracies of each. As a result it was determined that the lamp and elliptical reflector must be located within ±0.010 inches with reference to each other and that the parabolic reflector with reference to the hyperbolic reflector must be held within 0.010 inches of the hyperbolic reflector. The other elements did not have critical tolerances and would be satisfactory if held within 0.050 inches. The surface of the elliptical reflector must have a slope accuracy of 5 minutes.

It also became apparent that a new process would have to be developed for making elliptical reflectors in a manner suitable for quantity (127) production. Therefore it was decided to halt all hardware production until a module could be built and tested using all final design optical elements. These would all be made by the proposed production techniques. This would also verify the computer study and give much more assurance of final success.

REFLECTORS AND LENSES

The Ellipse

The elliptical reflector was unusually deep. So far as could be determined a reflector of this

contour had never been produced with the required accuracy. Since 127 of them were required, it was decided that a process which reproduced reflectors from a very accurate master would have to be employed.

First, difficulty was encountered in obtaining a glass master made to acceptable accuracies. Even getting a glass blank of the right shape proved difficult. An attempt was made to get pyrex type glass slumped to the approximate contour and then annealed. This proved to be a very slow process and in order to save several weeks annealing time, we agreed to only a rough anneal before grinding the master. This later proved to be a mistake as several masters were lost by a very gentle thermal shock at temperatures no higher than 140°F.

Later several optical shops were contacted for a second try at grinding the master. None would commit themselves to less than three months time for production. This turned out to be optimistic since delivery took more than five months.

In the meantime, we investigated other methods, one of which was making the master in metal on a tape controlled machine. This process is fast and apparently accurate. However, since this is not done in an optical shop, there is the added problem that mechanical tolerances are not easily translatable into optical tolerances. A master produced by this process does meet (according to initial tests) the optical tolerance requirements. Recently we learned of an optical shop with a new method of grinding glass rapidly and accurately using a diamond grinding process and have procured a master by this process and delivery was less than two months with accuracy much better than those previously produced.

After getting a master it was necessary to replicate the reflector. The first attempt was an ellipse made of cast epoxy which failed to hold its shape. Electroforming has shown considerable promise of success. To date we have been most successful with a thin film epoxy replication. For this an aluminum casting is machined approximately 0.030" oversize and a thin film of epoxy is forced between the aluminum and the master. It is hoped that the thin epoxy film will hold its accurate contour under the severe heating and cooling of the operational cycle: however, tests of up to 200 hours reveal no deterioration.

The Hyperbola and Parabola

The hyperboloidal reflector being made from quartz has presented no manufacturing problems: however, there have been temperature difficulties. The first design was for rigid contact with the LN₂ cooling coil and this created too much shock for quartz which fractured on each test. A new design using flexible copper leads to carry the heat to the LN₂ was used. This does not cause fracturing: however, heat destroyed the reflective coating. Latest tests indicate that this can be overcome by applying the proper emissive coatings on the back side of the reflector to assist in removing the heat by radiation to the cannister. As insurance an electroform hyperbola is being made and will be tested for heat transfer and satisfactory optical properties throughout the temperature range.

Electroforming has been used satisfactorily to make the paraboloidal reflector, although there was some early experience with lack of accuracy at the edges.

The Lenses

The lenses have presented no major problem once the allowable tolerances were determined. They are now being produced by Goddard Space Flight Center.

ASSEMBLY OF PROTOTYPE MODEL

Alignment

Fixtures have been designed to check and align each of the elements of the module within the tolerances required. It is believed that these will give satisfactory results when carefully used. To facilitate lamp replacement, one fixture will align the lamp to a predetermined surface. Thus the assembly can be replaced so that the arc of the new lamp will be in the same location as the replaced lamp.

Mounting Hardware

When the manufactured prototype components were completed, a module was assembled and properly aligned for test. The first tests indicated the temperature rise in the mounting hardware was causing excessive growth and not allowing the optical elements to hold their position within allowable tolerances. This was corrected by

increasing the reflectivity of the hardware with aluminum paint; and thereby decreasing the heat being absorbed and in turn its thermal growth.

Internal Reflector Cooling

LN₂ heat sinks have been designed behind each of the paraboloidal, as well as the hyperboloidal, reflectors to remove the heat absorbed by these elements.

Vacuum Seals

The vacuum seals from the relay lens to its flange and from the flange to the vessel have caused difficulties. The former is of viton and the latter was to be crushed aluminum. Temperature readings on these surfaces indicated probable trouble because of the broad temperature range (from maximum heat with the continuously-running solar simulator to the cold experienced with the LN₂ cooling when the solar simulator is not operating). Latest tests indicate that relocation of the LN₂ cooling unit and the use of viton seals throughout will probably correct this.

Relay Lens Protection

The temperature of the relay lens becomes low enough while the parabola is being cooled with LN₂ (the solar simulator not operating) to cause condensation if atmospheric air is in the area. To eliminate this, gaseous nitrogen which has evaporated from the operation of the cold wall

is heated, pumped up to, and distributed to each of the relay lenses and thereby prevents any condensation on them.

In addition to the prototype module mentioned above a test fixture has been made in which each element of the system may be individually adjusted. With this fixture (containing all production hardware except for the ellipse) intensities of 160 w/ft.² have been achieved. We feel that this performance gives great confidence in the basic capabilities of the system. The problem is how to make 127 equivalent modules.

The prototype system has shown performance as high as 142 w/ft.² and when reassembled with changes in lamp, ellipse, parabola, and hyperbola, has consistently shown results of more than 130 w/ft.²

As a result of testing the module and its components, we are confident of reaching our goal and are proceeding with the manufacture of all components. At present, expectations—as compared with the original specifications—are given in Table I.

In explanation of "fair match", this is a subjective opinion based on the use of wide-band filters corresponding to the manner in which the desired spectrum was specified, i.e. equivalent energy absorption for typical materials. These tests, carried out by Dr. Drummond of Eppley Laboratories, indicated that the ultraviolet had been filtered out by the system and that the infrared came through the system with a minimum

Table 1
Original and Present Specifications

Quantity	Specifications	
	Original	Present
Intensity	50 to 275 w/ft. ²	50-130 w/ft. ²
Uniformity	±10% over 1 ft. ²	±10% over 1 ft. ²
Collimation	±4° half angle	±4° half angle
Spectral Distribution	Johnson Curve	Fair Match
Area Covered	20 ft. diameter circle	17 ft. diameter uniform fall off to 50% at 20 ft.

loss. We chose HgXe lamps as a source because of the much lower maintenance requirements as compared with carbon arcs. Also, the HgXe lamps provided ultraviolet in excess of that found in the solar spectrum, which overcame the normal losses found in an optical system. This has been done with $2\frac{1}{2}$ kw HgXe lamps and the system is

designed to accept 5 kw HgXe lamps. As soon as the system is operational, we shall start work on the use of larger lamps which will be used to obtain Venus orbit intensity.

The construction of this simulator has been underway for two years and appears to be approximately one year from completion.

POST-IRRADIATION, ROOM TEMPERATURE, ELECTRICAL CHARACTERISTICS OF N-P SILICON SOLAR CELLS*

BRIAN T. CUNNINGHAM AND ERNEST MOSS

INTRODUCTION

An engineering investigation has been conducted to determine the post-irradiation, room temperature electrical characteristics of state-of-the-art, flight-quality, silicon N/P solar cells. The primary reason for this work was to obtain a complete and up-to-date evaluation of flight-quality solar cells being produced by the leading solar cell manufacturers in the United States. The present report contains the results of a 1 Mev electron bombardment to a flux level of 10^{16} electrons/cm² conducted at, or near, room temperature, combined with a cursory examination of post-irradiation annealing at elevated temperatures.

PROCEDURE

A "Request for Quotations" was forwarded in April 1964 to each of the leading solar cell manufacturers, requesting price information on production-type, flight-quality, 1×2 cm N/P solar cells in various quantities and efficiencies. The results of this request are shown in Table I. An order was placed with each of the respondents for 50 of the highest efficiency cells which they quoted in the RFQ. A tabulation of the purchase order is shown in Table I. It should be noted that no particular requirement was placed on base resistivity and hence the cells received varied from a nominal 1 to a nominal 10 ohm-cm material. Base resistivity ranges are given in Table IV. Nominal values are used throughout this presentation. In addition to the boron-doped cells received from all manufacturers, a group of 15 aluminum-doped 10 ohm-cm cells were forwarded by the Texas Instruments Company.

*Published as *Goddard Space Flight Center Document X-636-64-253*, August 1964.

TABLE I.—Summary of Cell Efficiency Order

Manufacturer	AMO Efficiency
Heliotek.....	11%
Hoffman.....	11%
International Rectifier.....	10.5%
Texas Instruments.....	10.5%
Radio Corporation of America.....	13%

Twenty-five cells from each manufacturer (with the exception of the aluminum-doped cells, where only fifteen were available) were selected at random for irradiation. E-I curves for all cells were obtained using a Spectrosun solar simulator. The simulator intensity was adjusted to an air mass zero equivalent with an aircraft-calibrated solar cell.† All measurements in this report were made at $32^{\circ}\text{C} \pm 2^{\circ}$ as measured on a thermocouple placed underneath the thermally conductive cell-mounting block. In the data to be presented, it will be noted that the results have been normalized to 30° centigrade, using results obtained previously.¹

After initial measurements were made, the cells were taken to the Naval Research Laboratory in Washington, D. C. for irradiation on the 1 Mev Van de Graaff generator. After each dose the cells were returned to the Goddard Space Flight Center for measurements.

Each group of 25 cells (again with the exception of the aluminum-doped units, where three were pulled after each irradiation) was allocated as shown in Table II.

†The calibrated solar cells were flown at different altitudes by the Lewis Research Center. A Langley plot was made of short circuit current vs. air mass and the extrapolated air mass zero short circuit current obtained.

TABLE II.—*Typical Allocation of Solar Cells from Each Manufacturer*

1 Mev Flux Electrons/cm ²	Number of Cells Irradiated	Number of Cells Withheld	Temperature Treatment
5×10 ¹² -----	25	4	4
5×10 ¹³ -----	21	4	4
5×10 ¹⁴ -----	17	4	4
5×10 ¹⁵ -----	13	4	4
10 ¹⁶ -----	9	—	4

It may be seen from this table that after each irradiation, four cells were removed from the group for the purpose of more extensive tests to be conducted later. Four cells from each manufacturer (noted in the column labeled Temperature Treatment) were measured after the irradiation, given a temperature cycle later modified to a high temperature soak—and then immediately

re-measured. These cells were then subjected to the next irradiation. The same group of cells was temperature-cycled throughout the experiment. No thermal tests were conducted on the aluminum-doped cells because of the limited number available. The post-irradiation measurement and thermal treatment schedule are shown in Table III.

TABLE III.—*Measurement and Heat Treatment Schedule*

1 Mev Flux (electrons/cm ²)	Nontreated Cells Post- Irradiation Measurement Commenced after (hours)	Treated Cells Post-Irradiation Measurement Commenced after (hours)	Temperature Treatment
5×10 ¹² -----	18	1	A
5×10 ¹³ -----	1	1	B
5×10 ¹⁴ -----	1	1	C
5×10 ¹⁵ -----	18	1	D
10 ¹⁶ -----	18	18	E

Explanation of Temperature Treatment:

- A—Room temperature to -70°C to 100°C in 1 hour, hold at 100°C for 0.5 hours.
- B—Same as A except temperature was held 100°C for 2 hours.
- C—Room temperature to 100°C in 0.25 hour, hold at 100°C for 2 hours.
- D—Same as C except cells were held for 4 hours at 100°C.
- E—Same as C except cells were held at 100°C for 3 hours.

After the temperature treatment, the cells were allowed to cool to room temperature before measurements commenced.

RESULTS

Average values have been used throughout the

data presentation because of the small amount of scatter and the convenience associated with data reduction.

Results of the initial (pre-irradiation) measurements are shown in Table IV.

TABLE IV.—Initial Comparison of Manufacturer's Cells

Manufacturer	Code Letter	Average Maximum Power (MW)	Average AMO Efficiency* (%)	†Base Resistivity Range (ohm-cm)		
				Low	Nominal	High
Heliotek.....	A	24.9	9.9%	7	10	14
Hoffman.....	B	25.3	10.0%	8	10	12
RCA.....	C	26.7	10.6%	0.7	1	1.2
TI (Boron).....	D	25.7	10.2%	7	10	13
IRC.....	E	25.2	10.0%	?	2	?
TI (Aluminum).....	F	24.5	9.7%	7	10	13

*Based on 1.8 cm² area
 †According to manufacturer

Figure 1 depicts the change in short circuit current as a function of irradiation. It may be seen that in drawing the curves the data points were followed quite closely so as not to obscure

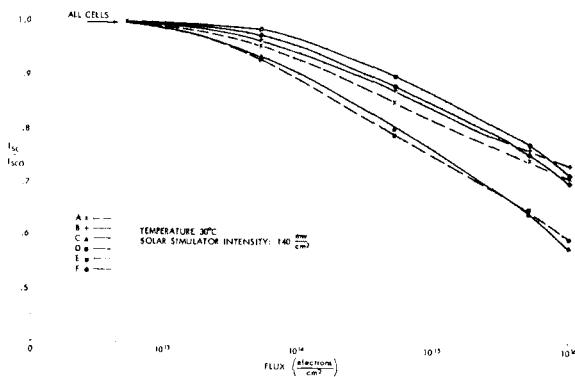


FIGURE 1.—Average short circuit current degradation as a function of 1 Mev electron flux.

any trends in the end points. Note the clear distinction between the lower (1 and 2 ohm-cm) and the higher (10 ohm-cm) base resistivity cells. Figure 2 shows the change in maximum power as a function of irradiation for all cells under investigation. Note that the TI cells show significantly less degradation than the other 10 ohm-cm cells, which in turn show significantly less degradation than the lower base resistivity cells. Figure 3 depicts the change in open circuit voltage as a

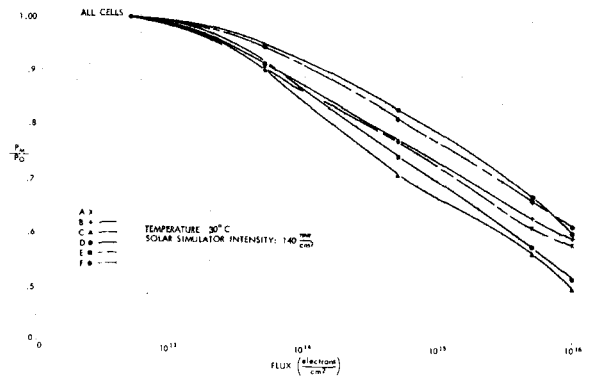


FIGURE 2.—Average peak power degradation as a function of 1 Mev electron flux.

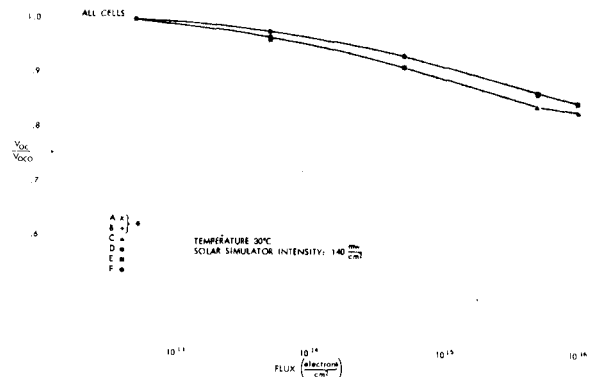


FIGURE 3.—Average open circuit voltage degradation as a function of 1 Mev electron flux.

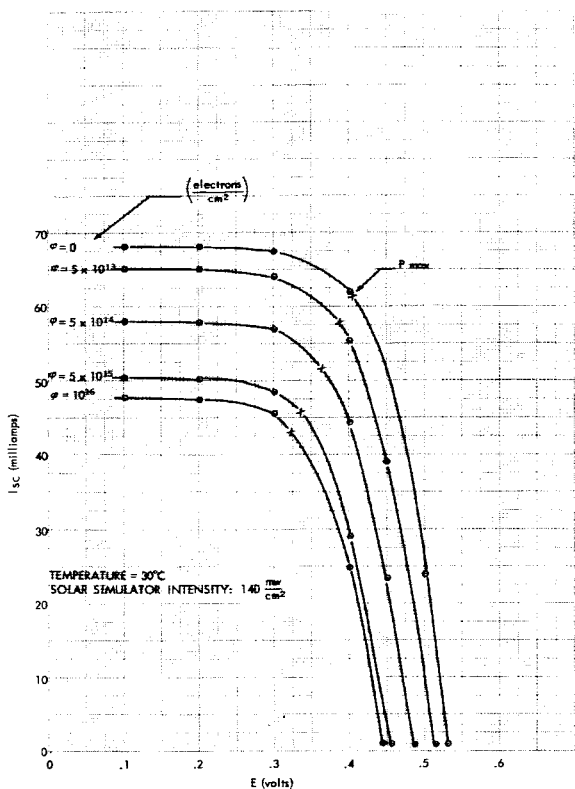


FIGURE 4.—Average degradation of group A solar cell I-V curves with 1 Mev electron bombardment.

function of irradiation. Figures 4 through 9 show the average E-I curve at different 1 Mev flux levels for each manufacturer. Figures 10 through 14 show the effect of thermal cycling. The smooth dashed curve in these figures is the same curve as appears in Figure 1 for the particular manufacturer. The solid sawtoothed degradation curve shows the effect of thermal cycling. The lower points at each flux level represent the average short circuit current readings of the cells prior to the thermal treatment described in the procedure. The upper value represents the post-thermal treatment average of the same cells.

CONCLUSIONS

1. Solar cell manufacturers are presently producing cells with a nominal 10 percent air mass zero efficiency.
2. The nominal 10 ohm-cm solar cell is decidedly more radiation resistant than the 1 ohm-cm cell at higher 1 Mev electron flux levels. This is

apparent in both the short circuit current and the maximum power.

3. Both the boron and the aluminum-doped TI cells show less radiation degradation, particularly in the maximum power, than any of the other manufacturers' cells under investigation.

4. Differences in radiation degradation observed between the aluminum and the boron-doped TI cells were small and not considered significant.

5. Changes in short circuit current due to a post-irradiation temperature treatment (sometimes called annealing) were definitely experienced in a limited experiment and can be as much as 5 percent. Measurements were not extensive enough to determine quantitatively how much of an effect the elevated temperature had on I_{sc} change; however, as noted in Table III, after the 10^{16} electrons/cm² dose, the cells were allowed to sit at room temperature overnight before any measurements were made. The data in Figures 9 through 14 at the 10^{16} e/cm² flux level seem to

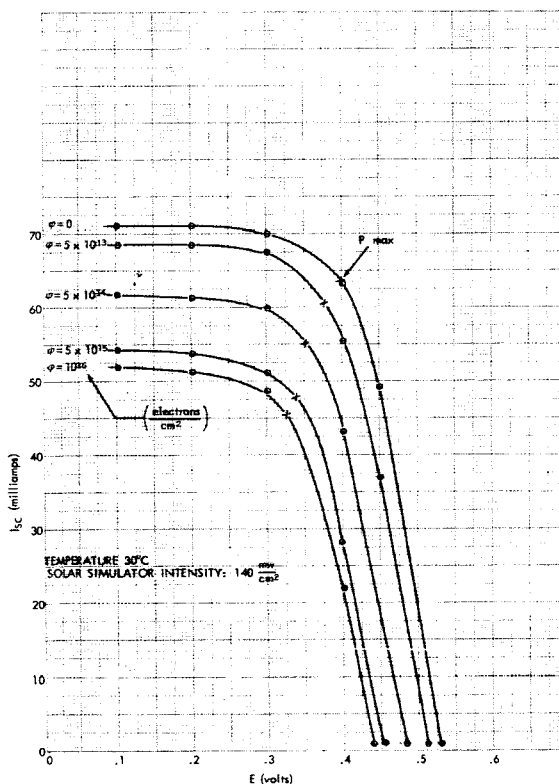


FIGURE 5.—Average degradation of group B solar cell I-V curves with 1 Mev electron bombardment.

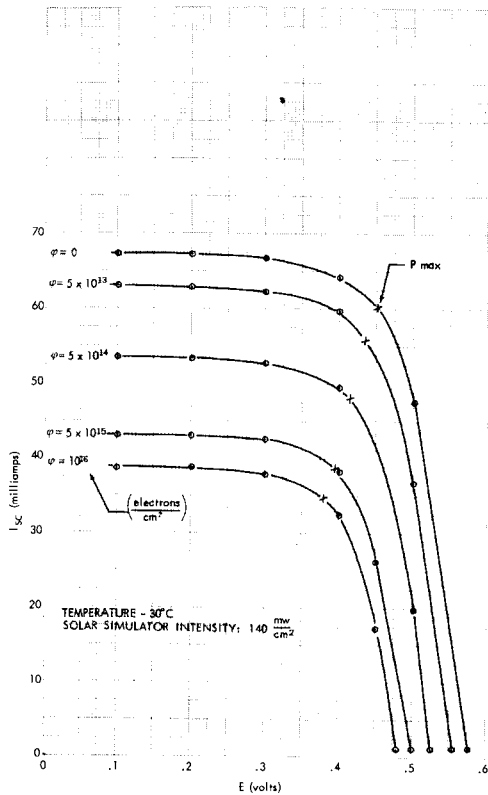


FIGURE 6.—Average degradation of group C solar cell I-V curves with 1 Mev electron bombardment.

indicate that an appreciable amount of improvement in I_{sc} (relative to the other post-irradiation, pre-heat treatment measurements) had taken place at room temperature in the intervening eighteen hours. The following general observations can be made relative to the change in I_{sc} :

- A. Subsequent irradiations degrade the cell more rapidly so that the end result essentially is the same as if no I_{sc} improvement had taken place.
- B. Cells from different manufacturers show different I_{sc} improvement characteristics.
 - (1) Initial temperature cycling seems to have degraded the TI cells so that subsequent post-irradiation treatment did not bring the post-thermal treatment value above the average of the nontreated cells.
 - (2) The low base resistivity (RCA and

IRC) treated cells are consistently above the average degradation line.

- (3) The 10 ohm-cm Heliotek and Hoffman cells oscillate about the average line as a result of the temperature treatment.

C. In order to compare the results of radiation damage studies performed in the laboratory, it will be necessary to find a means of monitoring or quenching the I_{sc} change so as to normalize the results to a given time after irradiation.

ACKNOWLEDGMENTS

Mr. William R. Cherry and Mr. Luther W. Slifer, Jr. of GSFC for assistance in the planning of the experiments.

Mr. Carmen Carosella and Mr. Richard Statler of NRL for conducting the irradiation.

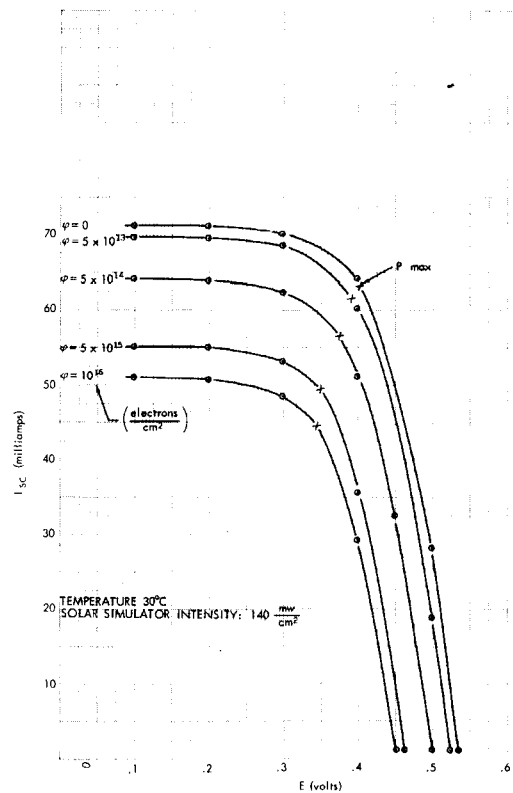


FIGURE 7.—Average degradation of group D solar cell I-V curves with 1 Mev electron bombardment.

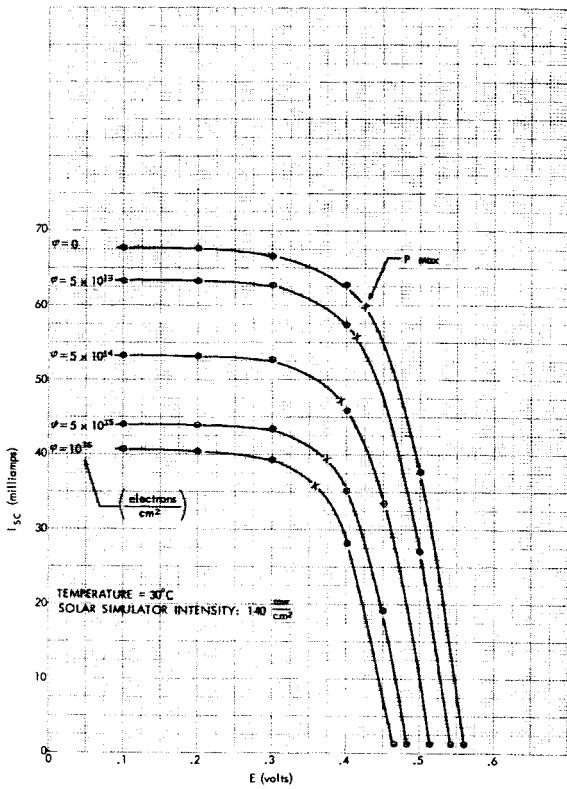


FIGURE 8.—Average degradation of group E solar cell I-V curves with 1 Mev electron bombardment.

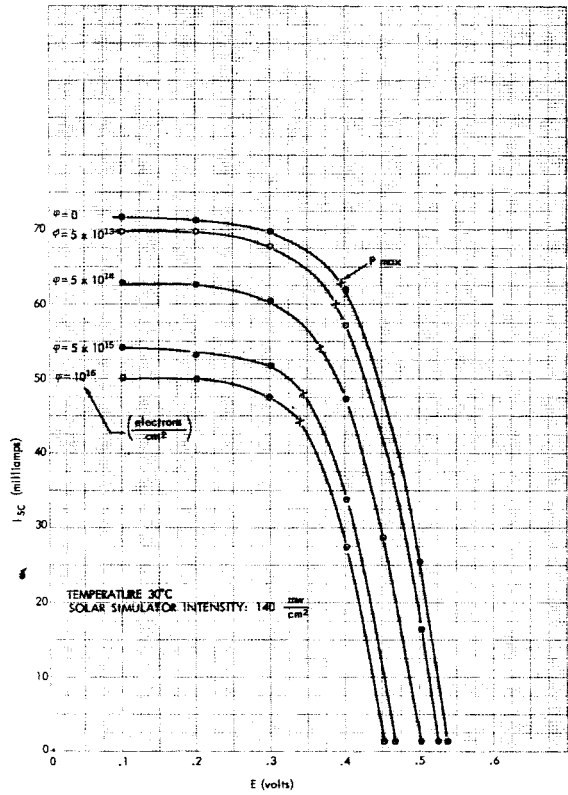


FIGURE 9.—Average degradation of group F solar cell I-V curves with 1 Mev electron bombardment.

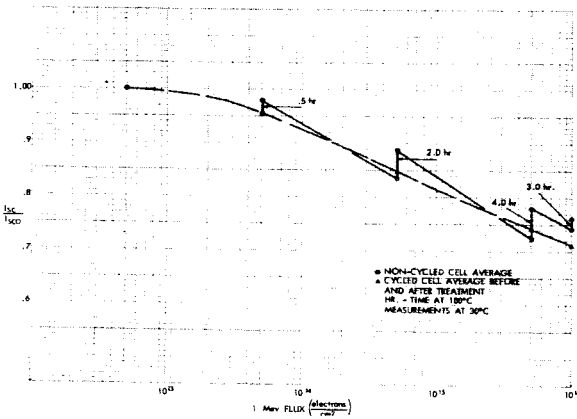


FIGURE 10.—Effects of temperature treatment on group A cells.

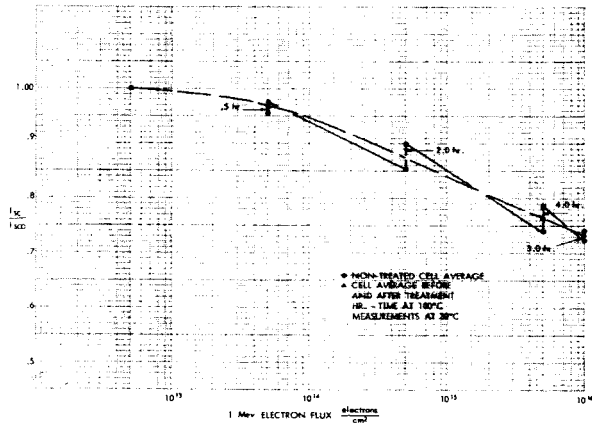


FIGURE 11.—Effects of temperature treatment on group B cells.

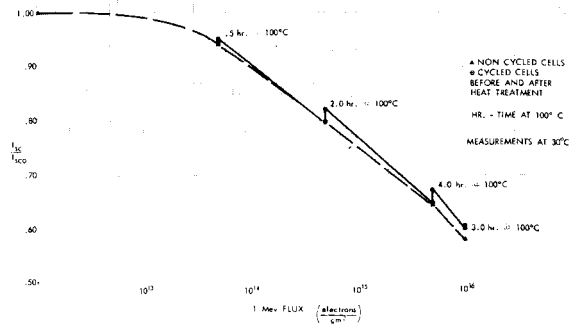


FIGURE 12.—Effects of temperature treatment on C-cells.

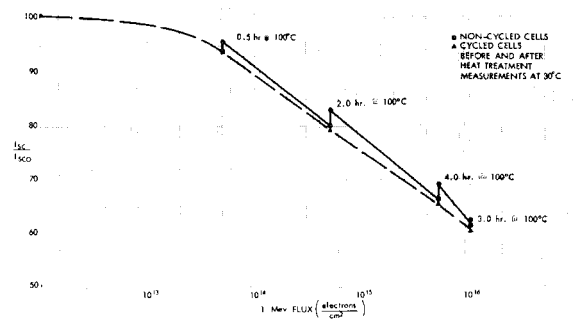


FIGURE 14.—Effects of temperature treatment on E-cells.

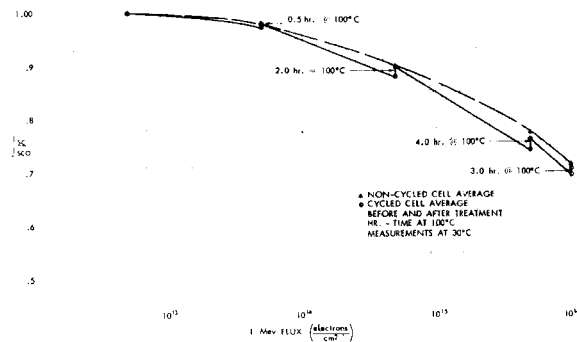


FIGURE 13.—Effects of temperature treatment on D-cells.

REFERENCE

1. BRIAN T. CUNNINGHAM, ROBERT L. SHARP, and LUTHER W. SLIFER, JR., "The Electrical Characteristics of Irradiated Silicon Solar Cells as a Function of Temperature," NASA-Goddard Space Flight Center, Greenbelt, Md., March 30, 1964.

AN ANALYSIS OF ERRORS IN THE GEOGRAPHIC REFERENCING OF NIMBUS CLOUD PICTURES*

EUGENE M. DARLING, JR.

Errors in the determination of satellite attitude from telemetered data are expected to cause a maximum overall error of 50 naut. mi. in the geographical referencing of Nimbus A cloud pictures. This error is tolerable for most operational purposes. For future Nimbus vehicles it is anticipated that this error can be reduced to less than 20 naut. mi.

INTRODUCTION

Both operational and research applications of satellite cloud data require that geographic referencing be superimposed on the pictures. This referencing—or gridding as it is usually called—may take the form of either latitude and longitude lines or gridpoints defined by latitude-longitude intersections.

If the orientation of the satellite relative to the earth were known exactly then it would be possible to compute perfect grids by transforming focal plane coordinates into spherical coordinates on the earth, or vice versa. However, in reality the satellite attitude is never known exactly, hence the coordinate transformations operate on erroneous input data and therefore produce incorrect grids.

This paper analyzes the effect of attitude stabilization errors on the accuracy of geographic grids computed for Nimbus cloud pictures. Two introductory sections contain technical material pertinent to the problem. The first section provides the necessary background information on the operation of the control system; in the second section the sources of mechanical, optical, and electrical error in the system are discussed and the maximum magnitude of each contributing factor is estimated in terms of its effect on satellite attitude. Next the computations of attitude errors resulting from various techniques for manipulating telemetry data are presented. Up to this point, the discussion will have been concerned entirely with the problem of satellite orientation.

However, in the following section the effect of attitude errors on geographic referencing is explored. Estimates of gridding errors for both Nimbus A and for future Nimbus satellites are presented.

In the final section the question of satellite data referencing is discussed in the broad context of location accuracy for related conventional meteorological data. Errors in the computation of mean cloud properties for grid squares are analyzed. An objective criterion is developed for determining the grid size compatible with a prescribed confidence level and a known gridding error.

DESCRIPTION OF THE NIMBUS ATTITUDE CONTROL SYSTEM

The portion of the Nimbus system which controls motions about the pitch and roll axes consists of three elements: *horizon scanners* which detect the earth and provide the required input to the *horizon attitude computer* which calculates pitch and roll errors[†] for input to the *control system*, consisting of flywheels and pneumatics, which generates the torques on the spacecraft

[†]Any angular departure of a satellite body axis from its nominal orientation produces an attitude error. The *nominal axis* positions are:

Yaw: in the orbital plane perpendicular to the orbit. Positive towards the earth

Roll: in the orbital plane perpendicular to the yaw axis. Positive in the direction of motion of the satellite.

Pitch: perpendicular to the orbital plane. Positive to the right of the satellite direction of motion.

The components of attitude error are defined as follows.

Yaw: rotation about the yaw axis so that the roll axis no longer lies in the orbital plane.

Roll: rotation about the roll axis so that the yaw axis no longer lies in the orbital plane.

Pitch: rotation about the pitch axis so that the yaw axis is no longer perpendicular to the orbit.

*Published as NASA Technical Note D-2137, January 1964.

required to correct these errors. The yaw attitude is detected by an integrating rate gyro which generates an error signal input to the yaw flywheel in the control system. See References 1 and 2 for a detailed discussion of the control system.

Pitch and Roll Attitude Control Scanner and Electronics

The detection device consists of two infrared scanners (IR1 and IR2) mounted on the control package (Figure 1). One scanner looks forward

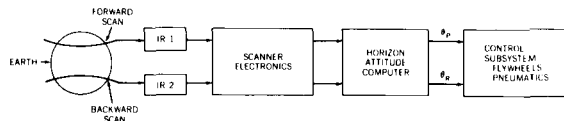


FIGURE 1.—Pitch and roll attitude control, where θ_p is the pitch error signal and θ_r the roll error signal.

in the direction of the positive roll axis; the other looks backward. The scanner beam is directed in a conical path by a rotating prism in such a way that the beam intercepts both the earth and outer space in each sweep. The detected energy is focused on a bolometer through an optical system sensitive to the 12.5 to 18μ band of the infrared spectrum. The bolometer output, being proportional to the incident energy, is essentially zero when the scanner sees outer space; increases rapidly at the leading edge of the earth; remains more or less constant while the beam sweeps across the earth; then drops sharply back to zero at the trailing edge of the earth.

Electronic circuits in the scanner system filter the bolometer output to produce a smooth wave form for input to a pulse reconstructor in which a square pulse is constructed from the original wave. Whenever the reconstructor detects an energy of 480 mv at the leading edge of a wave it marks the beginning of a pulse. The end of the pulse is marked when the trailing edge of the wave reaches zero output. The square pulse fed to the horizon attitude computer begins and ends at the points determined by the above criteria.

Horizon Attitude Computer

The horizon attitude computer operates on the square pulses received from the scanner system. Pitch error is obtained by properly weighting the

difference in width between the front and rear scanner pulses. Roll error is computed by weighing the difference in pulse length to the left and right of a magnetic reference which indicates the center of the pulse. Either the front or the rear pulse is used in this computation, depending upon the magnitude of the roll error. Pitch and roll errors are computed alternately and applied to a digital-to-analog converter which furnishes the proper error signal to the flywheels and pneumatics.

Control Subsystem

The pitch and roll attitude control loops both operate in the same manner. The error signal is fed to a flywheel amplifier which provides a control voltage to the momentum generator motor, thereby accelerating the flywheels by the amount required to produce the necessary restoring torque. If a specified threshold attitude error occurs then a signal is fed to a solenoid-controlled valve which permits gas to escape until the error falls below the threshold value. Also, in the event that either the pitch or roll flywheels reach 85 percent of maximum rated speed, pneumatic gating will occur (i.e. firing of a 0.52 second pulse of gas to dump flywheel momentum while at the same time providing the proper restoring torque).

Yaw Attitude Control

An integrating rate gyro (Figure 2) is mounted in the roll-yaw plane with its input axis inclined 12° with respect to the roll axis. This gyro senses yaw error as a component of the orbital pitch rate in the roll-yaw plane. The gyro electronics produces an error signal which is a linear function of yaw error, yaw rate, roll error and roll rate plus other small acceleration terms. The gyro output signal drives the yaw momentum motor which accelerates the flywheel by the proper amount. Pneumatic gating occurs at 65 percent of the maximum rated flywheel speed.

A yaw bias can be inserted by ground command in increments of 6° up to a maximum value of 30° . The purpose of this feature is to achieve

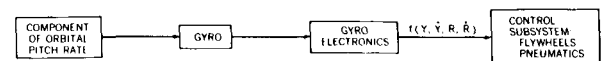


FIGURE 2.—Yaw attitude control, where Y is the yaw error; \dot{Y} the yaw rate; R the roll error; and \dot{R} the roll rate.

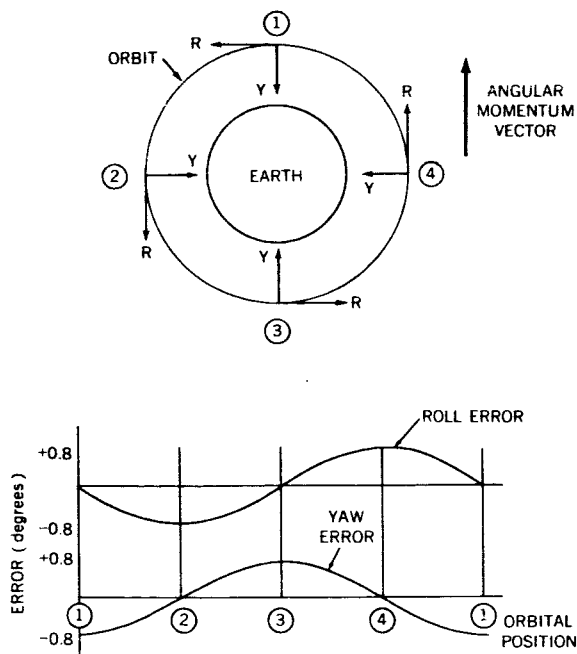


FIGURE 3.—Cyclic momentum exchange between the roll *R* and yaw *Y* axis.

maximum paddle insolation in the event that the satellite orbit departs from the sun-earth plane.

Cyclic Momentum Exchange

In the absence of pneumatic gating, meteorite impact, gravitational gradients or other external torques the resultant momentum vector of the satellite and its three flywheels remains fixed in inertial space. This condition is maintained by a cyclic interchange of momentum between the roll and yaw axes (Figure 3). The flywheel loops for roll and yaw are designed to take advantage of this momentum exchange in such a way that the removal of roll momentum by yaw substantially reduces the occurrence of roll gating. This design effectively minimizes the cross coupling between roll and yaw.

Figure 3 shows the normal cyclic variation in roll and yaw (in the absence of gating) resulting from flywheel operation. The error curves for roll and yaw each have an amplitude of about 0.8°; the two curves are out of phase by 90°.

Constant Pitch Error

At the end of the initial stabilization phase it is likely that the spacecraft will retain a certain

residual momentum which will be transferred to the flywheels. This initial flywheel momentum will be incorporated in the cyclic exchange between roll and yaw. However, the residual component along the pitch axis will result in a small, constant pitch error since the pitch momentum is not involved in the cyclic exchange mechanism. The maximum attitude error due to this effect is estimated to be about one half a degree.

SOURCES OF ERROR IN THE ATTITUDE CONTROL SYSTEM

Each component of the control system contributes to the total error in satellite attitude. The pertinent error-producing factors may be conveniently grouped into five categories. The maximum attitude error contributed by each category is listed in Table 1a.

Table 1
Sources of Attitude Error.

Error Source	a. Maximum Attitude Error (degrees)			
	Pitch	Roll	Yaw	
1. Input Signal	±0.58	±1.75	—	
2. Control System	±0.35	±0.35	±0.35	
3. Momentum	±0.45	±0.75	±0.82	
4. White Noise	±0.1	±0.1	±0.3	
5. Telemetry Scale	±0.06	±0.06	±0.06	
Mode	Error Computation	b. Total Attitude Error for Various Modes of Operation (degrees)		
		Pitch	Roll	Yaw
Disregard Telemetry	1+2+3+4	±1.5	±3.0	±2.0*
Use Telemetry	1+2+4+5	±1.1	±2.3	±2.0*
Scanner Pulse Reconstruction				
Assume 2° Yaw Error	2+4+5	±0.5	±0.5	±2.0*
Measure Yaw Error	2+4+5	±0.5	±0.5	±0.7

*It should be noted that the maximum yaw error for the first three methods has been entered as ±2.0° rather than the appropriate sum of component errors (see footnote page 6).

Input Signal

The effect of clouds on the earth's horizon is to place a cold radiating surface in the path of the scanner beam. As a result the energy level sensed

by the bolometer is intermediate between that of space and a cloud-free earth. Based upon theoretical calculations which have been verified by data from the Discoverer satellite the best estimate of the maximum energy ratio between a cloud-free and cloudy horizon is 2 to 1 (Reference 3). In calculating the pitch and roll errors shown in Table 1a this 2 to 1 ratio has been applied in the manner which produces the maximum error. Since yaw is under gyro control the above type of input signal error is not present.

Control System

This category includes the total error introduced by the scanner system (both optical and electronic) and the horizon attitude computer. Component errors averaging $.05^\circ$ have been root-mean-squared to obtain the tabulated values for each axis. In the case of both pitch and roll, the largest error (0.25° in each axis) is caused by mechanical alignment of the scanner system. The greatest individual contribution to yaw error (0.3°) is stability of the gyro drift.

Momentum, White Noise and Telemetry Scale

The errors tabulated for roll and yaw are a consequence of the momentum exchange mechanism previously described. It should be noted that the roll and yaw errors are out of phase by 90° so that the maximum error in one is always accompanied by zero error in the other (Reference 4).

The error listed under White Noise is that which would result from the presence of 100 mv RMS white noise in each control loop.

Attitude errors are telemetered in an 8-bit code (i.e. scale of 0 to 127). Thus there is a limit to the accuracy with which such errors can be read on the ground. This error is listed under Telemetry Scale.

ATTITUDE ERRORS FOR VARIOUS PROCEDURES

We shall now consider the errors which would result from several different procedures for determining satellite attitude. The total error for each method is computed by summing the values in Table 1a in a prescribed manner (Reference 5). Results are presented in Table 1b. It should be emphasized that errors computed in this way are maximum values. *Errors of this magnitude*

would occur only in the unlikely event that all contributing factors reached their maximum errors simultaneously and all with the same algebraic sign.

Disregard Telemetered Attitude Data

This is the null case where the satellite is assumed to remain in nominal attitude at all times. Accordingly, attitude errors are assumed to be zero throughout and all telemetry data on attitude is disregarded. The maximum error for this case is the sum of the first four items in Table 1a.*

Use Telemetered Attitude Data

This technique entails the use of telemetered attitude data without modification. The fine pitch and fine roll signals are interpreted as exact values of pitch and roll attitude, respectively. With this method the errors due to momentum are eliminated since the telemetered values are measured with respect to the cyclic momentum variations. However, the errors due to telemetry scale must be considered. The total maximum error in this case is the sum of times 1, 2, 4 and 5 in Table 1a. Note that the errors in pitch and roll are somewhat less than in the previous method.

Scanner Pulse Reconstruction

The salient feature of Table 1a is the large contribution of input signal error (Reference 6) to the total error in pitch and roll. For both axes this source accounts for *more than half* the total error. Thus, it is obvious that any expedient for substantially reducing this input error would result in a significant decrease in pitch and roll attitude error. A technique for reconstructing the earth input pulse from telemetered data on scanner output has been considered. The mathematics of this technique is presented in Appendix A. In applying the method (Reference 7), we must first reproduce the actual (i.e. cloudy) input wave. An ideal (i.e. no clouds) input wave of the same wavelength as the actual wave is then constructed and fed through an analog simulation

*Data on yaw attitude is provided with variable accuracy over a portion of the daylight orbit by the course sun sensor. The greatest accuracy (about $\pm 2^\circ$) is achieved at satellite sunset. From design considerations the initial yaw error is expected to be about $\pm 1^\circ$ with a possible maximum increase of 1° at the end of six months due to gyro drift. The ± 2.0 figure appearing in the table was selected as a pessimistic estimate of maximum yaw error, taking into account the degradation in yaw control expected at the end of the satellite life.

of the scanner system. Both the ideal and actual scanner outputs are used as inputs to a program simulating the horizon attitude computer to obtain:

1. Attitude errors relative to the actual earth; and
2. Attitude errors relative to the ideal earth.

The satellite control system responds to the former when it should be adjusting to the latter. Any difference between the attitude errors computed for (1) and (2) is due to the anomalous effect of clouds. With these two sets of attitude data it is a simple matter to determine the degree to which the actual satellite attitude departs from the attitude indicated by telemetry.

Assume 2° Yaw Error

In this case the error due to input signal has been removed from pitch and roll by the aforementioned technique. The $\pm 2.0^\circ$ error in yaw is retained. The total error in pitch and roll is now the sum of items 2, 4 and 5 in Table 1a.

Measure Yaw Error

This is the limiting case where pitch and roll errors are reduced as in the previous paragraph; and, in addition, it is assumed that yaw error is measured with the same order of accuracy as the other two axial components. The maximum error for all three axes is now the sum of items 2, 4 and 5 in Table 1a. This degree of accuracy could be achieved for the daylight orbit with a fine sun sensor of optimum design.

THE EFFECT OF ATTITUDE ERRORS ON THE ACCURACY OF GEOGRAPHIC REFERENCING

To be useful for either operational or research purposes the picture data from meteorological satellites must be provided with geographic referencing in the form of latitude-longitude grids. It is clear that uncorrected errors in satellite attitude will be reflected in location errors for picture elements relative to the surface of the earth. By means of spherical trigonometry it is possible to map angles measured at the satellite (i.e. attitude errors) into corresponding arcs on the surface of the earth representing the displacement of picture elements. Equations for this transformation are

derived in Appendix B. These equations have been used to calculate gridding errors corresponding to the attitude errors associated with each of the techniques listed in Table 1b. Results are summarized in Table 2. Complete tabulations and graphs are presented in Appendix B.

In Table 2 the heading "angle from PP" (principal point), referring to gridding errors, denotes an angle subtended at the satellite by an earth arc extending from the center picture principle point along the major axis (i.e., a perpendicular to the subsatellite point track bisecting the edges of the two side pictures) of the three picture area. The field of view of each camera is 37° with a 2° overlap, giving a total field of view of 53.5° on each side of the center camera principal point. In Table 2 gridding errors are shown for the center picture principal point, center picture edge, side picture center and side picture edge.

The mean maximum gridding error (see Appendix B, page 17) over the three picture area appears on the bottom line.

Note that gridding errors with the present design, disregarding telemetry, are about three times as large as those associated with the original design. When telemetry is used with the present design the errors are reduced to twice the original values. With the scanner pulse reconstruction technique the original design errors would be reduced by about 20 percent if a 2° yaw were experienced and by about 50 percent if yaw could be measured accurately.

It should be noted that the accuracy of the present control system would equal or exceed the original design specifications if the effect of input signal error could be removed.

Gridding Errors-Nimbus A

Initially, the picture grids to be generated by the computer complex at the CDA station in Alaska will be based upon attitude data obtained from telemetry which has been smoothed to remove large errors. Thus, the mean maximum error which may be anticipated at the outset of Nimbus operations will be about 50 naut. mi. (Table 2). As experience with telemetry data increases, we may find ways of reducing this error, but the degree of improvement cannot be estimated in advance. The manipulation of telemetry data to obtain refined attitude information

Table 2
Nimbus Attitude and Gridding Errors for an orbital altitude of 500 naut. mi.

Error Specifications		Original Design	Present Design		Scanner Pulse Reconstruction		
			Disregard Telemetry	Use Telemetry	Assume Yaw 2° Error	Measure Yaw Error	
Attitude Errors (degrees)	Pitch	±1	±1.5	±1.1	±0.5	±0.5	
	Roll	±1	±3	±2.3	±0.5	±0.5	
	Yaw	±1	±2	±2	±2	±0.7	
Gridding Errors (naut. mi.)	Angle from PP						
	0°	Location in Picture	9	31	21	5	5
	18.5°	Principal Point, Center Picture	11	38	27	8	7
	35°	Edge, Center Picture	20	56	45	18	13
	53.5°	Center, Side Picture	56	169	123	42	28
Mean Maximum Gridding Error (naut. mi.)			23	73	52	17	12

must be regarded as one of the experiments under the Nimbus R&D program.

Gridding Errors-Future Nimbus Satellites

Two approaches which show promise of providing more accurate attitude data in the future are the scanner pulse reconstruction method and the application of classical theoretical mechanics.

Scanner Pulse Reconstruction

Table 2 indicates that this technique offers a maximum possible reduction in gridding error of 60 percent or more compared with the "Use Telemetry" method of Nimbus A. The accuracy of this approach is contingent upon the development of a reconstruction technique with minimum sensitivity to noise in the scanner system transfer functions. So far this result has not been achieved.

If the development of this technique should prove successful, it would be necessary to make provisions on the satellite for routinely recording

and transmitting the scanner data. In the Nimbus A configuration high resolution infrared radiometer (HRIR) and scanner data are fed to the same tape recorder. However, the operation of this recorder is such that it is not feasible to time-share HRIR and scanner data on a single pass. As a consequence, it is necessary to record a full orbit of either HRIR or scanner data. Data selection is made by ground command. Due to the prime importance of HRIR data, the satellite tape recorder will routinely operate in the HRIR mode. An occasional orbit of scanner data will be recorded for the purpose of checking the operation of the control system. Only in the event of failure of the HRIR system would scanner data be recorded and transmitted routinely. Accordingly, the approach adopted for Nimbus A is to analyze, on a research basis, whatever scanner data becomes available with the aim of exploring the feasibility of developing an operational technique for extracting attitude from these data. In the event that such a technique were developed,

it would be necessary to modify the input equipment and input logic of the CDA computer system to accommodate the scanner data.

Theoretical Mechanics

The most powerful attack on the attitude problem entails the solution of equations governing the angular momentum vector of the satellite. As mentioned earlier, if there is no external torques acting on the vehicle, the resultant of flywheel momentum and satellite body momentum is a vector fixed in inertial space. In this case pitch, roll and yaw are defined by a third order system of differential equations which can be solved for the initial condition of zero error and zero rate in all three axes.* This set of equations represents a first approximation to the complete momentum equations, derived by omitting terms describing external torque contributions (1) at satellite separation, (2) due to pneumatic gating, and (3) caused by gravitational gradients. The solution to such a mathematical system is subject to error due to inaccurate specification of the satellite inertia tensor, uncertainties in flywheel response and pneumatic operation, as well as lack of precise knowledge about the initial motion of the vehicle.

Reduction of Yaw Error

Complete elimination of the yaw attitude error in the two present design cases (i.e. Appendix B, Table B1) would produce a *negligible* decrease in total attitude error. This is a consequence of the fact that the roll error accounts for at least 80 percent of the total error. Therefore it is not profitable to expend effort in reducing yaw error unless the input signal error affecting pitch and roll can be substantially reduced by some such technique as pulse reconstruction. (From Table B1 it is evident that yaw error becomes the prime contributor to total error once the input signal error has been eliminated).

There are two possible approaches to the problem of reducing yaw error, should this prove desirable: (1) incorporation of an improved sun sensor in the satellite control system, or (2) the development of techniques for solving the gyro equation. A sun sensor yields useful data for the

daylight orbit only, hence this approach does not offer a complete solution to the problem. On the other hand the gyro equation applies to the entire orbit so that a successful technique for solving it would represent a complete solution.

Yaw Bias

The effect of yaw bias had not been taken into account in this study. The introduction of a bias would result in another source of error in our knowledge of yaw orientation, over and above the uncertainty due to inaccuracies in the course sun sensor. No attempt has been made to estimate the magnitude of errors due to this effect.

METEOROLOGICAL ASPECTS OF GRIDDING ERRORS

The degree to which errors in geographic referencing affect the usefulness of meteorological satellite data is entirely a function of the requirements which the data must fulfill. For example the field forecaster who is interested in general cloud conditions over a sizable area has far less need for exact referencing than does the researcher in tropical meteorology who is concerned with the properties and motions of cloud elements of a size near the resolution capabilities of the AVCS system.

Winds and Clouds

Relation to Winds

Satellite cloud pictures will be used in conjunction with conventional meteorological data to analyze the existing state of the atmosphere as a preparatory step to prediction. The pictures will be related to certain aspects of the wind field which provide the key to future behavior of clouds. For example, a simple prediction technique might consist of advecting the clouds with the winds at cloud level. A more sophisticated method might make use of the vorticity field as an indicator of circulation development in conjunction with the divergence field as a predictor of precipitation.

It should be noted that the location error of upper level winds relative to the observing site is about 30-50 naut. mi. due to the trajectory of balloonborne radiosondes. This error is of the same magnitude as the gridding error and obtains when the cloud picture and the wind observation are taken at the same time. Since picture taking

*These equations have been derived by Dr. Poritsky of the GE Advanced Technology Center (Reference 8).

times do not in general coincide with wind observing times it will be necessary to displace the cloud data in order to obtain coincidence. On the average this adjustment will entail the displacement of clouds by an amount corresponding to an 8 hour cloud movement. Thus the average displacement will be about 250 naut. mi. By assuming that 50 percent of this displacement could be accounted for by extrapolation or prediction, the residual uncertainty in cloud location would be about 125 naut. mi.

Cloud Life Cycle

Another important consideration in the gridding problem is the life cycle of clouds identifiable in satellite pictures. It is convenient to classify clouds in three categories, each differing from the next by approximately an order of magnitude in cloud lifetime (Table 3).

Table 3
Cloud Life Cycle.

Type of Cloud	Order of Magnitude of Cloud Duration (hrs)
Individual cell	1
Cloud mass (e. g. cumulonimbus), squall line	10
Cyclonic and frontal cloudiness	100

In the prediction process we are mainly concerned with cloud features which appear in pictures taken on successive passes over a particular region. Since individual cloud cells will pass through their life cycle during the interval between observation and data distribution, this category is not of interest for prediction. Cloud masses tend to have a marked diurnal variation, hence they will be only marginally identifiable at 12 hour intervals—once by AVCS, and once by HRIR—by a single Nimbus. Cloudiness in this category will be of limited use for prediction.

Cyclonic and frontal cloudiness are observable by satellites for periods of several days and hence are of immense value in the prediction process. It is this type of cloudiness which will be associated with the wind field in the preparation of cloud

and precipitation forecasts. Cloud systems of this type cover areas of several tens of thousands of square miles, hence, there is obviously no need to locate individual elements of such systems with a high degree of accuracy.

Errors in Mean Cloud Properties

The life cycle of cloud elements, as well as the practical aspects of operational use, dictate that the field meteorologist concern himself with mean cloudiness over an area. We shall investigate the effect of gridding errors on his estimate of mean cloud cover. For this purpose it is convenient to consider a circle of uncertainty centered at each observed* cloud element location. The magnitude of the radius R is equal to the mean maximum gridding error (Table 2). The actual location of each observed cloud is assumed to be within this circle of uncertainty.

Suppose that we wish to estimate the mean cloudiness for a grid square of side S . It is obvious that the fraction of cloud elements observed within the square S which are actually there, is dependent upon the ratio S/R . The larger this ratio the higher the fraction of actual elements within S and vice versa. The higher the fraction the more reliable are the computations of mean properties within the square and vice versa. Thus, we may substitute errors in the computation of mean values at known locations for errors in the location of known elements. However, in the process we have registered a gain in usability since mean cloudiness expressed in 3 or 4 categories is sufficient for many purposes so that a 20 or 30 percent error in mean cloud amount is not significant.

The mean probability that elements observed within a square S actually belong there when the radius of uncertainty is R was computed by determining the fraction of the circle R which falls within square S . This fraction was integrated along the diagonal of the square and divided by the total distance to find the desired probability. The non-dimensional parameter S/R is plotted versus probability in Figure 4.

Note that the grid square side must be at least 50 percent larger than the radius of uncertainty

*The word "observed," in this context means the location computed by a gridding method which contains errors due to uncorrected satellite attitude deviations from nominal.

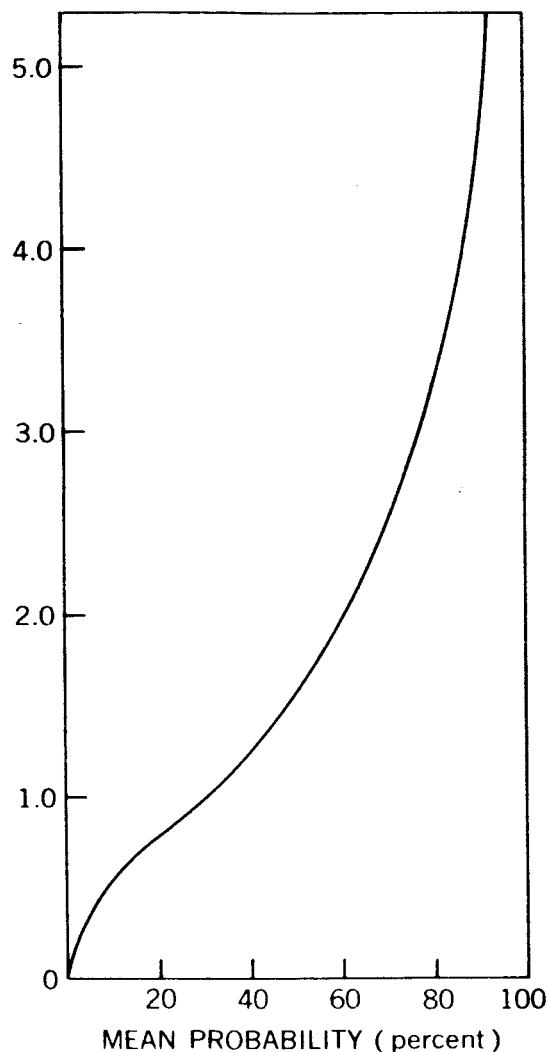


FIGURE 4.—Mean probability of actual cloud locations in a grid square, where S is a side of a grid square, and R the radius of maximum uncertainty in cloud element location.

in order to achieve a 50 percent probability that the observed cloud elements actually belong in the square. For 75 percent probability the side must be 3 times the radius.

Applying the above criterion to the gridding errors anticipated for Nimbus A (Table 2) we find that the mean maximum error of 50 naut. mi. is compatible with 75-naut. mi. grid squares for 50 percent probability or 150 naut. mi. squares for 75 percent. In the future when the gridding error has been reduced to 20 naut. mi. the associated grid squares will shrink to 30 naut. mi. for 50 percent or 60 naut. mi. for 75 percent.

Accuracy Requirements for Cloud Location

For most operational applications of satellite picture data a maximum gridding error of about 50 naut. mi. (Reference 9) would be adequate if the confidence level were at least 70 percent that observed elements in a square actually belonged there. This criterion (Figure 4) sets the grid square side at 125 naut. mi. (about 2° of latitude). With a grid of this size the adjustment for coincidence of cloud and wind data would require an average displacement of only one grid interval.

To summarize, a 125-naut. mi. grid imposed on Nimbus A picture data will satisfy the majority of operational users.

The Weather Bureau is planning to automatically process digitized Nimbus video data to produce a rectified product on a 10-naut. mi. grid (Reference 10). If the 50 naut. mi. mean maximum gridding error were ever to occur then each of these squares would contain only 2 percent of the data which actually belongs there. If the Weather Bureau photogrammetric system (Reference 11) for attitude determination were developed the mean maximum gridding error could be reduced to about 20 naut. mi. in which case the 10-naut. mi. squares would contain 8 percent of the proper data. In order to increase the location probability to 50 percent, the square size would have to be increased to 75 naut. mi. with the 50 naut. mi. error and to 30 naut. mi. with a 20 naut. mi. error.

The above figures only apply to the case where the satellite orientation departs from nominal by the maximum conceivable amount. If the average attitude remains close to nominal then each 10 naut. mi. square will actually contain a large percentage of the cloud elements observed there.

SUMMARY AND CONCLUSIONS

More than half of the total error in both pitch and roll attitude is contributed by erroneous signal inputs to the horizon scanner system. Elimination of this error source would result in a reduction of the mean maximum gridding error from 50 naut. mi. to less than 20 naut. mi. A technique involving reconstruction of the scanner input from telemetry data shows promise for removing this error, although initial numerical solutions have exhibited instability in the presence of noise.

Since the operational user is concerned with mean cloud properties over an area rather than the location of individual elements, it is reasonable to relate gridding errors to the computation of areal means. By adopting this point of view, it is possible to show that grid squares of 2° latitude on a side are compatible with the gridding errors expected for Nimbus A.

REFERENCES

1. BARCUS, R. and SPOLLEN, F. J., "A Functional Description of the Stabilization and Control Subsystem for the Nimbus Satellite," General Electric NIMCO Systems Memo No. 1, May 1962.
2. JENSVOLD, R. D., "Nimbus Attitude Inputs to the Gridding Process," General Electric NIMIT PIR No. 9733-010, November 1962.
3. SPOLLEN, F. J., "Revision A, Position Errors and Rates for NIMCO Subsystem," General Electric NIMCO PIR No. 9447-049, August 1962.
4. HANEL, R. A., BANDEEN, W. R., CONRATH, B. J., "The Infrared Horizon of the Planet Earth," No. X-650-62-164, Goddard Space Flight Center External Rept., August 1962.
5. BARRINGER, D. G., "Radius of Uncertainty due to Attitude Errors in the Nimbus Control System," General Electric NIMCO Memo No. 9752-CSA-02, January 1963.
6. STANHOUSE, R. W. and LECHMAN, H. W., "Review of the Earth Signal for the Nimbus Attitude Sensing System," General Electric TIS Rept. No. 62SD222, November 1962.
7. SCHALLER, E., "Determination of Gridding Attitude using the Reconstructed Earth Emittance Wave-shape," General Electric NIMIT PIR No. 9425-016, January 1963.
8. PORITSKY, "Differential Equations of Roll, Pitch and Yaw Error," General Electric NIMCO Note (unnumbered), January 1963.
9. DARLING, E. M., Jr., "Analysis of the Contribution of Meteorological Satellite Data to the Environmental Support of Air Force Weapon Systems," (UNCL TITLE), study prepared for the Nimbus project office, August 1962 (SECRET).
10. BRISTOR, C. L., "Processing Satellite Weather Data—A Status Report—Part I," Proceedings of the 1962 Fall Joint Computer Conference.
11. BARTLETT, R. B. and GREAVES, J., "Nimbus Display Subsystems for Vidicon Error Removal, Attitude Determination and Meteorological Editing," ARA-CON Labs. Final Rept. under U. S. Weather Bureau Contract No. Cwb-10489, January 1963.

APPENDIX A

SCANNER PULSE RECONSTRUCTION

The Nimbus scanner system, to a first approximation, may be regarded as a perfect linear system with transfer function $h(t)$, in the time domain. If the system output in response to the input function $f(t)$ is $g(t)$, then $g(t)$ may be expressed by the convolution integral.

$$g(t) = \int_0^t f(\tau) h(t - \tau) d\tau = f * h.$$

The problem is as follows: Given $h(t)$, the combined optical and electronic transfer function of the scanner system, and $g(t)$ the output wave of the scanner system, find $f(t)$, the system input. For a perfect linear system the solution is straight forward.

Taking the Laplace transform of $g(t)$

$$\zeta [g(t)] = \int_0^\infty e^{-st} g(t) dt = G(s). \quad (A1)$$

$$G(s) = F(s) \times H(s). \quad (A2)$$

where G , F and H are the frequency domain equivalents of g , f and h , respectively, $F(s) = \zeta[f(t)]$, $H(s) = \zeta[h(t)]$, we have, from Equation A2,

$$F(s) = G(s) \div H(s).$$

Thus the desired function $f(t)$ is merely the inverse

$$f\left[\frac{2t-1}{2}\right] = \frac{\frac{g(t)}{\Delta\tau} - \sum_{k=1}^{t-1} f\left[\frac{2k-1}{2}\right] g\left[\frac{2(t-k)+1}{2}\right]}{h\left(\frac{1}{2}\right)},$$

where $t \geq 2$.

The value of $g(t)$ in the numerator contains all the cumulative effects of noise from previous iterations. Furthermore, since $\Delta\tau$ is a very small time interval (about 1 millisecond) all the errors in $g(t)$ are greatly magnified. Weygandt found that this recursion formula works very well in reconstructing a simple input function (such as a

Laplace transform of $F(s)$:

$$f(t) = \zeta^{-1}[F(s)]$$

The uniqueness of this solution is assured by Lerch's theorem.

However the scanner system is not a perfect linear system. Mechanical, optical, and electronic noise are all present in the system transfer function. This noise is amplified by the following numerical solution which was used by Weygandt at General Electric in his first attempts to solve for $f(t)$.

The convolution was expressed as

$$g(t) = \Delta\tau \sum_{i=1}^t f\left(\frac{2i-1}{2}\right) h\left[\frac{2(t-i)+1}{2}\right].$$

By solving for $f(t)$ at successive values of t ; e.g.

$$g(1) = f\left(\frac{1}{2}\right) h\left(\frac{1}{2}\right) \Delta\tau$$

$$f\left(\frac{1}{2}\right) = \frac{g(1)}{\Delta\tau h\left(\frac{1}{2}\right)},$$

it is possible to establish the following recursion formula:

step) in the absence of noise. However the formula fails when applied to complex functions.

Dr. J. G. Jewell of General Electric has investigated that scanner pulse reconstruction problem in an attempt to find solutions which are less susceptible to deterioration in the presence of noise. At present three approaches appear promising.

APPENDIX B

MATHEMATICS OF GRIDDING ERRORS

A separate equation has been derived for the component of gridding error contributed by each axial component of satellite attitude error. The

total gridding error was obtained by combining these components as if they were independent. This procedure introduces negligible computational errors for the small angles involved (see Table 1). The gridding error geometry is shown in Figures B1, B2 and B3.

Pitch Component

The effect of the pitch error angle p is to rotate the plane $YPBN$ about TN the pitch axis so that PN rotates to QN (Figure B1). On the projection plane the pitch error is line PQ ; on the earth,

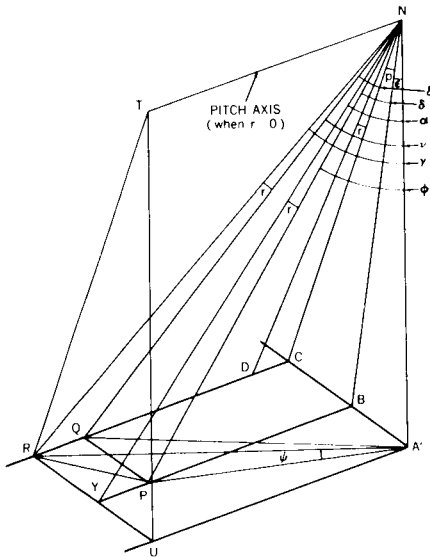


FIGURE B1.—Projection plane geometry for pitch and roll components of gridding error. N satellite position; A' , sub-satellite point; $h=NA'$ satellite height; H radius of the earth. $TNA'U$ is the plane containing the satellite pitch axis and perpendicular to the center camera focal-plane $RCA'U$ is the plane tangent to the earth at A' . For normal satellite attitude $RCA'U$ is perpendicular to $TNA'U$ as shown. $A'BCN$ is the satellite orbital plane. $A'BCN$ is normal to planes $TNA'U$, $RCA'U$, $YPBN$ and $TNCR$. In plane $RCA'U$ the points A' , B , C , D , U , P , Y , Q , R are intersections of rays emanating from the satellite and terminating at the earth's surface. Thus plane $RCAU$ is defined as the *projection plane* for points on the earth. The following projected quantities are defined on $RCAU$: P , the position of a cloud element for nominal satellite attitude, is defined by angles ϵ and δ in orthogonal planes $A'BCN$ and $YPBN$ respectively. PQ is the pitch error due to rotation of plane PNB about TN by pitch error angle p , QR is the roll error due to rotation of ray QN in plane $TNCR$ through roll error angle r (when $p=0$ the roll error is PY), RP is the resultant pitch-roll error.

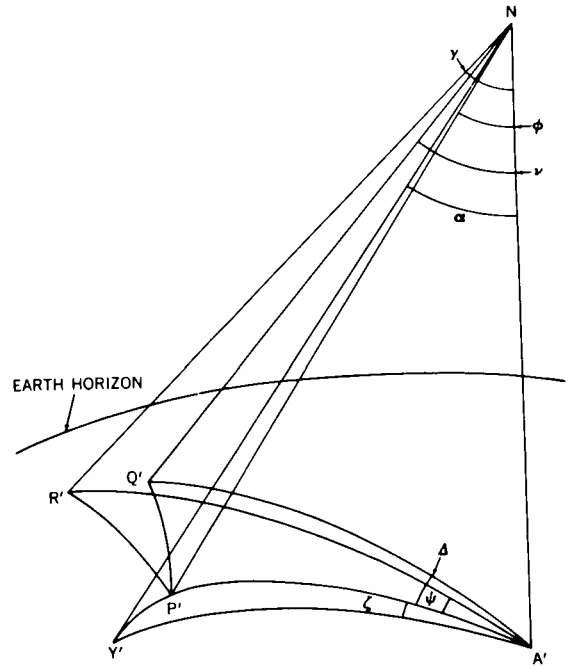


FIGURE B2.—Spherical earth geometry for pitch and roll components of gridding error. N satellite position; A' subsatellite point; $h=NA'$ satellite height; H radius of the earth. All points designated by primed letter are earth locations of like-lettered, unprimed, projected points on plane $RCA'U$ (Figure B1). The arcs connecting these earth points correspond to like-lettered lines in Figure B1.

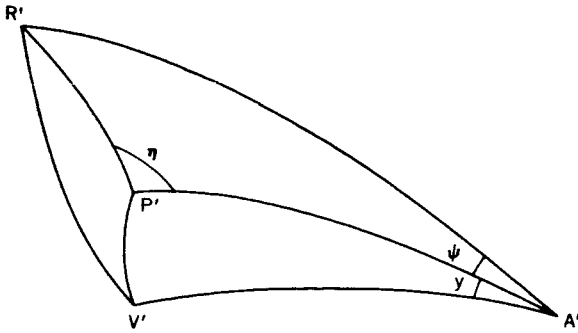


FIGURE B3.—Spherical earth geometry of yaw error $P'V'$ and total error $R'V'$. $P'V'$ is the yaw error due to rotation of arc $A'V'$ through yaw error angle γ ; $A'P' = A'V'$. $R'P$ is the resultant pitch-roll error; and $R'V'$ is the resultant pitch-roll-yaw error.

arc $P'Q'$ (Figure B2). $P'Q'$ may be found by applying the law of cosines to the spherical triangle $A'P'Q'$. In order to find $P'Q'$, the great circle arcs $A'P'$ and $A'Q'$ must be determined as well as angle Δ , where $A'P'$ subtends angle ϕ at the satellite, and $A'Q'$ subtends ν . The solution involves expressing ϕ and ν in terms of the known angles p , r , ϵ and δ .

From right triangles $A'CN$, QCN and QNA' (Figure B1),

$$\nu = \sec^{-1} [\sec (\epsilon + p) \sec \delta] ; \quad (\text{B1})$$

from right triangles $A'BN$, PBN and $A'BP$,

$$\phi = \tan^{-1} \left[\sec \epsilon \tan \delta \sec \left[\tan^{-1} \left(\frac{\sin \epsilon}{\tan \delta} \right) \right] \right]. \quad (\text{B2})$$

From right triangles $A'CQ$ and $A'BP$:

$$\Delta = \tan^{-1} \left[\frac{\sin (\epsilon + p)}{\tan \delta} \right] - \tan^{-1} \left(\frac{\sin \epsilon}{\tan \delta} \right). \quad (\text{B3})$$

In general the great circle arc XY (in degrees) which subtends an angle A at the satellite (measured from the local vertical) may be expressed as

$$XY = \sin^{-1} \left(\frac{h + H}{H} \sin A \right) - A, \quad (\text{B4})$$

where h is the satellite altitude and H is the earth's radius.

We can now solve spherical triangle $A'P'Q'$ for $P'Q'$:

$$P'Q' = \cos^{-1} (\cos A'Q' \cos A'P' + \sin A'Q' \sin A'P' \cos \Delta), \quad (\text{B5})$$

where, from Equation B4,

$$\begin{aligned} A'Q' &= \sin^{-1} \left(\frac{h + H}{H} \right) \sin \nu - \nu, \\ A'P' &= \sin^{-1} \left(\frac{h + H}{H} \right) \sin \phi - \phi, \end{aligned} \quad (\text{B6})$$

and Δ is given by Equation B3.

Roll Component

The roll error, PY Figure B1, results from the rotation of PN by roll error angle r to position YN . Proceeding as before the resultant gridding

Solving the spherical triangle $A'P'Y'$ for $P'Y'$:

$$P'Y' = \cos^{-1} (\cos A'Y' \cos A'P' + \sin A'Y' \sin A'P' \cos \zeta) \quad (\text{B9})$$

where, from Equation B4,

$$A'Y' = \sin^{-1} \left(\frac{h + H}{H} \sin \alpha \right) - \alpha,$$

and $A'P'$ is given by Equation B6; ζ by Equation B8.

error is arc $P'Y'$ of spherical triangle $A'P'Y'$ (Figure B2) which we find by expressing α and ζ in terms of p , r , ϵ and δ , using right triangles in Figure B1.

From right triangles $A'BN$, YNB , and YNA' , we have

$$\alpha = \sec^{-1} [\sec \epsilon \sec (\delta + r)] \quad (\text{B7})$$

and from $A'BP$ and $A'BY$

$$\zeta = \tan^{-1} \left(\frac{\sin \epsilon}{\tan \delta} \right) - \tan^{-1} \left[\frac{\sin \epsilon}{\tan (\delta + r)} \right]. \quad (\text{B8})$$

Yaw Component

P'V' (Figure B3) is the yaw error resulting from the rotation of arc A'V' through yaw error angle y ($A'V' = A'P'$).

Solving the spherical triangle P'V'A', we obtain

$$P'V' = \cos^{-1} (\cos^2 A'P' + \sin^2 A'P' \cos y) \quad (B10)$$

Pitch-Roll Resultant

Assuming that the satellite first pitches through angle p and then rolls through angle r , the resultant is RP in the projection plane and R'P' on

$$R'P' = \cos^{-1} (\cos A'R' \cos A'P' + \sin A'R' \sin A'P' \cos \psi) \quad (B13)$$

where, from Equation B4

$$A'R' = \sin^{-1} \left(\frac{h+H}{H} \sin \gamma \right) - \gamma$$

A'P' is given by Equation B6; and ψ by Equation B12.

Total Gridding Error

The total error is R'V' in Figure B3 which is the resultant of R'P' the combined pitch-roll error, and P'V' the yaw error.

$$R'V' = \cos^{-1} (\cos P'V' \cos R'P' - \sin P'V' \sin R'P' \sin \eta) \quad (B15)$$

where P'V' is given by Equation B10, R'P' by Equation B13 and η by Equation B14. Note the sign of the second term in Equation B15 is negative if ψ is positive and vice versa.

Computation of Gridding Errors

The gridding errors calculated by the above equations are listed in Table B1 where the three axial components and the total error are tabulated versus angular distance (measured at the satellite) from the center picture principal point. The dominance of the roll error is apparent, especially at the larger angular distances, for all methods except scanner pulse reconstruction where the yaw error predominates. Pitch error makes the smallest contribution to total error in all cases.

The total gridding error has been plotted versus the angle from the center picture principal point (Figures B4-B8) for the five attitude determination methods (Table 2). In each case the error

the earth. In this case we need angles γ , ϕ and ψ to solve spherical triangle R'P'A'.

In Figure B1 we have, from right triangles A'NC, CNR and RNA',

$$\gamma = \sec^{-1} [\sec (\epsilon + p) \sec (\delta + r)] \quad (B11)$$

And from right triangles A'BP and A'CR.

$$\psi = \tan^{-1} \left[\frac{\sin (\epsilon + p)}{\tan (\delta + r)} \right] - \tan^{-1} \left(\frac{\sin \epsilon}{\tan \delta} \right) \quad (B12)$$

Solving the spherical triangle R'P'A', we have

Solving first for angle η , we obtain

$$\eta = \cos^{-1} \left(\frac{\cos A'R' - \cos R'P' \cos A'P'}{\sin R'P' \sin A'P'} \right) \quad (B14)$$

Finally, solving spherical triangle P'V'R':

increases quite gradually out to about 35° (i.e. the center of the side picture) and then increases steeply from there to 53.5° , the edge of the field of view.

Mean Maximum Error

An approximation to the mean error along a diagonal from the center picture principal point to the side picture corner has been computed as follows. Equation B11 was solved for $\epsilon = 18.5^\circ$, $\delta = 53.5^\circ$, and the appropriate values of p and r for each method in Table 2 to obtain the angle γ corresponding to the diagonal in each case. The results were:

Original Design	56°49'
Present Design (Disregard telemetry)	58°46'
Present Design (Use telemetry)	58°02'
Scanner Pulse Reconstruction	56°14'

Table B1
Components of the Total Gridding Error.

Angle from Center Picture Principal Point (degrees)	Components of the Total Gridding Error (naut. mi.)			Total Gridding Error (naut. mi.)
	Pitch	Roll	Yaw	
Original Design				
0	9	9	0	9
10	9	9	2	9
20	9	10	4	11
30	10	12	7	16
40	10	18	10	25
50	11	32	13	43
53°30'	13	46	17	56
Present Design (Disregard TM)				
0	13	26	0	31
10	13	28	3	34
20	14	31	7	39
30	15	39	11	49
40	18	57	16	66
50	23	108	22	116
53°30'	26	160	28	169
Present Design (Use TM)				
0	9	20	0	21
10	9	21	3	23
20	9	23	7	28
30	10	29	11	38
40	11	43	16	54
50	12	80	22	88
53°30'	14	116	28	123
Scanner Pulse Reconstruction (Yaw Error 2°)				
0	4	4	0	5
10	4	5	3	6
20	4	5	7	9
30	5	6	11	15
40	5	9	16	23
50	6	16	22	35
53°30'	7	22	28	42
Scanner Pulse Reconstruction (Measure Yaw)				
0	4	4	0	5
10	4	5	1	6
20	4	5	2	8
30	5	6	4	11
40	5	9	6	15
50	6	16	8	23
53°30'	7	22	10	28

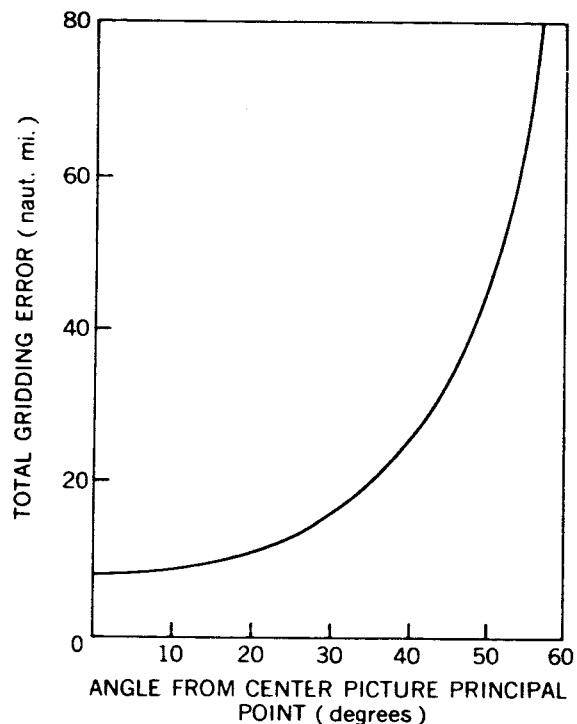


FIGURE B4.—Gridding errors (original design). The pitch, roll and yaw errors are each ± 1 degrees.

The mean error was computed for each case by numerically integrating the appropriate error curve (i.e. Figures B4-B8) from 0 to γ and dividing by the included area. This number has been designated as the *mean maximum error*. The value for each method is entered in the bottom row of Table 2.

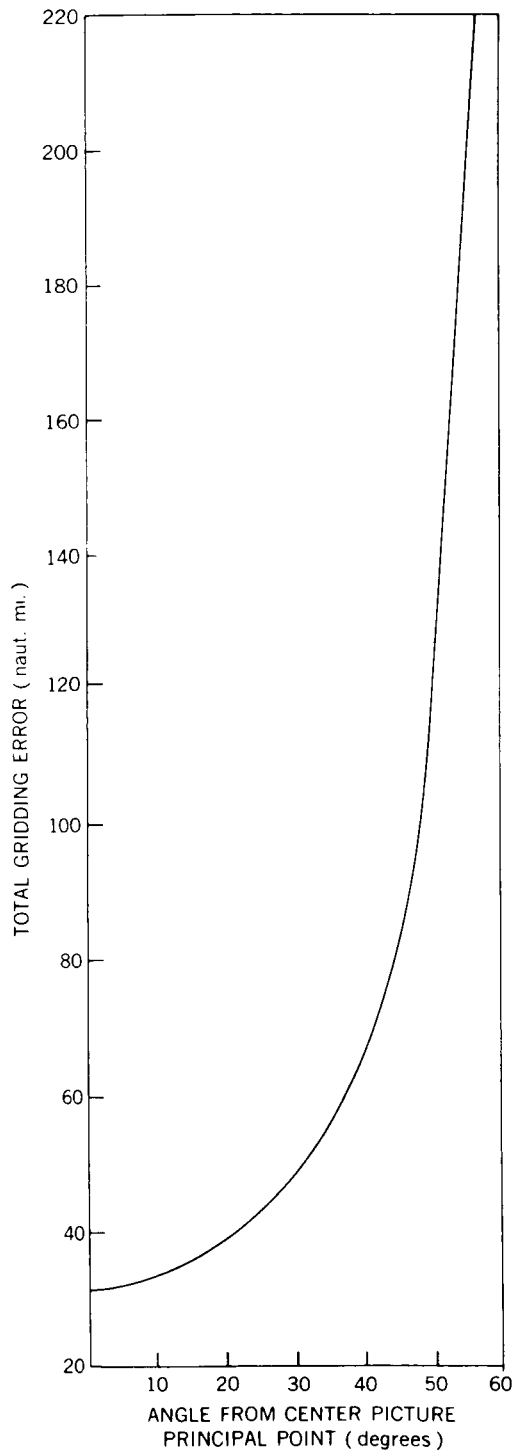


FIGURE B5.—Gridding errors, present design (disregard *TM*). The pitch error is ± 1.5 degrees; the roll error ± 3 degrees, and the yaw error ± 2 degrees.

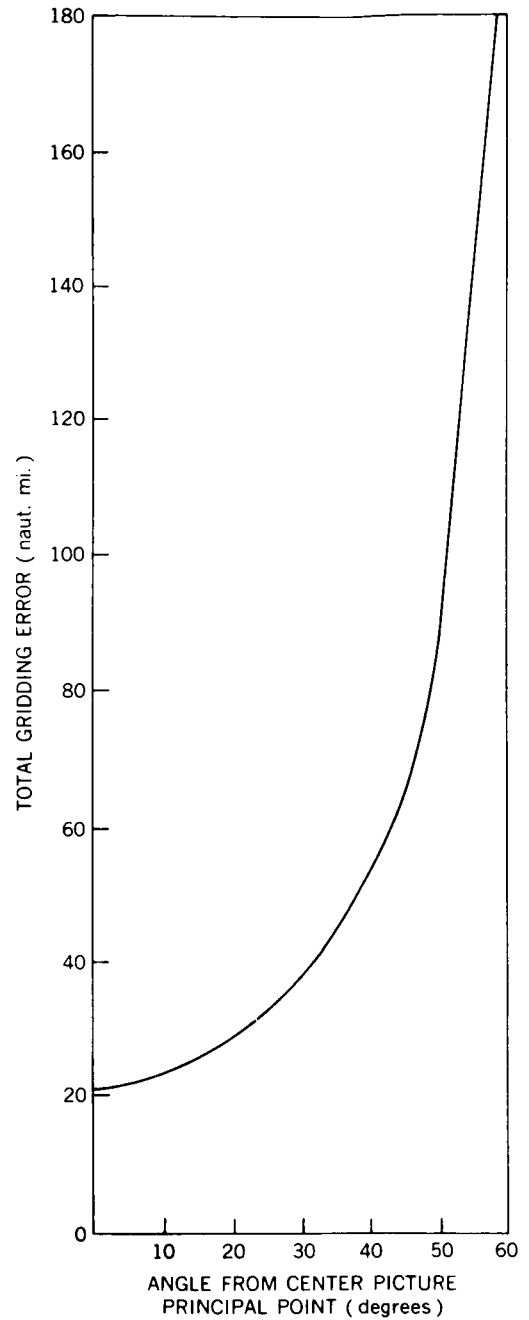


FIGURE B6.—Gridding errors, present design (use *TM*). The pitch error is ± 1.1 degrees, the roll error ± 2.3 degrees and the yaw error ± 2 degrees.

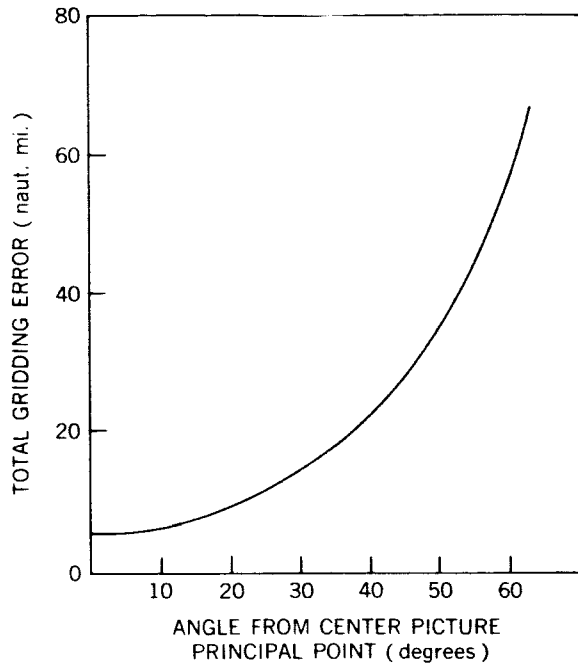


FIGURE B7.—Gridding error, scanner pulse reconstruction technique (disregard yaw). The pitch error is ± 0.5 degree, the roll error ± 0.5 degree, and the yaw error ± 2 degrees.

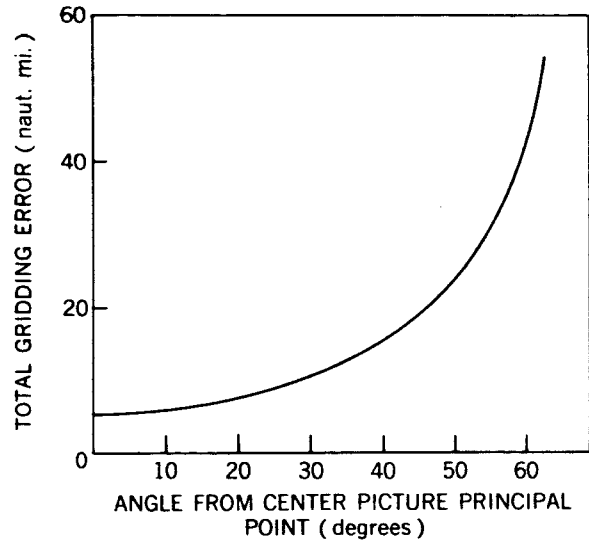


FIGURE B8.—Gridding errors, scanner pulse reconstruction technique (measure yaw). The pitch error is ± 0.5 degree, the roll error ± 0.5 degree, and the yaw error ± 0.7 degree.

A GEOGRAPHICAL GRID FOR NIMBUS CLOUD PICTURES*

EUGENE M. DARLING, JR.

This paper defines the geometrical, aesthetic and utilitarian factors which must be considered in selecting a geographical reference grid for use with Nimbus cloud pictures. A particular grid is proposed and critically examined with respect to these factors. This grid is found to be acceptable once minor changes are made to compensate for image foreshortening.

INTRODUCTION

The Nimbus meteorological satellite will carry an Advanced Vidicon Camera System (AVCS) consisting of three separate cameras in a trimetron array. This camera system views a 338×1600 naut. mi. area on the earth (the longer dimension being normal to the subsatellite track). Pictures are taken over the daylight portion of the orbit, stored on magnetic tape and transmitted to the ground station on command. These data are acquired in the form of FM signals which are recorded on tape. Satellite attitude and position at picture-taking times are automatically processed (Reference 1) by a computer program which generates a latitude-longitude grid for each picture. The gridding mathematics is presented in Reference 2.

The latitude and longitude lines consist of individual gridpoints which are mixed with the video signal, under control of a computer, during transmission over the long line to Washington, D.C. These gridpoints are composed of a sequence of white pulses followed by a sequence of black pulses, each pulse having a width of one picture element. There are provisions in the grid mixing equipment for varying the gridpoint size from a minimum width of 5 pulses (3 white, 2 black) to a maximum of 10 (5 white, 5 black). Sufficient gridpoints will be generated along each latitude and longitude to create the illusion of a continuous line when the picture is displayed through the kinescope system which produces a 70 mm positive transparency.

This paper examines the geometrical, aesthetic and utilitarian properties which a latitude-longitude grid suitable for Nimbus cloud pictures must exhibit. We start with a review of the mathematical characteristics of latitude and longitude lines on the earth; proceed to enumerate six factors pertinent to the Nimbus gridding problem; evaluate a particular grid in terms of these factors and summarize findings, including a recommendation for minor improvements.

It should be noted that this study is concerned solely with the problem of imposing geographical grids directly upon Nimbus cloud pictures. Rectification of data to a specific map projection is not considered. Furthermore, it has been assumed that errors due to lens distortion, vidicon raster shift and other AVCS aberrations have been removed prior to the mixing of gridpoints with the video data.

GEOGRAPHIC REFERENCING ON A SPHERICAL EARTH

The parallels of latitude are small circles of radius, $R \cos \phi$, normal to the axis of rotation where R is the radius of the earth and ϕ is the latitude (in degrees) measured north and south from zero at the equator to 90° at the poles. The parallels are equally spaced over the surface of the earth at a fixed interval of 60 naut. mi. per degree.

The meridians of longitude, θ are great circles formed by the intersection of planes containing the rotation axis of the earth. Since all meridians pass through the poles, the distance D corresponding to one degree of longitude diminishes from a

*Published as *NASA Technical Note D-2136*, February 1964.

maximum of 60 naut. mi. at the equator to zero at the poles. At any latitude ϕ ,

$$D = \frac{2\pi R \cos \phi}{360}$$

The function, D is shown in Table 1. Longitude is measured west and east from zero at the prime meridian (passing through Greenwich, England) to 180° at the International date line.

Table 1

Distance D Corresponding to One Degree of Longitude.

Latitude (degrees)	D (naut. mi.)	Latitude (degrees)	D (naut. mi.)
0	60.0	81	9.4
10	59.2	82	8.4
20	56.5	83	7.3
30	52.1	84	6.3
40	45.9	85	5.2
50	38.6	86	4.2
60	30.0	87	3.1
70	20.5	88	2.1
75	15.5	89	1.0
80	10.4	—	—

PROPERTIES OF THE NIMBUS GRID

The following factors must be taken into account in the design of a suitable grid for Nimbus cloud pictures: 1. Constancy of scale; 2. Image foreshortening; 3. Satellite attitude; 4. Obscuration of data; 5. Aesthetics; and 6. Data utilization. With the exception of scale constancy all of these considerations are either peculiar to satellite photography or have a special connection with the Nimbus system for constructing gridlines. A Nimbus grid has been selected which appears to represent a reasonable compromise between the conflicting demands of the above factors. Specifications for this grid are listed in Table 2 and sample grids for various portions of the earth are shown in Figure 1. Characteristics of the selected grid are evaluated in the following paragraphs.

Constancy of Scale

An ideal gridding scheme would be one where the interval between successive grid lines on the earth corresponds everywhere to a constant, fixed distance. Obviously such a result cannot be

Table 2

Nimbus Grid Specifications.

Latitude Range (degrees North or South)	Longitude Interval (degrees)	Latitude Interval (degrees)
0-70	2	2
70-84	10	2
84-88	40	2
88-90	No grid lines	

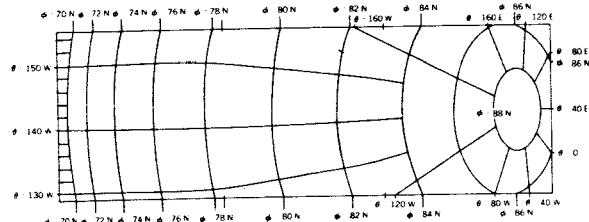


FIGURE 1a.—Nimbus grid, polar.

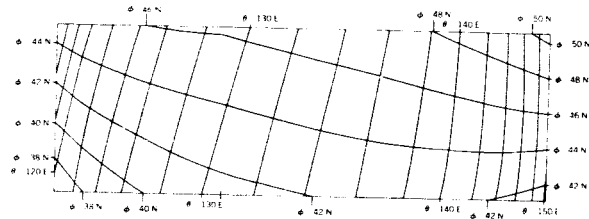


FIGURE 1b.—Nimbus grid, mid-latitude.

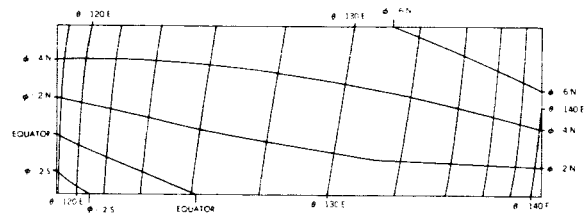


FIGURE 1c.—Nimbus grid, equatorial.

achieved in reality due to the convergence of meridians (Table 1). The best compromise is a fixed latitude scale used in conjunction with a variable longitude scale. The longitude interval must increase with increasing latitude in order to maintain an approximately constant scale.

The scale characteristics of the Nimbus grid are shown in Figure 2. The longitude scale changes by a factor of 3 within each of the three

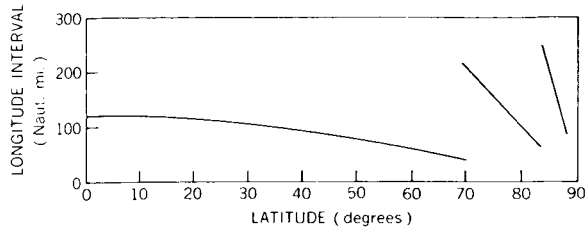


FIGURE 2.—Scale variability of the nimbus grid.

latitude bands. At the two discontinuities, 70° and 84°, the scale changes by 5 and 4, respectively. The latitude scale remains constant at 120 naut. mi. Although the variability in longitude distance is large when the entire globe is considered, it should be noted that the scale only varies by a factor of 2 between 60°S and 60°N latitude* which is the region of greatest meteorological interest. From this viewpoint the Nimbus grid represents an acceptable compromise in regard to scale constancy.

Image Foreshortening

The projection on the satellite camera focal plane of a fixed arc length on the earth's surface is a line segment which varies in length according to the location of the arc within the camera field of view. For an earth stabilized satellite the projected segment is largest when the arc appears in the center of the field of view. The segment diminishes in size as the arc shifts towards the edge of the field. This phenomena is called image foreshortening.

In order to develop the mathematics of foreshortening, we make use of the following two equations relating an angle β subtended at the satellite to the corresponding central angle α . Referring to Figure 3 it can be shown that

$$\beta = \tan^{-1} \left[\frac{R \sin \alpha}{h + R (1 - \cos \alpha)} \right], \quad (1)$$

and conversely

$$\alpha = \sin^{-1} \left[\frac{R + h}{R} \sin \beta \right] - \beta. \quad (2)$$

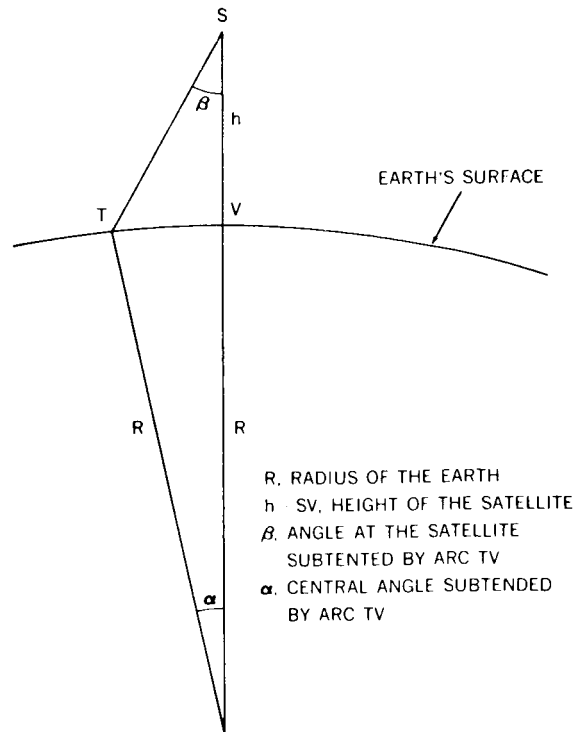


FIGURE 3.—Satellite geometry.

Single Focal Plane

Consider first the case where the earth is viewed by a single camera on an earth stabilized satellite (See Figure 4a). It is possible to express the angles β and γ , which define the degree of shortening, in terms h , R , α and ψ .

From Equation 2

$$\psi = \sin^{-1} \left[\frac{R + h}{R} \sin \delta \right] - \delta.$$

By successive application of Equation 1, γ is found as follows:

$$\gamma + \delta = \tan^{-1} \left\{ \frac{R \sin (\alpha + \psi)}{h + R [1 - \cos (\alpha + \psi)]} \right\},$$

$$\delta = \tan^{-1} \left[\frac{R \sin \psi}{h + R (1 - \cos \psi)} \right];$$

and by subtraction,

$$\gamma = \tan^{-1} \left\{ \frac{R \sin (\alpha + \psi)}{h + R [1 - \cos (\alpha + \psi)]} \right\} - \tan^{-1} \left[\frac{R \sin \psi}{h + R (1 - \cos \psi)} \right], \quad (3)$$

*This region constitutes 86.6 percent of the total surface area of the earth.

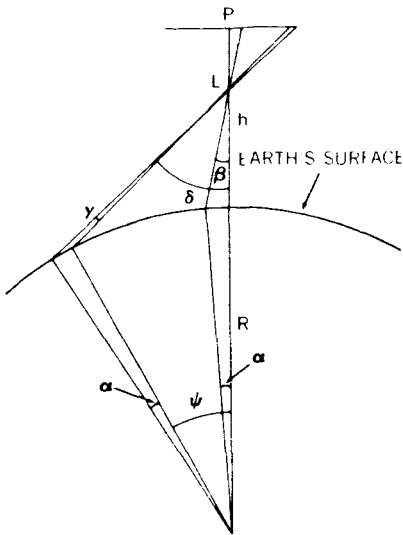


FIGURE 4a.—Foreshortening, single focal plane.

- R. RADIUS OF THE EARTH
- h. HEIGHT OF SATELLITE
- P. FOCAL PLANE CENTER (PRINCIPLE POINT)
- L. LOCATION OF LENS
- f. PL. CAMERA FOCAL LENGTH
- δ . ψ . SATELLITE ANGLE AND CORRESPONDING EARTH ANGLE ASSOCIATED WITH A POINT IN THE CAMERA FIELD OF VIEW
- α . A FIXED EARTH DISTANCE
- β . SATELLITE ANGLE CORRESPONDING TO α AT THE CENTER OF THE FIELD OF VIEW
- γ . SATELLITE ANGLE CORRESPONDING TO α AT AN ANGLE δ FROM THE PICTURE CENTER

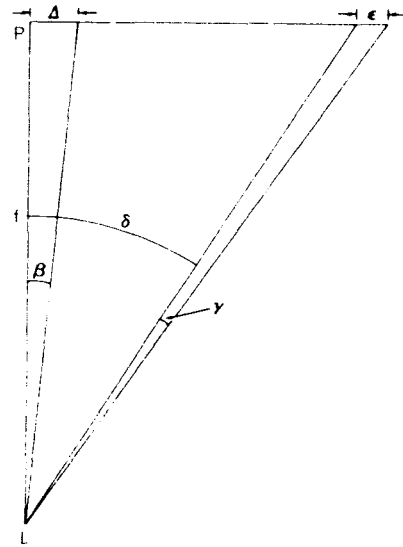


FIGURE 4b.—Focal plane detail.

where α is arbitrary. Referring to Figure 4b we see that the focal plane distances Δ and ϵ corresponding to angles β and γ , respectively, may be expressed as follows:

$$\Delta = f \tan \beta ; \tag{4}$$

$$\epsilon = f [\tan (\delta + \gamma) - \tan \delta] . \tag{5}$$

The foreshortening, F, in percent is

$$F = 100 \frac{(\Delta - \epsilon)}{\Delta} = 100 \left[1 - \frac{\tan (\delta + \gamma) - \tan \delta}{\tan \beta} \right] . \tag{6}$$

Trimetrogon Array

In the case of a trimetrogon camera array such as the Nimbus system the foreshortening in the center camera focal plane is defined by Equation 6. However, the projection geometry for the side camera is altered due to the angle of inclination ν of the side camera focal plane (see Figures 5a and b). The angle γ corresponding to α is

defined by Equation 3 as in the single focal plane case.

From Figure 5b we see that the focal plane distance ϵ corresponding to γ is

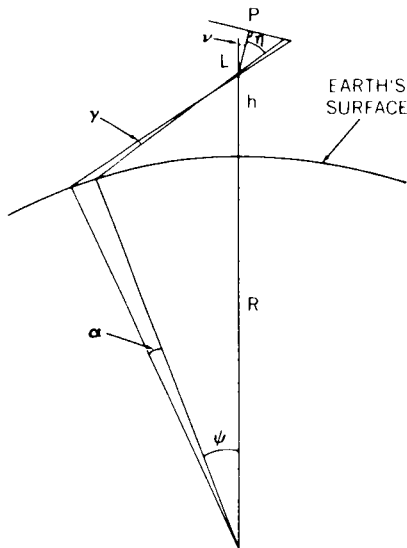
$$\epsilon = f [\tan (\eta + \gamma) - \tan \eta] . \tag{7}$$

In this case the foreshortening, F, in percent with respect to the center camera principal point is

$$F = 100 \frac{(\Delta - \epsilon)}{\Delta} = 100 \left[1 - \frac{\tan (\eta + \gamma) - \tan \eta}{\tan \beta} \right] . \tag{8}$$

In Figure 6 the percent foreshortening as defined by Equations 6 and 8 ($\nu=35^\circ$, i.e., the Nimbus System) is plotted versus angular distance δ from the center picture principal point. These curves

were computed for an earth increment α of one degree. Note that the image foreshortening is negligible within the center camera field of view, being only 3 percent at the edge ($\delta=18.5^\circ$). At



ν . ANGLE OF INCLINATION OF SIDE CAMERA PRINCIPLE AXIS WITH RESPECT TO CENTER CAMERA AXIS
 η . SATELLITE ANGLE CORRESPONDING TO A POINT IN THE SIDE CAMERA FIELD OF VIEW

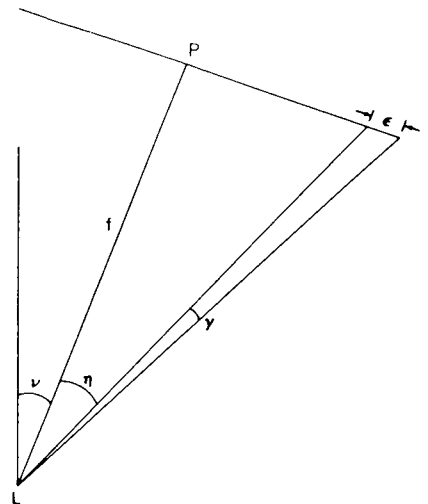


FIGURE 5a.—Foreshortening, side camera focal plane.

FIGURE 5b.—Side camera focal plane detail.

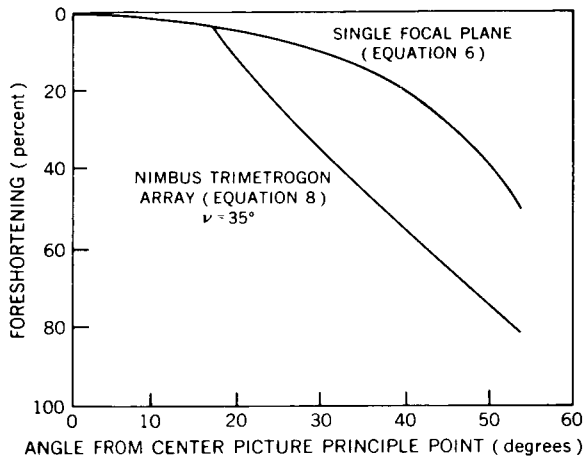


FIGURE 6.—Foreshortening in the image plane.

the 53.5° limit of the Nimbus field of view the foreshortening is 50 percent with a single focal plane and 80 percent with the trimetrogon array.

Thus we see that image plane foreshortening causes an apparent crowding together of earth objects near the edge of the camera system field of view, which has an effect similar to the actual convergence of meridians with increasing distance away from the equator. The combined result of convergence and foreshortening is illustrated in Table 3 which shows the length (mm) of 1° longitude on a 70 mm picture taken at 500 naut. mi. by the Nimbus camera system. In this case the

Table 3

Effect of Convergence and Foreshortening.

Latitude (degrees)	Length of 1° longitude, 70 mm picture (mm)	
	Center Picture Principal Point	Side Picture Edge
0	8.9	1.8
60	4.4	0.9

foreshortening is 80 percent, the convergence is 50 percent and the maximum combined effect is a reduction in length by an order of magnitude.

The effect of image foreshortening is further illustrated by the following example. On 70 mm film the spacing between meridians of the Nimbus grid at various points within the trimetrogon camera field of view is shown in Table 4. The interval of 1 to 4 mm between meridians at the side camera edge is too small for convenient viewing of the picture. In this extremely foreshortened portion of the image either the grid interval must be increased or gridding must be omitted in order to avoid crowding of the meridians. It would be desirable to modify the gridding scheme in order to maintain a minimum spacing of 5–10 mm throughout the 70 mm picture.

Table 4
Foreshortening of Meridians.

Latitude (degrees)	The Spacing of Meridians (mm)		
	Center Camera Principal Point	Side Camera Center	Side Camera Edge
0	17.7	9.9	3.6
30	15.4	8.6	3.1
50	11.4	6.4	2.3
70* (2° spacing)	6.1	3.4	1.2

*Beyond 70° the convergence of meridians has a greater effect than foreshortening on longitude spacing.

Satellite Attitude

In a previous study (Reference 3) the author has demonstrated that uncertainties in satellite attitude are expected to result in a mean maximum error of 50 to 70 naut. mi. in the grids computed for Nimbus cloud pictures. This type of error increases from 20–30 naut. mi. at the center picture principal point to 45–55 naut. mi. at the side picture center to 120–170 naut. mi. at the side picture edge. This effect constitutes an *actual location error* in contrast to the previous factors which merely contribute to variations in size. The reduction of gridding errors due to attitude uncertainties will be undertaken as an experiment using data from the first Nimbus satellite (i.e., Nimbus A).

In Reference 3 Darling shows that grid squares of about 125 naut. mi. in a side are optimum for determining mean cloudiness in the presence of a maximum gridding error of 50 naut. mi. due to anomalous satellite attitude. To a good approximation the Nimbus gridlines form squares of the desired size in the region between 40°S and 40°N latitude (see Figure 1b, c). However the convergence of meridians beyond 40° causes the "squares" to become increasingly trapezoidal (see Figure 1a). At 60° two adjacent boxes along a parallel must be combined to obtain squares of the proper size; at 70° three boxes must be joined. Beyond 70° the rapid changes in scale preclude the division of grid boxes into 120 naut. mi. squares.

Obscuration of Data

Every gridpoint imposed on a Nimbus cloud picture obliterates the data which originally occu-

ried the gridpoint position. The gridpoint sizes to be tested in early Nimbus experiments will range from a minimum width of 5 pulses (3 white, 2 black) to a maximum of 10 pulses (5 white, 5 black). These limits correspond, respectively, to earth distances of 2.0 and 4.1 naut. mi. at the center picture principal point and 10.2 to 20.5 naut. mi. at the side picture edge. Thus, it is clear that even the smallest gridpoint will obscure individual clouds of dimensions smaller than a mature cumulonimbus. It is equally apparent that the largest gridpoint will cover only an infinitesimal portion of the cloud field associated with a synoptic-scale disturbance. Therefore, we may conclude that individual gridpoint sizes within the range cited will not obscure those major cloud features which are most significant for operational applications.

The requirement that the density of points be sufficient to form a line conflicts with the need to minimize the number for the purpose of conserving data. In forthcoming experiments a range of gridpoint sizes and point densities along grid lines will be examined for readability and suitability. The following maximum sizes will be associated with the densities indicated:

Maximum Gridpoint Size	Density
5 elements;	1 point/15 elements
10 elements;	1 point/30 elements

These density values correspond to 0.1° and 0.2° latitude intervals at the center picture principal point. Note that both combinations result in the same *total* number of elements per triplet occupied

Table 5
Data Obscuration.

Geographic Region	Number of gridpoints per triplet Gridpoint size: 5 elements Gridpoint density: 1/15 elements	Percent of data obscured per triplet
Polar	1570	0.38
Mid Latitude	1930	0.46
Equatorial	1430	0.34

by gridpoints. The obscuration of data by gridpoints is documented in Table 5 where the percentage of data obscured was computed by multiplying the number of gridpoints per triplet times the gridpoint size and dividing by 2,085,000, the total number of elements. Since less than half of one percent of the total data is obscured in the worst case, it is clear that the degrading effect of gridlines on the picture is negligible.

Aesthetics

There are many problems involved in achieving a grid matrix of variable scale which is, at the same time, pleasing to the eye. Obviously, the number of discontinuities in gridlines between the equator and the poles must be kept to a minimum as the eye is disturbed by lines which end abruptly. However the demand for few discontinuities is at variance with the need for a quasi-constant scale which requires many changes in grid interval. In order to maintain a visual impression of continuity in the gridlines it is important that certain meridians be retained in their entirety. However, the number must be judiciously chosen since too many will result in the crowding of meridians near the poles while too few will provide an inadequate scale for locating features of meteorological interest.

The question of aesthetics also enters into the appearance of gridlines on the Nimbus cloud pictures. These lines are generated by a series of gridpoints consisting of several black and white pulses, each an element wide. These gridpoints must be sufficiently large to be clearly visible on 70 mm film and yet not so large as to be out of proportion on the greatest enlargement contemplated (presently undetermined, but probably not larger than one foot square). Furthermore, the

gridpoints should be spaced at sufficiently small intervals to cause the human eye to see a continuous line.

The appearance of the Nimbus grid will be generally pleasing to the eye. Visual continuity over the globe is achieved by retaining nine meridians in their entirety from the equator to the poles. Scale discontinuities have been kept to the irreducible number of two, both of which occur poleward of the Arctic Circle (i.e., at 70° and 84°). Thus there is no change in scale over the major continental areas of the world.

The width of a gridline varies as a function of orientation with respect to raster lines. This is another factor which affects the aesthetic quality of the grid. If θ is the angle between a raster line and a gridline consisting of gridpoints of ℓ elements length, then the width W of the gridline is

$$W = \ell \sin \theta \text{ for } \theta > 0$$

$$= r \text{ for } \theta = 0$$

where r is the width of a raster line. Thus the width of a gridline varies from ℓ when the line is normal to the raster line to r when it lies along a raster line. Table 6 shows gridline width as a function of gridpoint size, display size and gridline orientation. The 10 to 1 width ratio in the 10 element case appears to be aesthetically undesirable, especially in the one foot presentation where the maximum width is greater than one tenth of an inch. In this case the longitudes will look like ribbons and the latitudes like threads. The visible disparity in line width will be reduced by the tendency of the eye to see only the white portion of the line in dark areas and the black portion in light areas. However this optical

Table 6
Gridline Width.

Gridpoint Characteristics	Gridline Width (in)			
	5 elements		10 elements	
Display Size	70 mm	1 foot	70 mm	1 foot
Normal to raster line	.012	.051	.024	.103
Along raster line	.002	.010	.002	.010

effect will probably not be sufficient to render this large line size acceptable. On the other hand the five element line which seems to be suitable on both display sizes, will appear even more uniform in width as a result of this visual effect.

With a spacing of 15 elements between gridpoints each gridline will consist of spots separated by 0.9 mm (.036 in) on the 70 mm presentation. The eye should have no difficulty in supplying the additional information required to see the lines as continuous. On a one foot picture the same density of points yields a spacing of 1 every 0.15 in. which will adequate to define the line, but its appearance will be distinctly dotted rather than continuous.

Still another aesthetic factor is the ratio of the space between gridlines to the maximum width of a line. This ratio is shown in Table 7 for lines of 5 element width at three places in the trimetrogon field of view. An examination of sample grids indicates that a spacing to width ratio of less than 20 results in an undesirable crowding of gridlines. This confirms our previous

conclusion that an adjustment in the grid is required at the foreshortened edge of the image.

Data Utilization

In Reference 3, Darling suggested that field meteorologists are mainly concerned with mean cloud conditions rather than individual clouds. This being the case it is shown that the important operational parameter is the mean probability that clouds are actually located in the grid square where they appear. Taking into consideration the maximum gridding error due to attitude anomaly, it is deduced that a 70 percent mean probability (which corresponds to a 125 naut. mi. grid interval) would satisfy most operational requirements. This analysis is carried a step further in Figure 7 where the probability of determining the correct category of cloudiness is plotted against cloud location probability for two cases:

- (1) The categories are cloud (6/10-10/10 coverage) or no cloud (0-5/10),
- (2) The categories are overcast (8/10-10/10), broken (4/10-7/10) or scattered (0-3/10).

Table 7
Gridline Spacing.

Latitude (degrees)	Spacing Between Gridlines Divided by Gridline Width		
	Center Camera, Principal Point	Side Camera Center	Side Camera, Edge
0	59.1	33.1	11.9
30	51.3	28.7	10.3
50	38.1	21.3	7.6
70 (2° spacing)	20.2	11.3	4.1

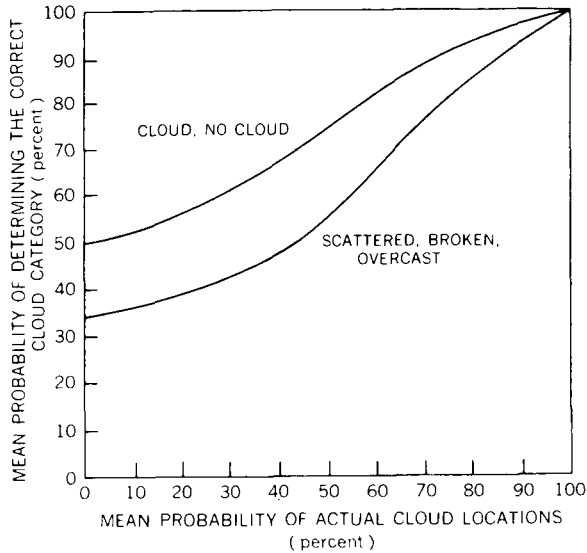


FIGURE 7.—Categorization of cloudiness in the presence of cloud location errors.

Note that 70 percent cloud location probability corresponds to 89 and 76 percent in the correct category for cases (1) and (2), respectively. From an operational point of view the discrimination between cloud and no cloud is probably the most important deduction which can be made from the pictures. The 89 percent confidence level associated with this type of binary decision is certainly adequate. Since this high confidence level corresponds to the maximum gridding error anticipated, we may conclude that the previously deduced 70 percent cloud location probability and attendant 125 naut. mi. grid interval are supported by this analysis.

We have indicated that the Nimbus grid is adequate to determine mean cloudiness in grid

squares 120 naut. mi. on a side, between 70°S and 70°N latitude, even in the case of maximum satellite attitude deviation from nominal. In fact the grid was specifically designed for this purpose. The grid is less adaptable to the location of individual cloud features due to the long distances over which interpolation must be performed.

The error to be expected in locating individual features may be estimated by making the following assumptions about the accuracy of interpolation as a function of grid spacing:

On 70 mm film the error due to interpolation is 1/10 the distance between gridlines for spacings ≥ 10 mm; 1/5 for spacings of 3 to < 10 mm; $1/2$ for < 3 mm.

Combining the latitude and longitude errors by root mean square yields the results shown in Table 8. Note that these errors are due to interpolation alone and do not include the effect of attitude errors. Table 8 shows that clouds may be interpolated with an error of less than 25 naut. mi. in the area between the side camera centers. The error is less than 45 naut. mi. elsewhere. Considering the 30–50 naut. mi. position error in upper level winds used in the analysis of cloud pictures, these inaccuracies in cloud location cannot be considered excessive.

Summary and Conclusions

The proposed Nimbus grid will serve its purpose admirably because of the following properties:

1. Reasonable constancy of scale over regions of prime meteorological concern.
2. A grid interval especially suited to the determination of mean cloudiness which, at the same time, permits the location of individual clouds with acceptable accuracy.

Table 8

Interpolation Errors.

Latitude (degrees)	Errors due to Interpolation (naut. mi.)		
	Center Camera Principal Point	Side Camera, Center	Side Camera, Edge
0	17.0	17.0	26.9
30	15.9	24.0	24.0
50	14.3	19.5	40.3
70 (2° spacing)	14.5	14.5	23.8
70 (10° spacing)	23.8	23.8	42.7

3. Negligible obscuration of data by gridlines.

4. A pleasing balance of line width and spacing over most of the three picture area.

A slight alteration of the grid interval is required in the extremely foreshortened edges of the side pictures in order to avoid crowding of meridians.

REFERENCES

1. "Nimbus Command and Data Acquisition (CDA) Station Data System Performance Requirements and Equipment Interface Specifications," General Electric Document No. 63SD557, Revision A, May 30, 1963 (Contract NAS 5-3116).
2. "Operational Gridding Methods for the Nimbus Vidicon System," by R. DEBIASE, et al. General Electric Document No. 63SD770, June 6, 1963 (Contract NAS 5-3116).
3. "An Analysis of Errors in the Geographic Referencing of Nimbus Cloud Pictures," by E. M. DARLING, JR. NASA Technical Note D-2137, 1963 (In Press).

HIGH SPEED VACUUM PERFORMANCE OF MINIATURE BALL BEARINGS LUBRICATED WITH COMBINATIONS OF BARIUM, GOLD, AND SILVER FILMS*

THOMAS W. FLATLEY

This report describes the second phase of a program involving metallic film lubrication of miniature ball bearings for vacuum use. Phase I involved bearings with gold-plated balls and raceways and was a retainer study which resulted in the selection of fully machined retainers of "S"-Inconel and silver-plated Circle "C" for use in this phase. Bearings with these retainer types and six ball and race plating combinations of barium, gold, and silver were studied in Phase II. They were run in pairs in small induction motors in a vacuum environment, with nominal test conditions of 10,000 rpm, no external loading, and an oil-free ambient pressure in 10^{-7} torr range.

Testing revealed only one bearing configuration worthy of further study—that involving gold-plated balls, silver-plated raceways, and the fully machined silver-plated Circle "C" retainer. In direct contrast the combination of silver-plated balls and gold-plated raceways gave consistently poor performance. Bearing lifetimes achieved with the other configurations, all involving barium plating, in general fell between these extremes.

The next and final program phase planned is a metallurgical study, investigating the effects of variations in plating thickness and procedures on the performance of the three best bearing configurations found to date.

INTRODUCTION

This report describes the second phase of a three part program involving metallic film lubrication of miniature ball bearings for vacuum use. Program philosophy and testing procedures were described in detail in a previous report covering Phase I of this study.†

Briefly, metallic film lubrication is being studied as a possible answer to the evaporation and radiation resistance problems experienced with conventional bearing lubricants in space. Bowden and Tabor‡ demonstrated that the use of low shear strength metallic films over relatively hard base metals can significantly reduce the coefficient of sliding friction between metals. The usefulness of this type of "lubrication" when applied to ball bearings is the subject under study in this program. This study is being conducted in con-

junction with New Hampshire Ball Bearings, Inc., Peterborough, New Hampshire. The program is set up in three parts:

1. Phase I—retainer study, all balls and races gold plated.
2. Phase II—plating study, involving various combinations of ball and race platings in conjunction with a retainer selected from Phase I.
3. Phase III—metallurgical study, investigation of the effects of plating procedures, thickness, alloying, etc. on the performance of a bearing configuration selected from Phase II.

The test program involves running pairs of $\frac{1}{8}$ in. bore bearings (Figure 1) in small 10,000 rpm induction motors (Figure 2) in a vacuum environment. The motors run until stalled and, since speed is a linear function of load, bearing performance can be observed by speed monitoring. Motor stall torque is on the order of $\frac{1}{4}$ oz-in. A

*Published as *NASA Technical Note D-2304*, June 1964.

†Evans, H. E., and Flatley, T. W., "High Speed Vacuum Performance of Gold Plated Miniature Ball Bearings, with Various Retainer Materials and Configurations," *NASA Technical Note D-2101*, December 1963.

‡F. P. Bowden and D. Tabor, "The Friction and Lubrication of Solids," Oxford, England: Clarendon Press, 1950.

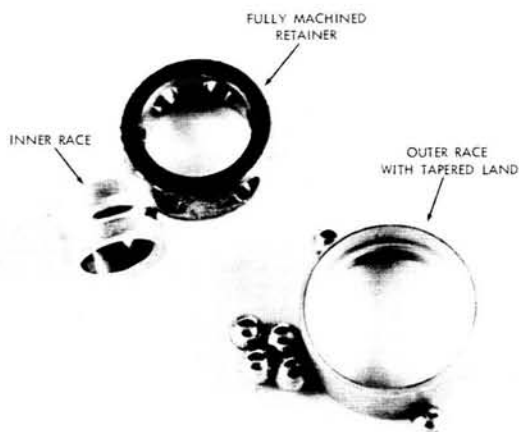


FIGURE 1.—Typical set of bearing components.

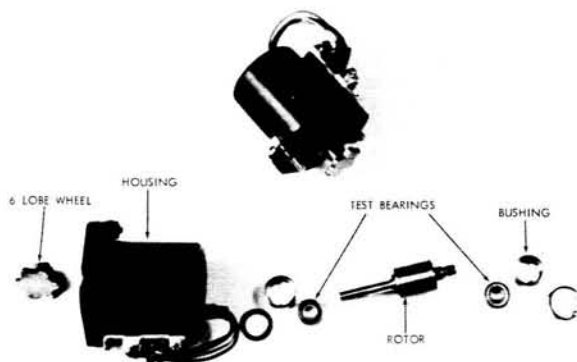


FIGURE 2.—Assembled and disassembled motor for test setup.

special multi-port vacuum system (Figure 3) was designed and built for use in this program. It employs cryogenic roughing and the ion-getter type of high-vacuum pumping in order to eliminate the possibility of bearing lubrication by diffusion pump oil.

Phase I resulted in the choice of fully machined retainers of "S"-Inconel and silver-plated Circle "C" for use in the second phase of the program.

PHASE II PROGRAM

Gold, silver, and barium were the metallic lubricants involved in this part of the study. They were plated on 440C stainless steel balls and races. The gold and silver were electrodeposited to a nominal thickness of 30 microinches and the



FIGURE 3.—Vacuum system, bearing test program.

barium was applied by vacuum deposition, nominally 10 microinches thick.

The six possible permutations of ball and raceway platings and the two retainer types used resulted in the testing of twelve basic bearing configurations. In addition, bearings with gold-plated balls and silver-plated raceways, with the plating done by an alternate plater (designated "LR" in this report) were tested. In Table 1 the test results are presented in fourteen groups, divided according to the basic bearing configuration.

RESULTS

The results are presented in Table 1 and graphically in Figures 4-9. In the latter case tests are grouped according to both ball plating and raceway plating, for comparison purposes. Each test therefore appears twice in these figures.

During testing, mounting difficulties sometimes resulted in the outer race being slightly "out of round." This condition unfortunately did not show up until the bearings were autopsied. Tests which were obviously invalid due to improper mounting have been culled from the data so that the results as shown represent valid tests of the metallic film lubricants involved.

Table 1

Bearing Configurations Tested and Phase II Test Results ("D" and "LR" designate two commercial plating sources).

Ball Plating	Raceway Plating	Retainer Type	Retainer Plating	Test	Running Time (hr)	Remarks
Barium	D Gold	S-Inconel	None	1	7	Gold build-up on balls
				2	13	Same
				3	10	Same
				4	5	Same
	D Silver	Silver-plated Annealed Circle C	D	5	39	Gold and silver build-up on balls
				6	42	Same
		S-Inconel	None	7	502	Heavy wear
				8	1012	Heavy wear and silver build-up on balls
				9	68	Silver flaking
				10	63	Circle C debris and silver build-up on balls
				11	84	Same
				12	103	Same
				13	45	Same
D Gold	Barium	S-Inconel	None	14	48	Gold and S-Inconel build-up on races
				15	66	Same
		Silver-plated Annealed Circle C	D	16	164	Heavy wear
				17	410	Severe ball wear with silver build-up
	D Silver	S-Inconel	None	18	180	Heavy wear, silver build-up on races
				19	203	Same
				20	20	Gold build-up on races
				21	380	Same
				22	79	Same
	Silver-plated Annealed Circle C	D	23	560	Severe wear	
			24	1307	Silver build-up on races	
			25	41	Silver flaking	
			26	757	Silver build-up on races	
LR Gold	LR Silver	S-Inconel	None	27	2	Silver build-up on balls
				28	34	Same
		Silver-plated Annealed Circle C	LR	29	2404	Severe wear
				30	63	Silver build-up on balls
				31	518+	No failure but heavy wear
D Silver	Barium	S-Inconel	None	32	703	Silver build-up on races
				33	111	Same
				34	343	Flaking of silver build-up in retainer pockets
		Silver-plated Annealed Circle C	D	35	161	Silver flaking
				36	266	Silver build-up on races
				37	373	Flaking of silver build-up in retainer pockets
	D Gold	S-Inconel	None	38	325	Silver build-up on races
				39	100	Gold build-up on balls
				40	189	Same
		Silver-plated Annealed Circle C	D	41	1	Silver flaking
				42	40	Gold build-up on balls
				43	16	Same
44	9	Same				
45	18	Same				

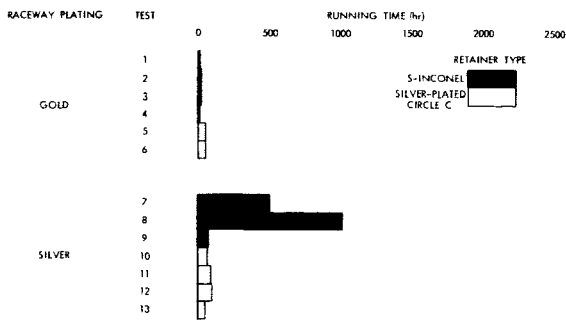


FIGURE 4.—Test results, barium-plated balls.

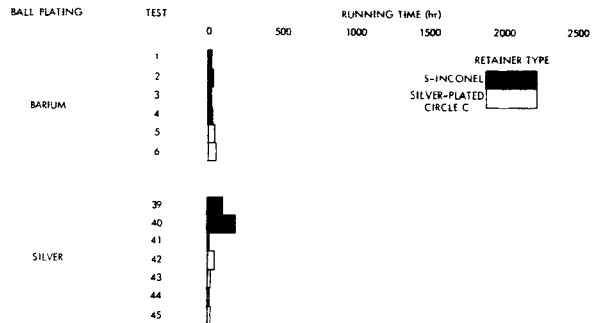


FIGURE 8.—Test results, gold-plated raceways.

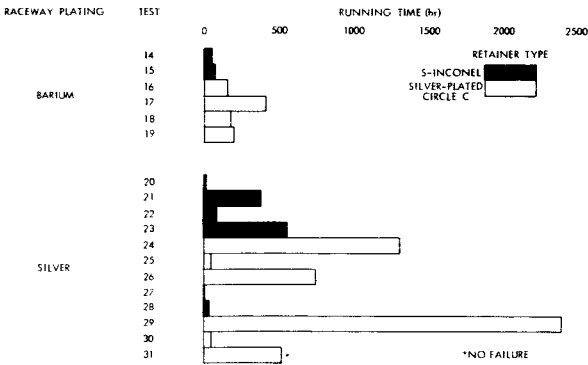


FIGURE 5.—Test results, gold-plated balls.

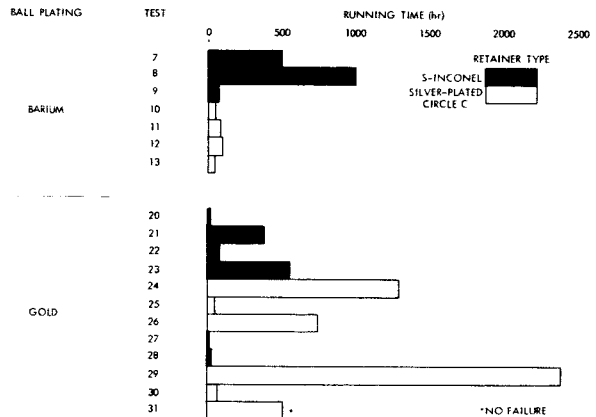


FIGURE 9.—Test results, silver-plated raceways.

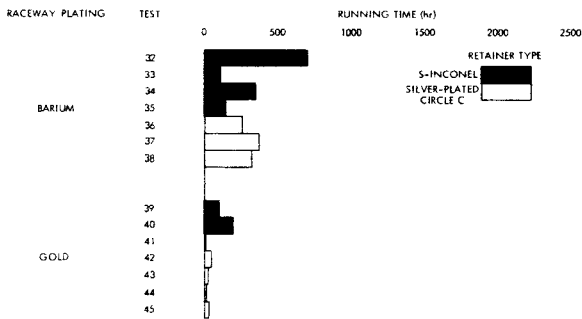


FIGURE 6.—Test results, silver-plated balls.

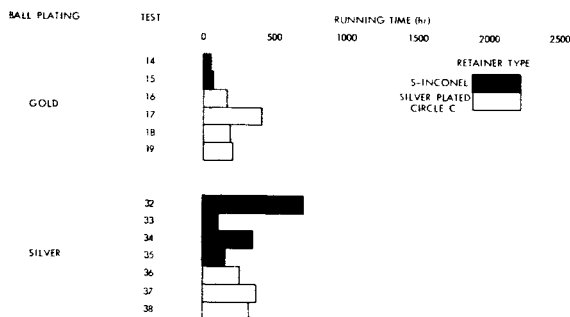


FIGURE 7.—Test results, barium-plated raceways.

Tests 1, 2, 3, and 4 involved the unplated S-Inconel retainer, barium-plated balls and gold-plated races. These bearings had an initial radial play of 0.0002 to 0.0004 in. and rapid failures occurred in each test. None of the bearings had any measurable radial play remaining when they were autopsied. The longest run was only 13 hr and the autopsies revealed a large amount of gold transfer from the races to the balls where it accumulated irregularly, making them rough and eventually causing seizure because of a loss of internal clearance. The retainers in these bearings were relatively unaffected by the testing.

The bearings in tests 5 and 6 were similar except that the silver-plated Circle C retainer was used. Initial radial play was again 0.0002 to 0.0004 in. and again none remained after testing. In these tests the retainer played an important part in extending the bearing life by a factor of 4, compared with those previously discussed (Figure 4). Gold transferred from the raceways to the balls again, but some of it then re-transferred to the

silver-plated retainer ball pockets, thus diminishing the rate of build-up, helping maintain some radial play, and extending the operating time. However, the longest run achieved was still only 42 hr.

Tests 7, 8, and 9 involved bearings with barium-plated balls, silver-plated raceways, and the S-Inconel retainer. Initial radial play was greater (0.0004 to 0.0006 in.) than in previous tests. In all cases where measurement was possible after testing, it was as before found to be zero. Two long runs (500 and 1000 hr) and one early failure were experienced with this configuration. The premature failure in test 9 was caused by silver flaking in a raceway of one of the bearings. The other bearing appeared to be as good as new when autopsied. This failure points out a basic reliability problem with metallic film lubricated bearings. A local defect in the plating, possibly imperceptible under visual inspection, can cause catastrophic failure.

Heavy wear took place during the longer runs on all the bearing balls and races, but retainer damage was surprisingly slight. Apparently the silver remained intact on the raceways for a long period, during which the barium on the balls lubricated their sliding in the retainer ball pockets, so that little metallic transfer took place.

Some silver transfer was evident in the failed bearings however. The balls had areas of build-up, possibly in places where base metal had been exposed. Either transfer of this type or eventual local flaking of the silver probably led to a roughness in the bearing which caused the generation of debris and then failure. The damage to the contact surfaces experienced prevents anything more than speculation in this regard.

Tests 10, 11, 12, and 13 involved the same type of bearings with the same radial play range, but with the silver-plated Circle C retainer. Figure 4 shows the poorer performance of this configuration—the longest run was only 103 hr. In each of these tests, radial play was lost in one bearing but maintained at its initial value in the other, indicating that seizure occurred soon after the failure process began. Failure was caused by an irregular build-up of silver and Circle C on the balls, and extensive retainer damage took place.

Although they were compatible with the previously discussed S-Inconel retainers and did well

when rolling on silver-plated raceways, the barium-plated balls, in sliding in these retainer ball pockets, caused considerable damage, generating both silver and Circle C debris. This debris caused removal of barium from the balls and permitted silver build-up and failure. In the better of the two bearings in each test, this process had not advanced so far.

Tests 14 and 15 involved gold-plated balls, barium-plated races, and S-Inconel retainers. Here again the initial radial play (0.0005 to 0.0006 in.) was maintained in one specimen and lost completely in the other in both tests. Retainer debris was evident in the bearings which failed, but not in those with internal clearance remaining—in one of the latter bearings a band of gold was found in each of the retainer ball pockets.

No visible gold remained on any of the balls in any case—it transferred to both the raceways and the retainers. The bearings ran well as long as some gold remained on the balls, but afterwards sliding contact between the unprotected balls and the S-Inconel retainer generated an abrasive debris which rapidly caused failure as gold and S-Inconel particles built up in the raceways.

Figure 5 shows the improved performance of this type of bearing in *tests 16, 17, 18, and 19*, where the silver-plated Circle C retainer was used. These bearings also had initial radial play levels of 0.0005 to 0.0006 in. and in three of the four tests this was maintained in one bearing and lost in the other. In the other test no measurable radial play remained in either specimen.

Heavy wear took place on the balls and races in all of these tests, but the retainers were relatively unaffected. Considerable amounts of gold transferred to the retainer ball pockets, providing effective sliding lubrication. These bearings ran well as long as enough protective barium remained on the raceways, but eventually, areas of exposed base metal picked up silver or came in contact with other base metal, leading to adhesive wear. Failure was caused by an irregular accumulation of wear debris and silver on the balls and raceways.

The fact that the initial radial play remained in some of the bearings again indicates that failure came shortly after this process began. This characteristic performance of the metallic film lubricated bearings means that catastrophic failure can come without much warning.

Tests 20, 21, 22, and 23 involved bearings with gold-plated balls, silver-plated races S-Inconel retainers, and 0.0003 to 0.0005 in. initial radial play. Figure 5 shows the wide variation in running times which was experienced.

In the case of the two shorter runs (20 and 79 hr) failure was obviously due to an accumulation of gold on the raceways which eliminated internal clearance, and the retainers were unaffected. No visible gold remained on the balls. In the longer runs (380 and 560 hr) raceway buildups also caused seizure, but much of the accumulated debris was due to retainer wear. These were not long, smooth runs. Some large speed fluctuations, indicating periods of high running torque, were observed—particularly in the early stages of the tests.

Apparently a large amount of gold transfer, primarily from the balls to the races, took place early in each of these tests, making the bearings rough. In two cases, this roughness was enough to cause early failure. In tests 21 and 23 the initial high torque period was survived, perhaps because of the generation of abrasive retainer wear particles which ground away some of the build-up. Eventually however retainer debris itself accumulated enough to cause failure.

It is perhaps significant that the two bearings of this group with the smallest initial radial play (0.0003 in.) were involved in the two short runs. The slightly greater internal clearance in the other bearings may account for their survival of the early period of rough operation.

Two very long runs and one short one were experienced when the silver-plated Circle C retainer was used in conjunction with this basic bearing in tests 24, 25, and 26. All of these bearings had 0.0004 or 0.0005 in. initial radial play. The early failure (41 hr) was caused by silver flaking in the raceway, again pointing out the reliability problem with this type of lubrication. The bearings were running smoothly, but failure came suddenly and almost instantaneously. In the longer runs (1307 and 757 hr) a dip in speed again indicated considerable early gold transfer to the raceways, but here transfer to the retainer ball pockets was also taking place so that the raceway build-up was limited.

The natural migration of gold during the operation of this bearing configuration resulted in

very effective metallic film lubrication. The retainer ball pockets were well protected, resulting in very little wear, and raceway build-up was not excessive. Eventual failure seemed to be the result of one or a combination of the following: Minor flaking in the retainer ball pockets after many hours of sliding contact with the balls; some Circle C debris from the edges of the retainer; raceway flaking due to work-hardening of the plating. Here again the actual failure process must have been rapid since one bearing in each case had good radial play remaining when autopsied.

Bearings similar to those of the two previous groups, but with an alternate plating source were also run. In tests 27 and 28 the S-Inconel retainer was used and two short runs resulted (see Figure 5). These failures were caused by a transfer of silver from the races to the balls of one bearing in each case. In the other specimen, the usual gold-to-silver transfer took place. With the "LR" gold and silver plating, the previously observed strong tendency for gold migration to silver-plated surfaces was not present. The reason for this basic difference in the platings produced by the two sources will be a very important item in Phase III testing. Since metallic build-up on the balls is more serious than raceway accumulation, in the cases where silver transferred to the gold, running time was short.

Tests 29, 30, and 31 involved in the same ball and raceway platings and the silver-plated Circle C retainer. An early failure (63 hr) resulted from silver build-up on the balls similar to that described above, but two very long runs were attained with this configuration. The longest run of the entire program (2404 hr) left one pair of bearings severely damaged, so that one can only speculate that gold-to-silver transfer took place, forming a beneficial lubricating film on the contact elements and permitting a long life.

Test 31 involved only one bearing of this type and it was not the cause of the failure after 518 hr. It had radial play levels of 0.0007 in. before test and 0.0004 in. when autopsied. None of the original plating was visible on the balls or raceways and a powdery debris was prevalent, but the retainer ball pockets seem to be well protected and showed little wear.

Tests 32, 33, 34, and 35 involved silver-plated balls, barium-plated races, and S-Inconel retainers.

The initial radial play of each of these bearings was either 0.0004 or 0.0005 in. These bearings ran well as long as barium remained on the raceways, but whenever base metal was exposed areas of silver build-up began. Silver also accumulated in the retainer pockets providing protection, but after a time the deposits there began to flake, and in one case (test 34) a loosened flake caused seizure of a bearing.

This configuration provided fairly good performance (lifetimes of 100 to 700 hr) but does not seem to offer as much promise as other types. Original radial play levels were maintained in some of the bearings, again indicating that the failure process was rapid once base metal was exposed to the silver.

When the silver-plated Circle C retainer was used, in tests 36, 37, and 38, similar running times were experienced (Figure 6). Two of the three failures here were traced to areas of silver build-up on the raceways, and the third was caused by a silver flake from a retainer ball pocket similar to that described above. All visible silver was lost from the balls in each of these tests, much of it forming flaking deposits in the retainer pockets. Even if raceway build-up could be prevented completely with this configuration, loose flakes would eventually cause catastrophic failure.

Tests 39, 40, and 41 involved silver-plated balls, gold-plated races, S-Inconel retainers, and initial radial play levels of 0.00025 to 0.0004 in: One very rapid failure (test 41) occurred because of excessive flaking of the silver plate on the balls, again demonstrating the reliability problem. In the other cases, the bearings ran for 100 and 189 hr before a gold and retainer wear debris build-up on the balls eliminated radial play. Since retainer wear debris was generated the running time was probably lengthened by a grinding away of some of the gold build-up. All gold was removed from both raceways during these tests, and no measurable radial play remained in any when autopsied.

Finally, tests 42, 43, 44, and 45 involved silver-plated balls, gold-plated races, and silver-plated Circle C retainers. The bearings had initial radial play of 0.0002 to 0.0004 in. and their performance was uniformly poor (Figure 6). Here gold built up rapidly on the balls, and to some extent in the retainer ball pockets. No retainer debris was generated so the gold build-up quickly caused

failure. The same strong tendency for gold migration to silver-plated surfaces, which previously resulted in some long runs, quickly led to seizure with this configuration.

CONCLUDING REMARKS

The only Phase II type of bearing worthy of further consideration in the bearing program is that involving silver-plated races, gold-plated balls, and a silver-plated Circle "C" retainer. It will join two Phase I bearing configurations* which provided comparable performance as the basic Phase III bearing types. Since the relative merits of plating sources D and LR have not been settled experimentally, both will again be employed in early Phase III testing. Table 2 shows the bearing configurations still under consideration.

Table 2

Phase III Bearing Configurations.

Retainer	Ball Plating	Race Plating	Plating Source
Silver-plated fully machined Circle C	Gold	Gold	D and LR
None (full complement)	Gold	Gold	D and LR
Silver-plated fully machined Circle C	Gold	Silver	D and LR

Among the aims in Phase III will be reliability improvement by the control of plate adhesion, and the precision measurement and control of plating thickness. Phase III will involve several changes in the test procedure. To broaden the scope of the program, a radial load on the bearings will be introduced, and the test speed will be reduced to 1800 rpm. A magnetic coupling drive will be used to eliminate the drive motor from the vacuum chamber and permit lower ambient pressures. The bearings will run at synchronous speed until they reach a fixed torque level.

*1. Gold-plated balls and races and a silver-plated Circle C fully machined retainer.

2. Gold-plated balls and races, full complement (no retainer).

ANALYSIS OF DISSIMILAR SATELLITE APPENDAGES DURING ERECTION*

RICHARD W. FORSYTHE

By using the Lagrangian formulation of dynamics, equations of motion are developed to describe the deployment process of unlike swinging appendages erecting simultaneously on a spinning satellite. With the Ariel I International Ionosphere Satellite (1962 01) used as the example, an analog computer solution for the variables involved in the erection of the solar paddles and the spin stabilization booms is shown.

INTRODUCTION

Satellite missions often dictate that the payload have appendages extending from the main body. Because of physical limitations it may be necessary to open the booms from a closed position after the nose fairing surrounding the satellite has been released. A common erection procedure is: (1) The last stage with the satellite mounted on it is spin-stabilized, (2) the last stage is fired, (3) the appendages are released after last-stage burnout, and (4) separation of the satellite from the last stage occurs. The energy that erects the appendages is derived from the spinning of the entire system.

The engineer who designs the appendages is necessarily concerned with formulating and solving the equations of motion describing the system so that he may properly design the appendages to withstand the loads during erection. If the appendages are all *similar*, the problem is relatively simple. However, if the erection involves the simultaneous release and swinging erection of *dissimilar* appendages, then the formulation and solution of the equations of motion becomes more difficult. This report presents a Lagrangian formulation and analog computer solution of these equations and shows how they were used to obtain the forces on the solar paddles and inertia booms of the Ariel I International Ionosphere Satellite during their simultaneous erection.

LIST OF SYMBOLS

The symbols used in this report are defined as follows:

- a, f = Distances from the system spin axis to the pivot points of the two types of appendages
- b, d = Total length of the uniformly distributed arms of the two types of appendages
- c = Distance from the pivot point to the center of concentrated mass of appendages 2 and 4
- e = Total length of the uniformly distributed mass on the end of appendages 1 and 3
- I_b = Total moment of inertia about the spin axis of the satellite, spent last stage, and attachments minus the inertia of the appendages
- I_0 = Total moment of inertia about the spin axis of the satellite, spent last stage, and attachments before the appendages erect
- I_2 = Total moment of inertia about the spin axis of the satellite, spent last stage, and attachments after the appendages erect
- I_r = Total moment of inertia about the spin axis of the satellite, spent last stage, and attachments at any time, $t = \tau$
- m_b = Sum of the masses of all arms of length b
- m_d = Sum of the masses of all arms of length d
- m_e = Sum of the masses of all paddles of length e
- m_w = Sum of the concentrated masses on the end of appendages 2 and 4

*Published as NASA Technical Memorandum X-938, January 1964.

T = Total energy of the satellite, spent last stage, and attachments

t = Time

dx = Incremental distance along appendages 2 and 4

dy = Incremental distance along appendages 1 and 3

θ = Angle of appendages 2 and 4 from the folded position where $\theta = 0^\circ$

$\dot{\theta}$ = Component of the angular velocity of appendages 2 and 4 in a direction parallel to the system spin axis

$\ddot{\theta}$ = Component of the angular acceleration of appendages 2 and 4 in a direction parallel to the system spin axis

λ = Angular velocity of the entire system about the intended spin axis

ρ = Mass per unit length of any uniformly distributed part

ϕ = Angle of appendages 1 and 3 from the folded position where $\phi = 0^\circ$

$\dot{\phi}$ = Component of the angular velocity of appendages 1 and 3 in a direction parallel to the system spin axis

$\ddot{\phi}$ = Component of the angular acceleration of appendages 1 and 3 in a direction parallel to the system spin axis.

DYNAMIC ANALYSIS

Figure 1 shows the system to be discussed.

To properly design the appendages, the engineer will need to know each boom's impact energy at erection.* Therefore, this report will show a solution to obtain $\dot{\theta}$ and $\dot{\phi}$ as functions of time. Other parameters of interest, such as λ , $\ddot{\theta}$, and $\ddot{\phi}$, are obtained as byproducts.

SOLUTION

The total kinetic energy of the system T is given by the following expression:

$$\begin{aligned} \Sigma T = & \frac{1}{2} m_w c^2 \dot{\theta}^2 + \frac{1}{2} m_w (c \sin \theta + a)^2 \dot{\lambda}^2 + \frac{1}{2} \int_0^b \rho x^2 \dot{\theta}^2 dx \\ & + \frac{1}{2} \int_0^b \rho (x \sin \theta + a)^2 \dot{\lambda}^2 dx + \frac{1}{2} \int_d^{d+e} \rho y^2 \dot{\phi}^2 dy + \frac{1}{2} \int_d^{d+e} \rho (y \sin \phi + f)^2 \dot{\lambda}^2 dy \\ & + \frac{1}{2} \int_0^d \rho y^2 \dot{\phi}^2 dy + \frac{1}{2} \int_0^d \rho (y \sin \phi + f)^2 \dot{\lambda}^2 dy + \frac{1}{2} I_b \dot{\lambda}^2 \quad (1) \end{aligned}$$

in which I_b is the total moment of inertia of the entire system around the spin axis less the inertia of the four booms.

Carrying out the indicated integration, Equation 1 reduces to

$$\begin{aligned} T = & \frac{1}{2} m_w c^2 \dot{\theta}^2 + \frac{1}{2} m_w (c \sin \theta + a)^2 \dot{\lambda}^2 + \frac{1}{6} m_b b^2 \dot{\theta}^2 \\ & + \frac{1}{2} m_b \left(\frac{1}{3} b^2 \sin^2 \theta + ab \sin \theta + a^2 \right) \dot{\lambda}^2 + \frac{1}{6} m_e \dot{\phi}^2 (3d^2 + 3ed + e^2) \\ & + \frac{1}{2} m_e \left[\left(\frac{3d^2 + 3ed + e^2}{3} \right) \sin^2 \phi + (2fd + fe) \sin \phi + f^2 \right] \dot{\lambda}^2 \\ & + \frac{1}{6} m_d d^2 \dot{\phi}^2 + \frac{1}{2} m_d \left(\frac{d^2}{3} \sin^2 \phi + fd \sin \phi + f^2 \right) \dot{\lambda}^2 + \frac{1}{2} I_b \dot{\lambda}^2 \quad (2) \end{aligned}$$

*Forsythe, R. W., "A Method for Simulating Zero Gravity Erection of Satellite Appendages," NASA TN D-1852, 1963 (In press).

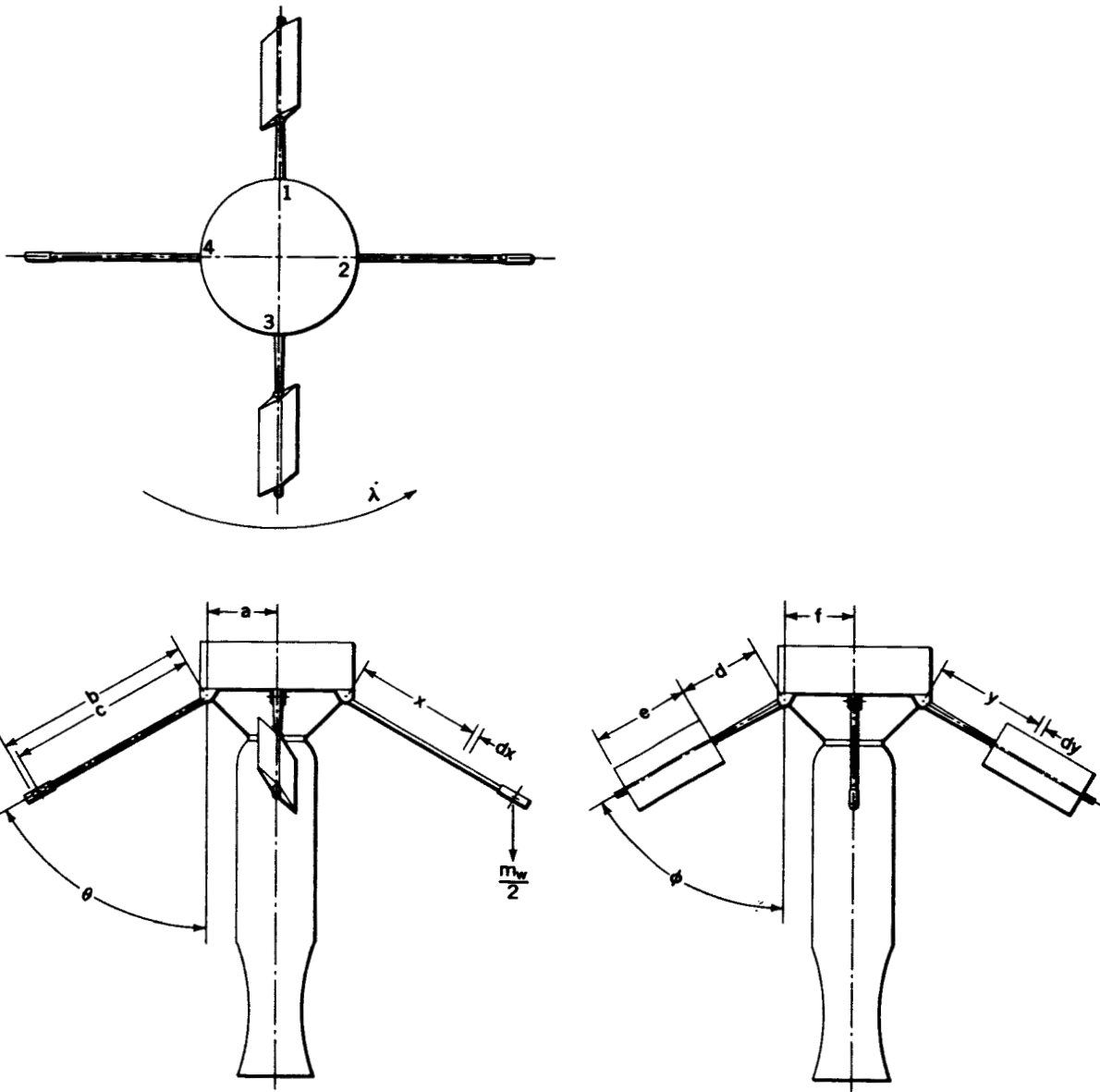


FIGURE 1.—Spacecraft and last stage.

where

m_b = sum of the individual masses of all arms of length b ,

m_d = sum of the individual masses of all arms of length d ,

m_e = sum of the individual masses of all paddles of length e .

It is seen that T is a function of the q_n generalized coordinates θ , ϕ , and λ and their time derivatives. The equations of motion may then

be written in the Lagrangian form as follows:

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_n} \right) - \frac{\partial T}{\partial q_n} = Q_n \quad (n = 1, 2, 3) .$$

In general the erection of appendages occurs during a coast phase of the launch and is completed in such a short time that all externally applied forces are negligible. Thus, since the internal constraints do not contribute to the

external work done, $Q_n = 0$; and for our system the equations of motion are

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_n} \right) - \frac{\partial T}{\partial q_n} = 0. \quad (3)$$

Performing the indicated differentiation of the kinetic energy equation results in the following expressions:

$$\frac{\partial T}{\partial \theta} = m_w (c \sin \theta + a) (c \cos \theta) \dot{\lambda}^2 + m_b \dot{\lambda}^2 \left(\frac{1}{3} b^2 \sin \theta \cos \theta + \frac{1}{2} ab \cos \theta \right), \quad (4)$$

$$\begin{aligned} \frac{\partial T}{\partial \phi} = & \frac{1}{2} m_e \dot{\lambda}^2 \left[\frac{2(3d^2 + 3ed + e^2)}{3} \sin \phi \cos \phi + (2fd + fe) \cos \phi \right] \\ & + \frac{1}{2} m_d \dot{\lambda}^2 \left[\frac{2}{3} d^2 \sin \phi \cos \phi + fd \cos \phi \right], \end{aligned} \quad (5)$$

$$\frac{\partial T}{\partial \lambda} = 0; \quad (6)$$

and

$$\frac{\partial T}{\partial \dot{\theta}} = m_w c^2 \dot{\theta} + \frac{1}{3} m_b b^2 \dot{\theta}, \quad (7)$$

$$\frac{\partial T}{\partial \dot{\phi}} = \frac{1}{3} (3d^2 + 3ed + e^2) m_e \dot{\phi} + \frac{1}{3} m_d d^2 \dot{\phi}, \quad (8)$$

$$\begin{aligned} \frac{\partial T}{\partial \dot{\lambda}} = & (c \sin \theta + a)^2 m_w \dot{\lambda} + \left(\frac{1}{3} b^2 \sin^2 \theta + ab \sin \theta + a^2 \right) m_b \dot{\lambda} \\ & + \left[\left(\frac{3d^2 + 3ed + e^2}{3} \right) \sin^2 \phi + (2fd + fe) \sin \phi + f^2 \right] m_e \dot{\lambda} \\ & + \left(\frac{1}{3} d^2 \sin^2 \phi + fd \sin \phi + f^2 \right) m_d \dot{\lambda} + I_b \dot{\lambda}; \end{aligned} \quad (9)$$

and

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{\theta}} \right) = m_w c^2 \ddot{\theta} + \frac{1}{3} m_b b^2 \ddot{\theta}, \quad (10)$$

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{\phi}} \right) = \frac{1}{3} (3d^2 + 3ed + e^2) m_e \ddot{\phi} + \frac{1}{3} m_d d^2 \ddot{\phi}. \quad (11)$$

Substituting the expressions into the Lagrangian form as indicated by Equation 3,

By integrating Equation 12,

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{\lambda}} \right) - 0 = 0. \quad (12) \qquad \frac{\partial T}{\partial \lambda} = K. \quad (13)$$

where K is the constant of integration. Thus,

$$\begin{aligned} & (c \sin \theta + a)^2 m_w \dot{\lambda} + \left(\frac{1}{3} b^2 \sin^2 \theta + ab \sin \theta + a^2 \right) m_b \dot{\lambda} \\ & + \left[\left(\frac{3d^2 + 3ed + e^2}{3} \right) \sin^2 \phi + (2fd + fe) \sin \phi + f^2 \right] m_e \dot{\lambda} \\ & + \left(\frac{1}{3} d^2 \sin^2 \phi + fd \sin \phi + f^2 \right) m_d \dot{\lambda} + I_b \dot{\lambda} = K . \end{aligned} \quad (14)$$

Examination of Equation 14 reveals it to be simply the general conservation of momentum equation describing the system.

Evaluation of K is easily accomplished by evaluating the left-hand side of Equation 14 at

any time $t = \tau$ where the relation between θ , ϕ , and $\dot{\lambda}$ is known, such as the initial point or the end point.

Therefore, the three developed equations necessary to completely describe the internal force

system are

$$m_w c^2 \ddot{\theta} + \frac{1}{3} m_b b^2 \ddot{\theta} - (c \sin \theta + a) (c \cos \theta) m_w \dot{\lambda}^2 - \left(\frac{1}{3} b^2 \sin \theta \cos \theta + \frac{1}{2} ab \cos \theta \right) m_b \dot{\lambda}^2 = 0 , \quad (15)$$

$$\begin{aligned} & \frac{1}{3} (3d^2 + 3ed + e^2) m_e \ddot{\phi} + \frac{1}{3} m_d d^2 \ddot{\phi} \\ & - \frac{1}{2} \left[\frac{2(3d^2 + 3ed + e^2)}{3} \sin \phi \cos \phi + (2fd + fe) \cos \phi \right] m_e \dot{\lambda}^2 \\ & - \frac{1}{2} \left[\frac{2}{3} d^2 \sin \phi \cos \phi + fd \cos \phi \right] m_d \dot{\lambda}^2 = 0 , \end{aligned} \quad (16)$$

$$\begin{aligned} & (c \sin \theta + a)^2 m_w \dot{\lambda} + \left(\frac{1}{3} b^2 \sin^2 \theta + ab \sin \theta + a^2 \right) m_b \dot{\lambda} \\ & + \left[\left(\frac{3d^2 + 3ed + e^2}{3} \right) \sin^2 \phi + (2fd + fe) \sin \phi + f^2 \right] m_e \dot{\lambda} \\ & + \left(\frac{1}{3} d^2 \sin^2 \phi + fd \sin \phi + f^2 \right) m_d \dot{\lambda} + I_b \dot{\lambda} = I_\tau \dot{\lambda}_\tau . \end{aligned} \quad (17)$$

Two of the equations are nonhomogeneous, second degree, second order, nonlinear differential equations; and the third is a nonhomogeneous, second degree, first order, nonlinear differential

equation. The solution for $\ddot{\theta}$ and $\ddot{\phi}$ is not "straightforward" by hand calculation. The computer solution of the equations is simple, however, and provides an exact description of the

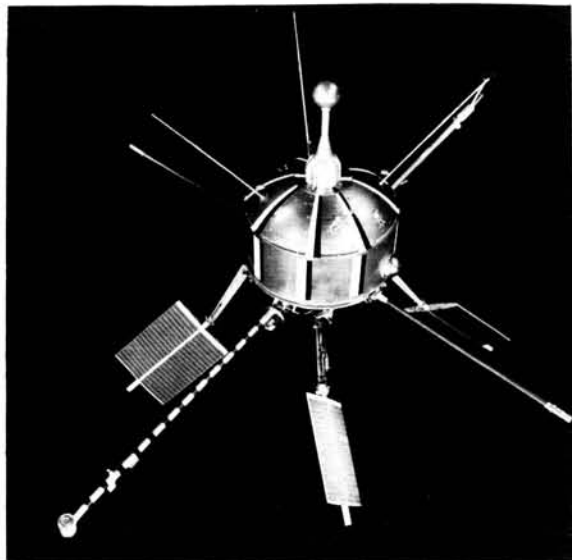


FIGURE 2.—Ariel I International Ionosphere Satellite.

parameters involved. The next section illustrates the solution of the equations of motion applied to the Ariel I satellite as solved on an analog computer.

APPLICATION OF DERIVED EQUATIONS TO THE ARIEL I SATELLITE

The Ariel I International Ionosphere Satellite is shown in Figure 2. The payload has two experimental booms, two mass booms used for spin stabilization, and four solar paddles. In actual flight prior to any erection the appendages were folded down beside the last-stage bottle. The planned erection sequence was as follows: (1) The system was to be spin-stabilized at 161 rpm prior to last-stage ignition; (2) 15 minutes after last-stage ignition, a stretch yo-yo de-spin device would unwind to reduce the spin rate of the system to 76 rpm; (3) 60 seconds after yo-yo release, the two experimental booms were to be released and, under the constraint of a governor, erect in approximately 3.0 seconds; (4) 60 seconds after the previous action, with the system at approximately 52 rpm, the four solar paddles and the two mass booms would be released simultaneously and allowed to swing freely to their erected positions; and (5) 60 seconds later, with the system at approximately 35 rpm, the separation of the spacecraft from the last-stage bottle would occur and the satellite would be in orbit as seen in Figure 2. This section presents the solution of

the equations of motion describing the simultaneous erection of the solar paddles and the mass booms.

A sketch of the system to be analyzed is shown in Figure 3. On the actual spacecraft solar paddles 1 and 3 were slightly different from paddles 2 and 4; however, to simplify the problem the average of the two paddle parameters will be used.

The three equations of motion describing the simultaneous erection action of dissimilar appendages are, as given before, Equations 15, 16, and 17.

It is desirable to solve the equations for θ and ϕ so that the design engineer may obtain values for the kinetic energies of the appendages at impact.

The following quantities describe the Ariel I satellite and the last stage after burnout, just prior to the erection of the solar paddles and mass booms:

$I_0 = 18,599 \text{ lb-in.}^2$	$0^\circ \leq \theta \leq 90.00^\circ$
$I_2 = 27,609 \text{ lb-in.}^2$	$0^\circ \leq \phi \leq 36.76^\circ$
$I_b = 16,483 \text{ lb-in.}^2$	$a = 10.125 \text{ in.}$
$m_w/2 = 0.775 \text{ lb}$	$b = 34.000 \text{ in.}$
$m_b/2 = 0.750 \text{ lb}$	$c = 33.500 \text{ in.}$
$m_c/4 = 2.200 \text{ lb}$	$d = 14.600 \text{ in.}$
$m_d/4 = 1.890 \text{ lb}$	$e = 16.500 \text{ in.}$
$\lambda_0 = 52 \text{ rpm}$	$f = 10.500 \text{ in.}$

Before Equations 16 and 17 can be used, an adjustment to ϕ must be made. On the Ariel I satellite the solar paddles did not swing out to their erected positions in the plane made by the spin axis and the paddles' initial positions. The main hinge on the paddles constrained the travel so that a top view of the actual path was as seen in Figure 4, where ϕ is approximately the angle the paddles would move through if they swung straight out and ϕ' the total angle the paddles actually swung through.

A comparison of the actual c.g. path length along ϕ' to the straight path length to the point where $\phi = 36.76$ degrees gave a ϕ' of 40.36 degrees. When the length of time involved for erection is considered, the limits for the angle travel should therefore be 0 to 40.36 degrees. However, the moment of inertia of the entire system and thereby the system spin rate is independent of ϕ' and depends on ϕ . Therefore as an approximation describing the action, the limits 0 to 40.36 degrees were used; however, to satisfy inertia conditions, the substitution $\phi = (36.76/40.36) (\phi')$, or $\phi = k\phi'$,

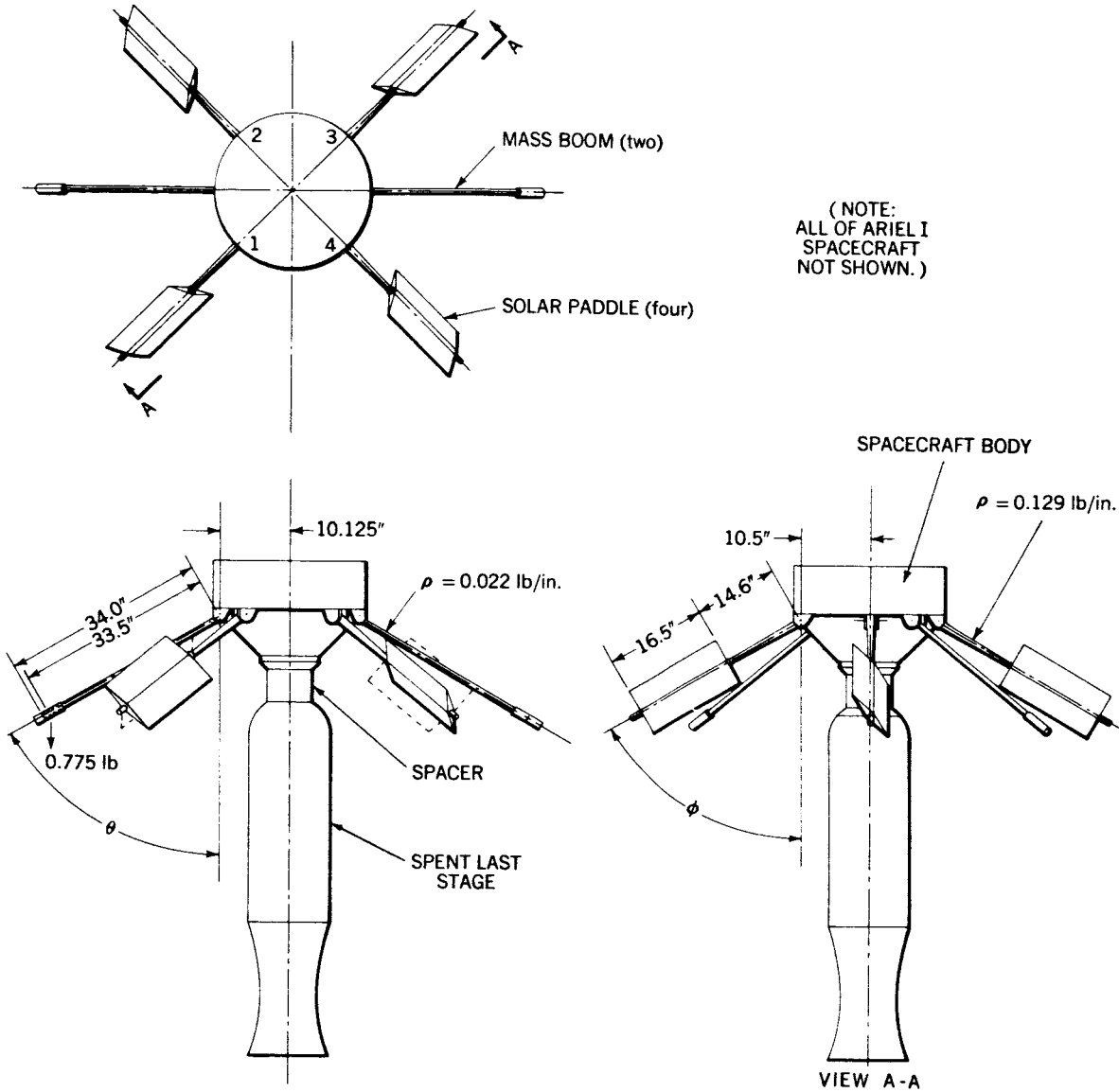


FIGURE 3.—Ariel I spacecraft and last stage.

was made. The terms $\dot{\phi}'$ and $\ddot{\phi}'$, when obtained from the computer solution, will be the angular velocity and acceleration of the paddles along the

0 to 40.36 degree path. By making the substitution and going through the derivation again, Equations 16 and 17 are changed to

$$\frac{1}{3}(3d^2 + 3ed + e^2) m_e \ddot{\phi} + \frac{1}{3} m_d d^2 \ddot{\phi}$$

$$- \frac{1}{2} k \left[\frac{2}{3} (3d^2 + 3ed + e^2) \sin k\phi' \cos k\phi' + (2fd + fe) \cos k\phi' \right] m_e \dot{\lambda}^2$$

$$- \frac{1}{2} k \left[\frac{2}{3} d^2 \sin k\phi' \cos k\phi' + fd \cos k\phi' \right] m_d \dot{\lambda}^2 = 0 \quad (18)$$

and

$$\begin{aligned}
 & (c \sin \theta + a)^2 m_w \dot{\lambda} + \left(\frac{1}{3} b^2 \sin^2 \theta + ab \sin \theta + a^2 \right) m_b \dot{\lambda} \\
 & + \left[\left(\frac{3d^2 + 3ed + e^2}{3} \right) \sin^2 k\phi' + (2fd + fe) \sin k\phi' + f^2 \right] m_e \dot{\lambda} \\
 & + \left(\frac{1}{3} d^2 \sin^2 k\phi' + fd \sin k\phi' + f^2 \right) m_d \dot{\lambda} + I_b \dot{\lambda} = I_r \dot{\lambda}_r. \quad (19)
 \end{aligned}$$

By applying the constants, Equation 15 reduces to

$$\ddot{\theta} - f_2(\theta) \dot{\lambda}^2 = 0, \quad (20)$$

where

$$f_2(\theta) = \sin \theta \cos \theta + 0.338 \cos \theta;$$

Equation 18 reduces to

$$\ddot{\phi} - f_2(k\phi') \dot{\lambda}^2 = 0, \quad (21)$$

where

$$f_2(k\phi') = k(\sin k\phi' \cos k\phi' + 0.505 \cos k\phi');$$

and Equation 19 reduces to

$$f_1(\theta) \dot{\lambda} + f_1(k\phi') \dot{\lambda} + 46.498 \dot{\lambda} = 252.948, \quad (22)$$

where

$$f_1(\theta) = 5.793 \sin^2 \theta + 3.919 \sin \theta$$

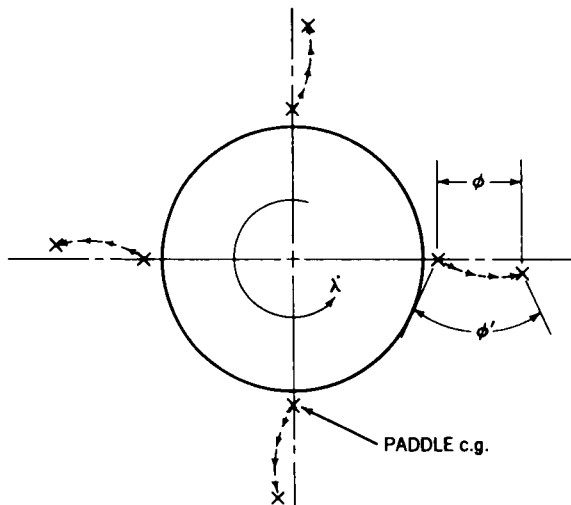


FIGURE 4.—Solar paddle paths.

and

$$f_1(k\phi') = 13.331 \sin^2 k\phi' + 13.455 \sin k\phi'.$$

Two analog computers, each with 20 amplifiers and approximately a ± 10 volt limit, were used to solve the three equations. The variables were estimated and scaled as follows:

Variable	Scaled variable
θ	$\left[\frac{\theta}{9} \right]$
$\dot{\theta}$	$[\dot{\theta}]$
$\ddot{\theta}$	$\left[\frac{\ddot{\theta}}{2.5} \right]$
ϕ'	$\left[\frac{\phi'}{4.036} \right]$
$\dot{\phi}'$	$[\dot{\phi}']$
$\ddot{\phi}'$	$\left[\frac{\ddot{\phi}'}{2.5} \right]$
λ	$[2\dot{\lambda}]$
t	$[20t]$
$f_1(\theta)$	$[f_1(\theta)]$
$f_1(k\phi')$	$\left[\frac{f_1(k\phi')}{1.5} \right]$
$f_2(\theta)$	$[10 f_2(\theta)]$
$f_2(k\phi')$	$[10 f_2(k\phi')]$

The scaled equations take the form:

$$\left[\frac{\ddot{\theta}}{2.5} \right] = \frac{1}{100} \left[10 f_2(\theta) \right] \left[2\dot{\lambda} \right]^2,$$

$$\left[\frac{\ddot{\phi}}{2.5} \right] = \frac{1}{100} \left[10 f_2(k\phi') \right] \left[2\dot{\lambda} \right]^2,$$

$$\left[2\dot{\lambda} \right] = 10 \left(\frac{5.059}{0.1 \left[f_1(\theta) \right] + 0.15 \left[\frac{f_1(k\phi')}{1.5} \right] + 4.650} \right).$$

The computer program for the three equations is shown in Figure 5.

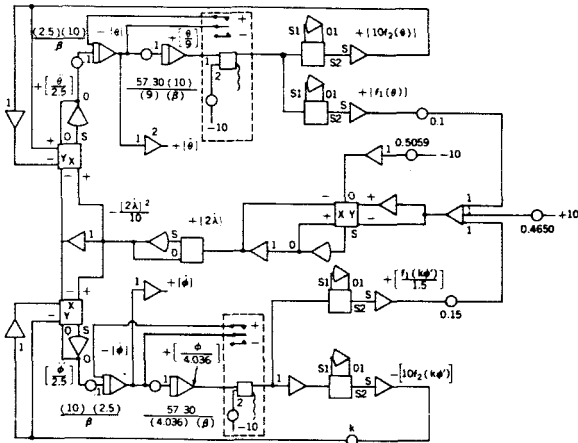


FIGURE 5.—Analog computer diagram.

With the program shown and using a time scale factor of $\beta=100$, the erection variables were plotted on an X-Y plotter. The simultaneous solution for the unknown quantities versus time and also versus the appendage angle from the folded position is shown in Figures 6, 7, and 8.

Inspection of the curves reveals the following:

1. At full mass boom erection ($\theta=90$ degrees), $\dot{\theta}=5.35$ rad/sec; and at full paddle erection ($\phi'=40.36$ degrees), $\dot{\phi}'=4.61$ rad/sec. If the kinetic energy of all appendages is calculated considering these impact angular velocities and the inertia around the hinge points of the appendages

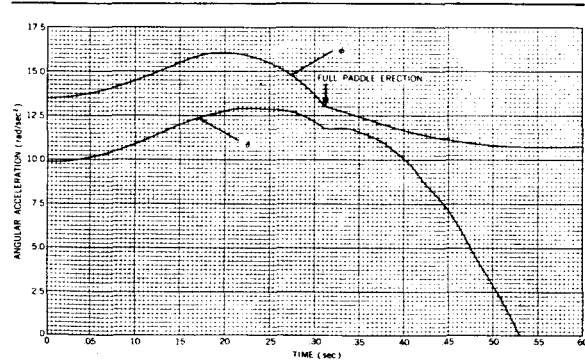


FIGURE 6.—Solar paddle and mass boom angular accelerations versus time.

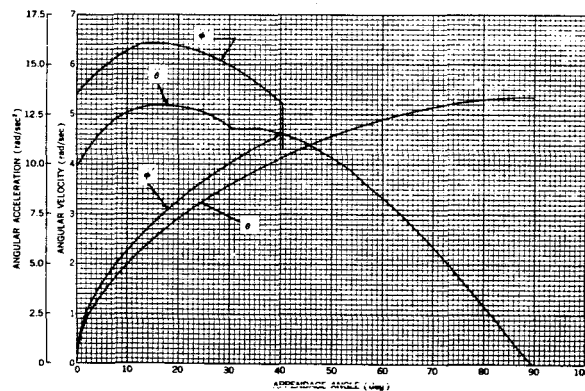


FIGURE 7.—Solar paddle and mass boom angular accelerations and velocities versus appendage angle.

and if this energy is compared with the energy difference given by $(1/2 I_0 \dot{\lambda}_0^2 - 1/2 I_2 \dot{\lambda}_2^2)$ where all of these quantities are accurately known, it will be

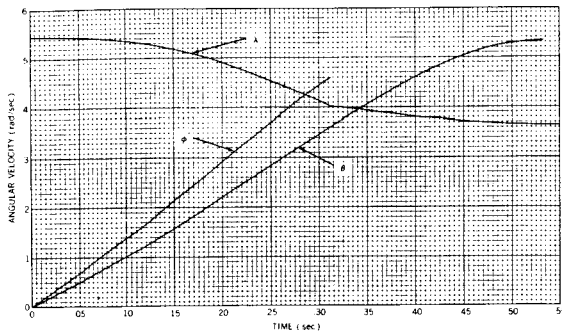


FIGURE 8.—Solar paddle, mass boom, and satellite spin axis angular velocities versus time.

seen that agreement is good to within 0.03 percent.

2. Maximum $\ddot{\theta} = 12.90$ rad/sec², and maximum $\ddot{\phi}' = 16.05$ rad/sec².

3. With an initial spin rate λ_0 of 5.44 rad/sec (52 rpm), the de-spin of the entire system to the

proper final spin rate of 3.66 rad/sec (35 rpm) was plotted out.

By using the above values for $\dot{\theta}_{\max}$ and $\dot{\phi}_{\max}$, the mass booms and solar paddles were designed to withstand erection on the Ariel I satellite.

CONCLUSIONS

The conclusions drawn from the calculations may be summarized as follows:

1. By the Lagrangian formulation, three equations that describe the simultaneous erection process of unlike satellite appendages may be developed.

2. The simultaneous solution of the equations on a computer will provide a complete path history of the variables θ , $\dot{\theta}$, $\ddot{\theta}$; ϕ , $\dot{\phi}$, $\ddot{\phi}$; and λ .

3. By using the magnitudes of the variables, the appendages may be designed for the erection action.

A METHOD FOR SIMULATING ZERO-GRAVITY ERECTION OF SATELLITE APPENDAGES*

RICHARD W. FORSYTHE

Equations are developed which show an analytical approach to simulating the 0-g impact loading of satellite swinging booms. The equations are obtained by simple manipulation of the equations of motion and the kinetic and potential energies involved and are verified by data from the actual spin tests of the International Ionosphere Satellite Ariel I (1962 01).

LIST OF SYMBOLS

- | | | | |
|-------|--|------------------|---|
| a | Distance from the system spin axis to the boom pivot points. | V_θ | Total potential energy of the test system for any value of θ . |
| E | Energy contained in all appendages due only to the component of boom angular velocity parallel to the satellite spin axis, that is, $\dot{\theta}$. | W | Total weight of each uniformly distributed boom arm. |
| I | Total moment of inertia about the spin axis of the satellite, spent last stage, and attachments. | w | Weight on the end of each boom. |
| I_b | Total moment of inertia about the spin axis of the satellite, spent last stage, and attachments minus the inertia of the booms. | δ | Total deflection at impact of the end of each appendage. |
| J | Total moment of inertia about the spin axis of the test model and attachments. | θ | Boom angle from the folded position where $\theta=0^\circ$. |
| J_b | Total moment of inertia about the spin axis of the test model and attachments minus the inertia of the booms. | $\dot{\theta}$ | First derivative of θ , or the component of the angular velocity of all appendages in a direction parallel to the spin axis. |
| L | Total torque on the booms during erection. | $\dot{\lambda}$ | Angular velocity of entire system about the intended spin axis. |
| l | Arm length, or the distance from the boom pivot point to the boom weight m. | $\ddot{\lambda}$ | First derivative of $\dot{\lambda}$, or the angular acceleration of the entire system about its spin axis. |
| M | Total mass of each uniformly distributed boom arm. | ρ | Weight per unit length of the uniformly distributed boom arm; obtained by dividing the density of the boom arm by the cross-sectional area. |
| m | Mass on the end of each boom. | | |
| n | Number of booms on the satellite. | | |
| r | Distance from the boom pivot point to the center of the boom weight m. | | |
| T | Total energy of satellite, spent last stage, and attachments. | | |
| V | Total potential energy of the spinning system in a 1-g conservative field. | | |

General Notation

Subscript 0: Initial system condition prior to boom erection.

Subscript 2: Final system condition after boom erection.

Subscript g: Refers to an expression written for the 1-g force field condition.

Subscript z: Refers to an expression written for the 0-g force field condition.

*Published as NASA Technical Memorandum X-937, January 1964.

INTRODUCTION

Satellite missions often dictate that the payload have appendages, or booms, extending out from the main structure. Because of physical limitations, it may be necessary to open the booms from a closed position after nose-fairing release. The booms usually will be erected either by a telescoping mechanism or, if space is limited in the main satellite body, by a swinging boom technique in which the folded booms swing out when released. Prior to the last-stage ignition normally the system will be spin stabilized and energy for the boom erection will result from the spin.

The main problem involved in developing an erectable swinging boom is to devise testing procedures which simulate actual erection conditions in a 0-g field.

This report gives a method used during development of the International Ionosphere Satellite Ariel I (1962 01) for simulating boom erection conditions in a 0-g field.

ANALYSIS

Figure 1 is a sketch of the spin system and coordinate system which is discussed. The parameters are:

- a Distance from the system spin axis to the boom pivot points.
- x Distance from the boom pivot point to a point on the boom arm.
- m Mass on the end of each boom.
- l Arm length, or the distance from the boom pivot point to the boom weight m.
- r Distance from the boom point to the center of the boom weight m.
- θ Boom angle from folded position where $\theta=0^\circ$.
- $\dot{\lambda}$ Angular velocity of the entire system about the symmetrical axis.

It is assumed that there are two booms 180 degrees apart. The free swinging booms will normally be designed to take the following two types of loads:

1. The total torque L on the boom system during raising of the boom.
2. The energy E_z that must be absorbed to stop the booms in the erected position.

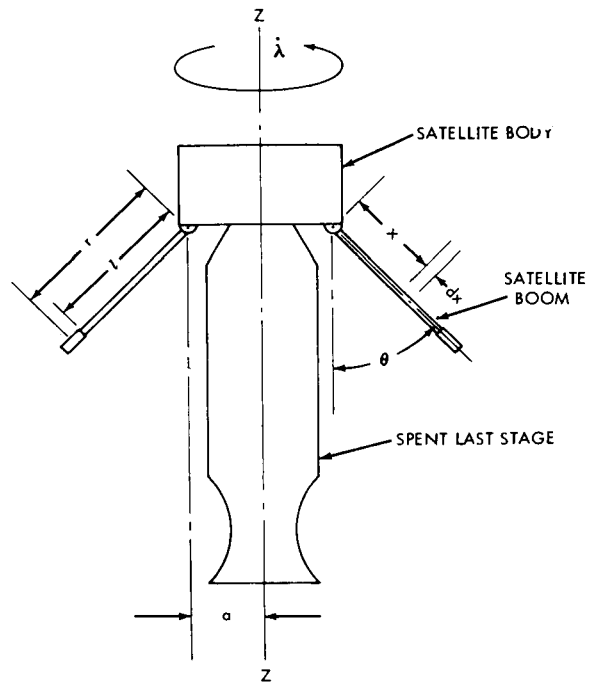


FIGURE 1.—Satellite and last stage.

Normally at the time of a satellite boom erection the system is spinning essentially in a 0-g field and moving at a constant velocity; thus, the system can be examined analytically as if it were stationary and spinning. It is relatively easy to show that the total torque on the two boom hinges during swing-up is given by the expression

$$L = nmr\ddot{\lambda}(r \sin \theta + a) + 2nmr^2\dot{\lambda}\dot{\theta} \cos \theta + \frac{1}{3}nM\ddot{\lambda}\left(l \sin \theta + \frac{3}{2}a\right) + \frac{2}{3}nMl^2\dot{\lambda}\dot{\theta} \cos \theta, \quad (1)$$

where

n is the number of appendages,

M is the total mass of each boom arm, and

$\ddot{\lambda}$ is the system's angular acceleration about its spin axis

The total kinetic energy of the whole satellite system is given by the expression

$$T_z = \frac{1}{2} nmr^2\dot{\theta}^2 + \frac{1}{2} n\rho\dot{\theta}^2 \int_0^l x^2 dx + \frac{1}{2} nm\dot{\lambda}^2 (r \sin \theta + a)^2 + \frac{1}{2} n\rho\dot{\lambda}^2 \int_0^l (x \sin \theta + a)^2 dx + \frac{1}{2} I_b \dot{\lambda}^2 \quad (2)$$

where ρ is the uniformly distributed weight per unit length (lb/in) of the boom arm and I_b is the total moment of inertia of the entire structure less that of the booms. Equation 2 reduces to

$$T_z = \frac{1}{2} nmr^2\dot{\theta}^2 + \frac{1}{6} nMl^2\dot{\theta}^2 + \frac{1}{2} nm\dot{\lambda}^2 (r \sin \theta + a)^2 + \frac{1}{2} nM\dot{\lambda}^2 \left(\frac{1}{3} l^2 \sin^2 \theta + al \sin \theta + a^2 \right) + \frac{1}{2} I_b \dot{\lambda}^2 \quad (3)$$

Both momentum and energy are conserved in the system. In the Lagrange formulation, the partial derivative of Equation 3 with respect to $\dot{\lambda}$ gives the momentum equation for the entire system, since $\partial T/\partial \lambda = 0$;

$$\frac{\partial T}{\partial \dot{\lambda}} = nm\dot{\lambda} (r \sin \theta + a)^2 + nM\dot{\lambda} \left(\frac{1}{3} l^2 \sin^2 \theta + al \sin \theta + a^2 \right) + I_b \dot{\lambda} \quad (4)$$

Since $\theta = 0^\circ$ and $\lambda = \lambda_0$ at time $t = 0$, the momentum equation becomes

$$na^2\dot{\lambda}_0(m + M) + I_b\dot{\lambda}_0 = \text{Constant} \quad (5)$$

From the conservation of momentum it follows that

$$\frac{\dot{\lambda}}{\dot{\lambda}_0} = \frac{na^2(m + M) + I_b}{nm(r \sin \theta + a)^2 + nM \left(\frac{1}{3} l^2 \sin^2 \theta + al \sin \theta + a^2 \right) + I_b} \quad (6)$$

When Equation 6 is differentiated with respect to θ and multiplied by $\dot{\theta}$, an expression for $\dot{\lambda}$ is obtained. With the expression for $\dot{\lambda}$ and Equations 1 and 3, it is easy to plot the energy contained in the booms due to $\dot{\theta}$ and the torque on the hinges. If this procedure is followed, it will become obvious that *the energy due to $\dot{\theta}$ contained in the booms at full erection governs the structural design of the boom far more than does the torque during erection.* For acceptance testing of a boom system it is therefore important to simulate the kinetic energy of the boom at impact in the erected position. This paper presents a method for doing this and neglects any attempted simulation of the torsional forces involved.

SIMULATION PROCEDURE

To simplify calculations it is assumed that all tests will be run with the spin axis of the system parallel to the 1-g force field. Also, it is assumed that full erection occurs at $\theta = \pi/2$.

For a correct simulation at $\theta = \pi/2$, the kinetic energy E_z contained by the booms due to $\dot{\theta}$ when tested in a 1-g field must equal the kinetic energy E_z in the booms due to $\dot{\theta}$ when erection occurs in space, that is, $E_z = E_g$.

Solution for E_z

Because of the conservation of energy, $T_z = \text{constant}$. When T_z is evaluated at $t = 0$ ($\theta = 0^\circ$) and the result is equated to Equation 3, an expres-

sion is obtained which may be solved for E_z .

$$\begin{aligned} E_z &= \frac{1}{2} n\dot{\theta}^2 \left(mr^2 + \frac{1}{3} M^2 \right) \\ &= \frac{1}{2} na^2\dot{\lambda}_0^2 (m + M) + \frac{1}{2} I_b\dot{\lambda}_0^2 - \frac{1}{2} nm\dot{\lambda}^2 (r \sin \theta + a)^2 \\ &\quad - \frac{1}{2} nM\dot{\lambda}^2 \left(\frac{1}{3} l^2 \sin^2 \theta + al \sin \theta + a^2 \right) - \frac{1}{2} I_b\dot{\lambda}^2 . \end{aligned} \quad (7)$$

Evaluated at $\theta = \pi/2$, the expression is:

$$\begin{aligned} E_z &= \frac{1}{2} na^2\dot{\lambda}_0^2 (m + M) + \frac{1}{2} I_b\dot{\lambda}_0^2 \\ &\quad - \frac{1}{2} nm\dot{\lambda}_2^2 (r + a)^2 - \frac{1}{2} nM\dot{\lambda}_2^2 \left(\frac{1}{3} l^2 + al + a^2 \right) - \frac{1}{2} I_b\dot{\lambda}_2^2 , \end{aligned} \quad (8)$$

which is easily simplified to

$$E_z = \frac{1}{2} I_0\dot{\lambda}_0^2 \left(1 - \frac{I_0}{I_2} \right) . \quad (9)$$

Solution for E_g

The term E_g is obtained by first writing Equation 3 for a system in a conservative field of 1-g, that is,

$$\begin{aligned} T_g &= \frac{1}{2} nmr^2\dot{\theta}^2 + \frac{1}{6} nM^2\dot{\theta}^2 + \frac{1}{2} nm\dot{\lambda}^2 (r \sin \theta + a)^2 \\ &\quad + \frac{1}{2} nM\dot{\lambda}^2 \left(\frac{1}{3} l^2 \sin^2 \theta + al \sin \theta + a^2 \right) + \frac{1}{2} J_b\dot{\lambda}^2 + V , \end{aligned} \quad (10)$$

where V equals the total potential energy of the entire system. At $t=0$, $\theta=0^\circ$, this becomes

$$T_g = \frac{1}{2} na^2\dot{\lambda}_0^2 (m + M) + \frac{1}{2} J_b\dot{\lambda}_0^2 + V_0 = \text{Constant} . \quad (11)$$

This expression can now be substituted for T_g in Equation 10, and the result solved for E_g :

$$\begin{aligned} E_g &= \frac{1}{2} n\dot{\theta}_2^2 \left(mr^2 + \frac{1}{3} M^2 \right) \\ &= \frac{1}{2} na^2\dot{\lambda}_0^2 (m + M) + \frac{1}{2} J_b\dot{\lambda}_0^2 + V_0 - \frac{1}{2} nm\dot{\lambda}^2 (r \sin \theta + a)^2 \\ &\quad - \frac{1}{2} nM\dot{\lambda}^2 \left(\frac{1}{3} l^2 \sin^2 \theta + al \sin \theta + a^2 \right) - \frac{1}{2} J_b\dot{\lambda}^2 - V_\theta , \end{aligned} \quad (12)$$

where V_θ denotes the potential energy of the entire system for any value of θ .

When Equation 12 is evaluated at $\theta = \pi/2$ and simplified, the following is obtained:

$$E_g = \frac{1}{2} J_0\dot{\lambda}_0^2 \left(1 - \frac{J_0}{J_2} \right) - (V_2 - V_0) , \quad (13)$$

where the term $(V_2 - V_0)$ expresses the change in potential energy of the entire system due to the erection of the booms.

Solution for λ_{0g}

If the initial angular velocity λ_0 and all of the inertias, masses, and dimensions for the entire system are the same in both the 0-g and 1-g states, then by comparing Equations 9 and 13 it can be seen what must differ between the two conditions. Since in a 1-g field the term $(V_2 - V_0)$

is not zero, E_g will be less than E_x by the amount $(V_2 - V_0)$. The boom angular velocity $\dot{\theta}_{2g}$ at $\theta = \pi/2$ in the 1-g field must be less than $\dot{\theta}_{2g}$ in the 0-g field. Therefore, to achieve the result $E_x = E_g$, the right-hand sides of Equations 9 and 13 must be made equal at the point of full erection. Thus,

$$\frac{1}{2} I_0 \lambda_0^2 \left(1 - \frac{I_0}{I_2}\right) = \frac{1}{2} J_0 \lambda_{0g}^2 \left(1 - \frac{J_0}{J_2}\right) - (V_2 - V_0) \quad (14)$$

Inspection of Equation 14 shows that the inertias, the initial system spin rate, or both may be altered to make the identity hold. Since normally any testing will be accomplished with a model of the satellite, it usually is easy to "slug up" the system so that $I_0 = J_0$ and $I_2 = J_2$. Therefore, with a known initial design spin rate of λ_0 it is easy to solve Equation 14 for the only unknown, the initial spin rate of the test system λ_{0g} :

$$\lambda_{0g} = \sqrt{\lambda_0^2 + \frac{2I_2(V_2 - V_0)}{I_0(I_2 - I_0)}} \quad (15)$$

Test Arrangement

The spin test setup should be a simulation of the hardware shown in Figure 1 mounted on a suitable spin table. A typical setup is shown in Figure 2.

It is imperative that the entire system be "free-wheeling" during the boom erection, that is, that there be as little as possible externally applied torque (drive power, friction, air drag, etc.) acting on the system. Some type of clutching arrangement, as is shown in Figure 2, is therefore necessary to remove the drive motor from the system just prior to boom release. A well engineered bearing design will reduce the friction torque to a negligible amount. The torque induced by wind resistance generally may be neglected unless the swinging appendages present a large area; in such a case it might be necessary to conduct the tests in a vacuum chamber.

Use of λ_{0g}

With the setup shown in Figure 2, the only mass to undergo a potential energy change will be the swinging booms.

By solving for $V_2 - V_0$ and applying this to Equation 15 for an actual initial spin rate of λ_0 , a

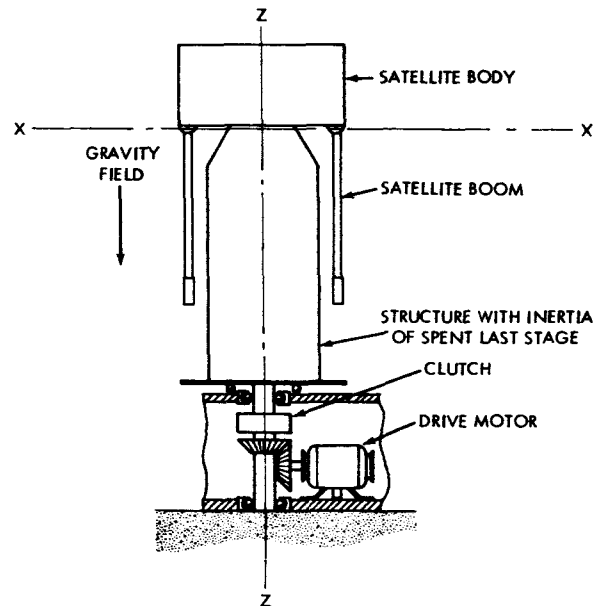


FIGURE 2.—Spin test setup.

new initial spin rate λ_{0g} is obtained. If in a force field test the booms are released at a system angular velocity of λ_{0g} , the booms' angular velocity $\dot{\theta}$ will equal the 0-g boom angular velocity at one point, that is, at $\theta = \pi/2$, and the simulation will be sufficient.

If the booms are swung in the up direction during the tests, then λ_{0g} will be higher than the actual flight λ_0 . Also, because of the conservation of momentum, the final test angular velocity λ_{2g} naturally will be higher than the actual flight λ_2 . The change in system energy during erection in the tests will be greater than in space flight by the amount of potential energy gained; however, the angular velocity of the booms will be flight level at the instant of erection.

It should be remembered that the tests may be

run with the appendages swinging down rather than up. In this case the potential energy to be dissipated by the booms would have to be subtracted from the initial system energy.

The procedure described, with the appendages swinging up, was used for acceptance tests of the satellite Ariel I. The solar paddles and the mass booms, which were used for spin stabilization, were tested in this manner. Appendix A gives the dynamic calculations for simulated flight erections of the mass booms.

CONCLUSION

This method of testing has distinct advantages: it can be quickly incorporated into a test program;

no uncertain quantities are introduced into the test setup; and no elaborate test setup is necessary.

Examination of the equations involved will reveal the disadvantages or limitations. It is always necessary to know accurately and to control all parameters (spin rates, inertias, and energy changes) in the test system. If the change in potential energy ($V_2 - V_0$) is large in comparison to the planned initial spin rate λ_{0z} , then precision control of all test parameters becomes necessary to prevent the introduction of a large simulation error. Correctly used, however, this technique provides a proper means of simulating the space erection of satellite appendages, and tests the entire system for impact.

APPENDIX A

CALCULATIONS FOR SIMULATING THE ZERO-GRAVITY ERECTION OF THE ARIEL I MASS BOOMS

As is shown in Figure A1, the Ariel I satellite had eight erectable appendages: two experiment-carrying booms, two mass booms for spin stabilization, two solar paddles with rigid arms, and two solar paddles with secondary hinges in the arms. In actual flight, prior to any erection, the append-

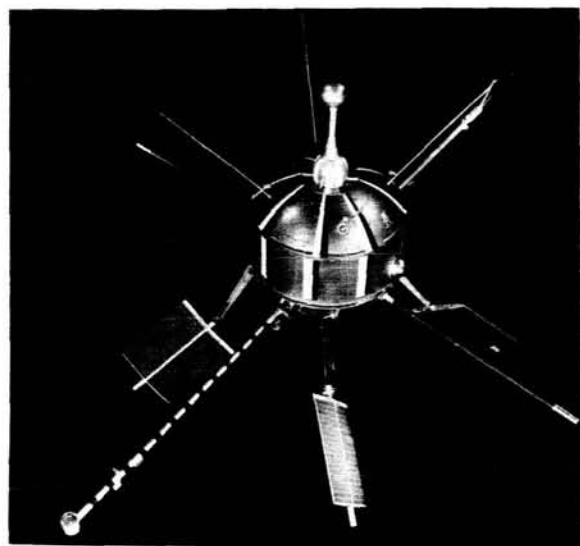


FIGURE A1.—The Ariel I satellite.

ages were folded down beside the last-stage rocket bottle. The planned erection sequence was as follows: (1) the system was to be spin-stabilized at 161 rpm prior to the last-stage ignition; (2) 15 minutes after last-stage ignition, a stretch yo-yo de-spin device would unwind, reducing the spin rate to 76 rpm; (3) 60 seconds after yo-yo release the two experimental booms were to be released and, under the constraint of a governor, erect in approximately 3.0 seconds; (4) 60 seconds after the previous action, with the system at approximately 52 rpm, the four solar paddles and two mass booms would be released simultaneously and allowed to swing freely to their erected positions. At this point the system would be at approxi-

mately 35 rpm; (5) 60 seconds later, separation of the spacecraft from the last-stage bottle would occur and the satellite would be in orbit as seen in Figure A1.

In the design of the solar paddles and mass booms it was necessary to solve for the impact energy at erection of each type appendage. Since they were all released at the same instant but did not all reach their erected positions at the same instant, the analytical solution of the process was considerably more involved than the case where only one type of appendage erects at a time. In the last case, the impact energy at erection in a force free field is:

$$E_x = \frac{1}{2} I_0 \lambda_0^2 \left(1 - \frac{I_0}{I_2} \right), \quad (A1)$$

where

E_x is the energy that must be absorbed to stop the booms in the erected position,

I_0 is the initial total moment of inertia about the spin axis of the satellite, spent last stage, and attachments,

I_2 is the final moment of inertia about the spin axis of the satellite, spent last stage, and attachments,

λ_0 is the initial angular velocity of the entire system about the symmetrical axis.

The more involved and actual process of simultaneous erection of dissimilar appendages was solved by writing Lagrangian expressions involving the three types of appendages and solving the equations on a computer. The resultant actual impact energy load on the two mass booms due to a flight erection of the six appendages at an initial λ_{0z} of 52 rpm was calculated to be 8.51 ft-lb. During development testing of the appendages, it was necessary to test the solar paddles and mass booms separately. For a proper acceptance test of the mass booms it was then necessary to conduct tests simulating the booms containing 8.51 ft-lb of kinetic energy at the point of erection. When the mass booms were tested separately,

this was accomplished by first calculating an initial 0-g system angular velocity λ_{0z} , which corresponds to 8.51 ft-lb of kinetic energy in the mass booms at erection. For the tests it was assumed the experimental booms were open prior to the mass booms release and the solar paddles remained folded. The Ariel I moments of inertia for these conditions were: I_0 , 4.01 slug-ft²; and I_2 after mass boom full erection, 4.84 slug-ft². With these values, Equation A1 gives $\lambda_{0z} = 47.5$ rpm. The test, therefore, was to simulate a 0-g mass boom erection at an initial rotation rate of 47.5 rpm.

From Equation 15 in the body of the report, that is,

$$\dot{\lambda}_{0g} = \sqrt{\dot{\lambda}_0^2 + \frac{2I_2(V_2 - V_0)}{I_0(I_2 - I_0)}} \quad (A2)$$

an initial spin rate λ_{0g} was calculated to simulate a space erection at 47.5 rpm. Figures A2 and A3 show the test system before and after erection.

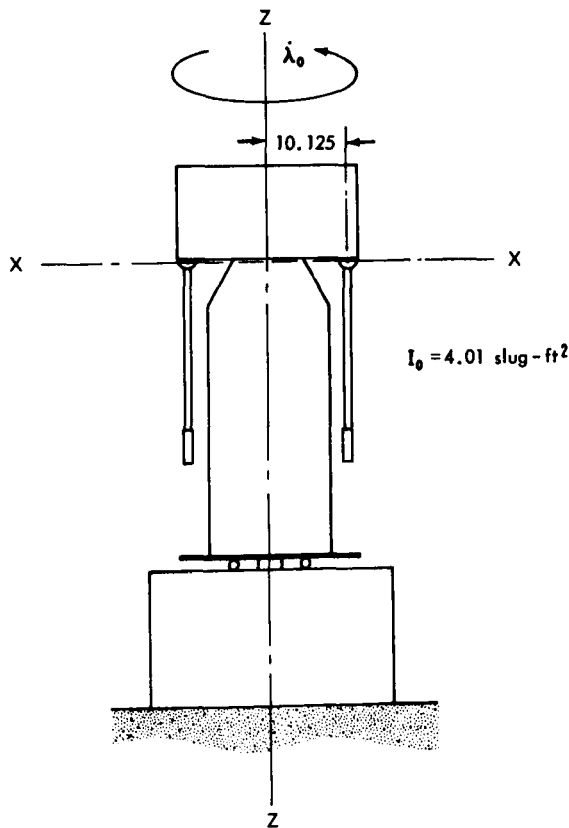


FIGURE A2.—Test system before boom erection.

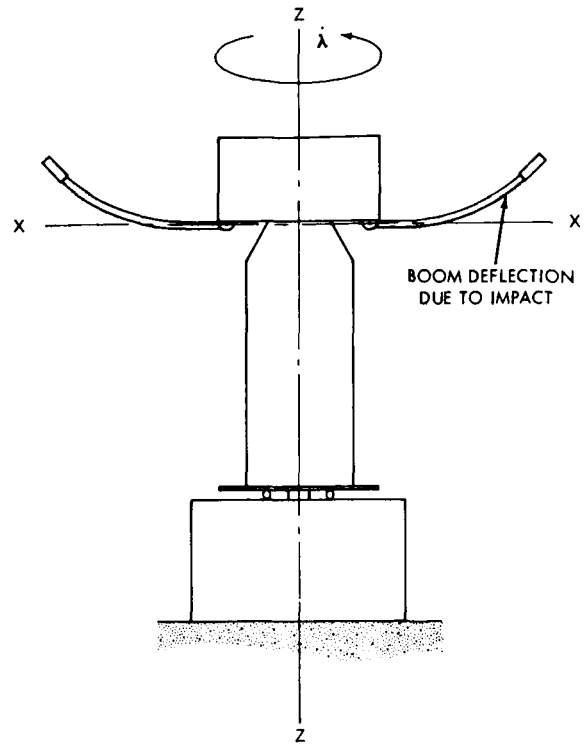


FIGURE A3.—Test system at boom erection.

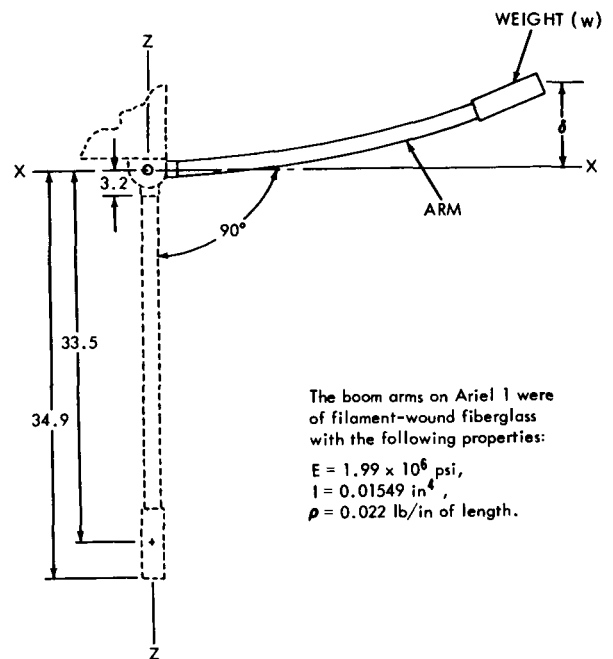


FIGURE A4.—Mass boom displacement.

To use Equation A2, the potential energy change $V_2 - V_0$ between the configurations shown in Figures A2 and A3 was calculated. Obviously,

the mass booms were the only consideration. Figure A4 shows the extreme positions of each mass boom.

The weight w on the end of each arm was

$$V_2 - V_0 = n\rho \int_0^l (x - x \cos \theta) dx + nw \int_0^l (1 - \cos \theta) dx + n\rho \int_0^l z dx + nw \int_0^z dz, \quad (A3)$$

where the last two terms in the expression consider the deflection. For full erection at $\theta = \pi/2$ and

$$n\rho \int_0^l (x - x \cos \theta) dx + nw \int_0^l (1 - \cos \theta) dx = n\rho \frac{l^2}{2} + nwr$$

$$= 77.3 \text{ in-lb.}$$

A close approximation for the third term in Equation A3 was obtained by approximating the curve of the deflected boom arm in Figure A4 by a quarter cosine wave function,* that is,

$$z = \delta \left(1 - \cos \frac{\pi x}{2l}\right)$$

and

$$\begin{aligned} n\rho \int_0^l z dx &= n\rho \delta \int_0^l \left(1 - \cos \frac{\pi x}{2l}\right) dx \\ &= n\rho \delta l \left(1 - \frac{2}{\pi}\right) = 3.2 \text{ in-lb.} \end{aligned} \quad (A4)$$

The last term in Equation A3 is simply:

$$nw \int_0^z dz = 9.0 \text{ in-lb.} \quad (A5)$$

and so the total potential energy change is

$$\begin{aligned} V_2 - V_0 &= 77.3 + 3.2 + 9.0 = 89.5 \text{ in-lb} \\ &= 7.46 \text{ ft-lb.} \end{aligned}$$

All the known values are then applied to Equation A2, and

$$\begin{aligned} \dot{\lambda}_{0R} &= \sqrt{(4.974)^2 + \frac{2(4.84)(7.46)}{4.01(4.84 - 4.01)}} \\ &= 65.1 \text{ rpm.} \end{aligned}$$

0.775 lb. For $E_z = 8.51$ ft-lb the deflection δ was calculated to be approximately 5.9 inches. This δ was large enough to warrant consideration when calculating $V_2 - V_0$. Therefore.

with $l \approx 34.0$ (arm extended into the weight), the first two terms of Equation A3 are:

Thus, the actual configuration shown in Figure A5 was spun up to approximately 67.0 rpm, the drive disengaged, and the spin allowed to decay to 65.1 rpm, at which point the booms were released and allowed to swing freely to the erected position.

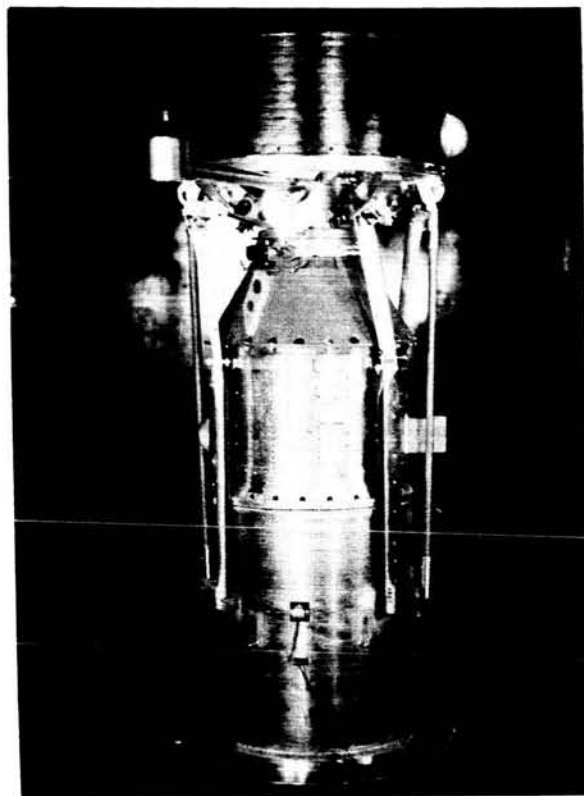


FIGURE A5.—The Ariel I test configuration.

*Den Hartog, J. P., "Mechanical Vibrations," 4th Ed., New York: McGraw-Hill, 1956, p. 153.

Since deflection was considered, $\dot{\theta}_g$ was slightly higher than the actual flight $\dot{\theta}_z$ at $\theta = \pi/2$. However, the total potential energy in the booms at full deflection for the tests, when $\dot{\lambda}_{0g} = 65.1$ rpm, equaled the potential energy in the booms at full deflection for a space erection when $\dot{\lambda}_{0z} = 47.5$ rpm. Figure A6 shows the calculated plots of the kinetic energy in the booms versus the boom angle θ for the actual space and test conditions.

From Figure A6 it is seen that at the instant when $\theta = 90^\circ$, or full erection, the booms had about 12 in-lb more kinetic energy in the tests than during the actual flight. However, the change in potential energy in the tests due to the upward deflection of the booms subtracted from the kinetic energy. From Equations A4 and A5 it will be seen that the calculated potential energy change due to deflection equals 12.2 in-lb. From high-speed motion pictures of the erection tests, it was ascertained that the total deflection δ of each boom end was approximately 6 inches. Therefore, at full deflection, the strain of each boom was approximately equal to that anticipated in the actual flight and simulation was completed.

If the erecting appendages had been fairly rigid and the deflection was negligible, then a graph of the variables in Figure A6 would have shown that the two kinetic energy curves terminate on the

same value at $\theta = 90^\circ$. The angular velocities $\dot{\theta}_z$ and $\dot{\theta}_g$ would have been equal at that point.

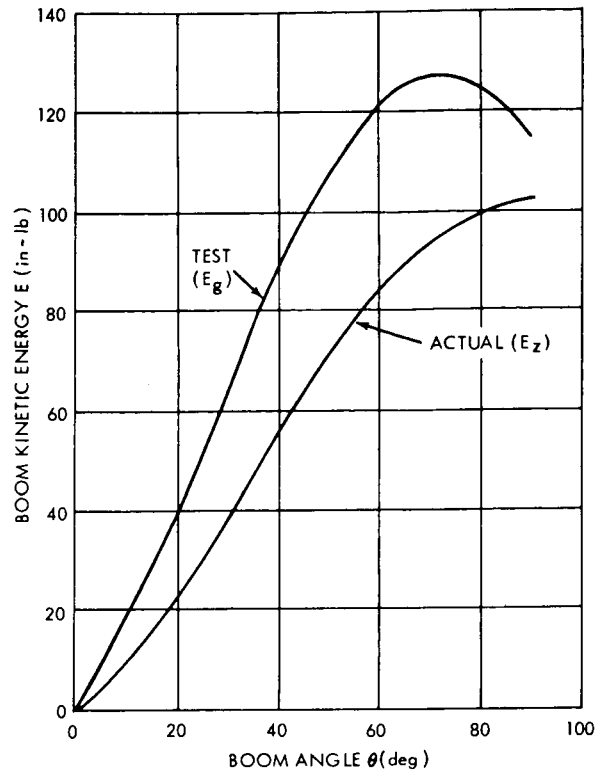


FIGURE A6.—Calculated plots of the kinetic energy in the Ariel I mass booms under both test and actual flight conditions.

MODELLING OF SPACECRAFT UNDER RANDOM LOADING

JOSHUA E. GREENSPON

INTRODUCTION

As spacecraft become larger and more complicated it becomes more difficult and expensive to test the full scale structures. In addition, if specifications for a spacecraft are to be obtained before the vehicle and payload are constructed, the only way that predicted response can be obtained is for scale model tests to be conducted and calculations performed on the proposed design. If small models could be constructed of materials which were easier to handle in manufacturing and less expensive than the full scale, a great deal of information could be obtained efficiently on a proposed space vehicle. The purpose of this report is to present a derivation of the modelling laws for random dynamic loading on an elastic structure such as a spacecraft-booster combination which is constructed of a different material than the prototype.

DERIVATION OF THE MODELLING LAWS

A. Forces and Motions

The theory presented here is an application and extension of the basic theory of modelling presented by Weber.¹ Let the quantities associated with the full scale be primed and those associated with the model unprimed. Thus if ℓ' is the characteristic length in the full scale, t' the full scale time, and k' the force in the full scale.

$$\ell' = \ell\lambda \quad t' = t\tau \quad k' = k\chi \quad (1)$$

where λ is the linear dimension scale factor, τ is the time scale factor, and χ is the force scale factor.

The phenomenon that is to be scaled is random loading and response of an elastic structure. The only assumption that will be made is that the model is geometrically similar to the prototype

and is subjected to a geometrically similar load distribution.

The basic equations governing the elastic motions of the structure in the absence of body forces are the three dimensional equations of elasticity.²

$$\begin{aligned} (\bar{\lambda} + \bar{\mu}) \left(\frac{\partial \Delta}{\partial x} \frac{\partial \Delta}{\partial y} \frac{\partial \Delta}{\partial z} \right) + \bar{\mu} \nabla^2(u, v, w) \\ = \rho \left(\frac{\partial^2 u}{\partial t^2}, \frac{\partial^2 v}{\partial t^2}, \frac{\partial^2 w}{\partial t^2} \right) \end{aligned} \quad (2)$$

where $\bar{\lambda}$ and $\bar{\mu}$ are the Lamé constant, Δ is the cubical dilatation and ρ is the mass density of the material. The displacements are u, v, w .

The cubical dilatation and the Laplacian are

$$\begin{aligned} \Delta = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \\ \nabla^2(u, v, w) = \frac{\partial^2(u, v, w)}{\partial x^2} + \frac{\partial^2(u, v, w)}{\partial y^2} \\ + \frac{\partial^2(u, v, w)}{\partial z^2} \end{aligned} \quad (3)$$

The Lamé constants $\bar{\lambda}$ and $\bar{\mu}$ can be written in terms of the modulus of elasticity E and Poisson's ratio ν as follows

$$\bar{\lambda} = \frac{E\nu}{(1+\nu)(1-2\nu)} \quad \bar{\mu} = \frac{E}{2(1+\nu)} \quad (4)$$

So the x component of the equations of motion becomes

$$\begin{aligned} \frac{E}{2(1+\nu)(1-2\nu)} \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 v}{\partial x \partial y} + \frac{\partial^2 w}{\partial x \partial z} \right) \\ + \frac{E}{2(1+\nu)} \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) = \rho \frac{\partial^2 u}{\partial t^2} \end{aligned} \quad (5)$$

and there are two other equations for the y and z components. The scaling laws can be derived by merely working with one of these equations. In accordance with our notation, eq. (5) is for the model since it contains unprimed quantities. The equivalent equation for the prototype, assuming a different material with approximately the same Poisson's ratio but different E and ρ^* is

$$\frac{E'}{2(1+\nu)(1-2\nu)} \left(\frac{\partial^2 u'}{\partial \chi'^2} + \frac{\partial^2 v'}{\partial \chi' \partial y'} + \frac{\partial^2 w'}{\partial \chi' \partial z'} \right) + \frac{E}{2(1+\nu)} \left(\frac{\partial^2 u'}{\partial \chi'^2} + \frac{\partial^2 u'}{\partial y'^2} + \frac{\partial^2 w'}{\partial z'^2} \right) = \rho' \frac{\partial^2 u'}{\partial t'^2} \quad (6)$$

Let the scale factors for E and ρ be

$$\frac{E'}{E} = \sigma_E \quad \frac{\rho'}{\rho} = \sigma_\rho \quad (7)$$

By the choice of model material and linear scale this will fix σ_E , σ_ρ and λ . The rest of the scale factors will then be derived from these basic quantities. Thus the time factor τ will be determined as follows:

$$\frac{E'}{E} = \sigma_E = \frac{k'/\ell'^2}{k/\ell^2} = \frac{k\chi/(\ell^2\lambda^2)}{k/\ell^2} = \frac{\chi}{\lambda^2}$$

$$\frac{\rho'}{\rho} = \sigma_\rho = \frac{k'\ell'^2/\ell'^4}{k\ell^2/\ell^4} = \frac{k\chi t^2 \tau^2 / (\ell^4 \lambda^4)}{k\ell^2/\ell^4} = \frac{\chi \tau^2}{\lambda^4} \quad (8)$$

thus

$$\chi = \sigma_E \lambda^2$$

and

$$\tau = \sqrt{\frac{\sigma_\rho \lambda^4}{\chi}} = \lambda \sqrt{\frac{\sigma_\rho}{\sigma_E}}$$

(frequency is therefore modelled as

$$\frac{1}{\tau} = \frac{1}{\lambda} \sqrt{\frac{\sigma_E}{\sigma_\rho}})$$

Equation (6) for the prototype can then be written

$$\frac{\sigma_E E}{2(1+\nu)} \frac{1}{\lambda} \left[\frac{\partial^2 u}{\partial \chi^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} + \frac{1}{1-2\nu} \left(\frac{\partial^2 u}{\partial \chi^2} + \frac{\partial^2 v}{\partial \chi \partial y} + \frac{\partial^2 w}{\partial \chi \partial z} \right) \right] = \sigma_\rho \rho \frac{\lambda}{\tau^2} \frac{\partial^2 u}{\partial t^2} \quad (9)$$

*Usually the Poisson's ratio is of secondary importance. Moreover the Poisson's ratio for the model probably will not vary greatly from the full scale material.

Collect all scaling terms into a single coefficient on the left. In order for the model to obey the same dynamic equation as the prototype, the scaling relation must be as follows

$$\frac{E'/E}{\rho'/\rho} \frac{t'^2/\ell^2}{\ell'^2/\ell^2} = 1 \quad (10)$$

or

$$\frac{E'}{f'^2 \rho' \ell'^2} = \frac{E}{f^2 \rho \ell^2}$$

or finally

$$\sqrt{\frac{f' \ell'}{E'}} = \sqrt{\frac{f \ell}{E}}$$

where f is the frequency. The physical number

$$\sqrt{\frac{f \ell}{E}} \frac{1}{\rho}$$

is a type of Strouhal number. If the same material were used in the model and prototype then Junger's relations³ would hold, i.e.

$$f \ell = f' \ell' \quad (11)$$

For Junger's case frequency is scaled inversely to length. Equations (8) and (10) are the basic relations for modelling with different materials.⁴ The other quantities involved in the problem can be derived from these relations. Thus

Pressure:

$$p'/p = \chi/\lambda^2 \quad p' = \sigma_E p$$

Stress:

$$\sigma'_x/\sigma_x = \chi/\lambda^2 \quad \sigma'_x = \sigma_E \sigma_x \quad (12)$$

Velocity:

$$\frac{v'}{v} = \frac{\ell'/t'}{\ell/t} = \frac{\lambda \ell / t \tau}{\ell/t} = \frac{\lambda}{\tau} = \sqrt{\frac{\sigma_E}{\sigma_\rho}}$$

(for the same material and velocities scale directly)

Acceleration:

$$\frac{a'}{a} = \frac{\ell'/t'^2}{\ell/t^2} = \frac{\lambda \ell / t^2 \tau^2}{\ell/t^2} = \frac{\lambda}{\tau^2} = \frac{1}{\lambda} \frac{\sigma_E}{\sigma_\rho}$$

The statistical quantities will then be scaled as follows:

Mean square pressure:

$$\frac{\langle p'^2 \rangle}{\langle p^2 \rangle} = \sigma_E^2$$

Correlation function:

$$\left(\text{i.e. } \frac{c'}{c} = \frac{\text{Correlation full scale}}{\text{Correlation model}} \right)$$

(a) Pressure:

$$\langle p_1' p_2' \rangle / \langle p_1 p_2 \rangle = \sigma_E^2$$

(b) Displacement:

$$\langle w_1' w_2' \rangle / \langle w_1 w_2 \rangle = \lambda^2$$

(c) Velocity:

$$\langle v_1' v_2' \rangle / \langle v_1 v_2 \rangle = \sigma_E / \sigma_p$$

(d) Acceleration:

$$\frac{\langle a_1' a_2' \rangle}{\langle a_1 a_2 \rangle} = \frac{1}{\lambda^2} \left(\frac{\sigma_E}{\sigma_p} \right)^2$$

Power spectral density:

$$\text{i.e. } \frac{(p.s.)'}{(p.s.)} = \frac{\text{Power spectral density full scale}}{\text{Power spectral density model}}$$

(a) Pressure:

$$\sigma_E^2 \hat{\tau} = \sigma_E^2 \lambda \sqrt{\frac{\sigma_p}{\sigma_E}} = \lambda \sqrt{\sigma_p \sigma_E^3}$$

(b) Displacement:

$$\lambda^2 \hat{\tau} = \lambda^2 \lambda \sqrt{\frac{\sigma_p}{\sigma_E}} = \lambda^3 \sqrt{\frac{\sigma_p}{\sigma_E}} \tag{13a}$$

(c) Velocity:

$$\frac{\sigma_E}{\sigma_p} \hat{\tau} = \frac{\sigma_E}{\sigma_p} \lambda \sqrt{\frac{\sigma_p}{\sigma_E}} = \lambda \sqrt{\frac{\sigma_E}{\sigma_p}}$$

(d) Acceleration:

$$\frac{1}{\lambda^2} \left(\frac{\sigma_E}{\sigma_p} \right)^2 \hat{\tau} = \frac{1}{\lambda^2} \left(\frac{\sigma_E}{\sigma_p} \right)^2 \lambda \sqrt{\frac{\sigma_p}{\sigma_E}} = \frac{1}{\lambda} \sqrt{\frac{\sigma_E^3}{\sigma_p^3}}$$

B. Damping

One of the big problems in using scale models is the fact that damping does not scale. However, this fact should not hold us back if we analyze how Coulomb damping and material damping enter into the problem. Coulomb damping will arise as external damping through rubbing between attached pieces. The basic relation for Coulomb damping states that the frictional shearing stress is proportional to the normal stress, i.e.

$$(s.s.) = \mu(N.S.) \tag{14}$$

where μ is the coefficient of sliding friction between the two surfaces. For the prototype this relation will be

$$(s.s.)' = \mu'(N.S.)'$$

where

$$\mu' = \sigma_\mu \mu \tag{15}$$

Rewriting this equation for the model

$$\sigma_E (s.s.) = \mu \sigma_\mu \sigma_E (N.S.) \tag{16}$$

Thus σ_μ must be unity, or the coefficient of sliding friction in the model must be adjusted to equal that in the full scale. This could possibly be done by lubrication methods or roughing the surface depending on whether the coefficient of friction has to be increased or decreased.

For material damping the problem is not as direct and we must resort to some assumptions regarding the response of a structure to random loading. For different damping laws the response will take different forms. Take the case in which the damping is viscous. The power spectral density of the displacement of a structure under random loading can, for practical purposes, be written as⁵

$$(p.s.)_d = \sum_n \sum_m \frac{\gamma_n \gamma_m}{|Y_n| |Y_m|} Cmn(\omega) \tag{17}$$

where γ is associated with the mode shape of the structure, Cmn is the correlation integral which for all practical purposes can be considered independent of material damping, and Y is the response function which is critically dependent on

damping. The response function Y can be written

$$Y_n(\omega) = M_n \sqrt{(\omega_{0n}^2 - \omega^2)^2 + \beta_n^2 \omega^2} \quad (18)$$

where M_n is the generalized mass for the n th mode, and β_n is the damping coefficient for this mode. The logarithmic decrement δ_n can be written.

$$\delta_n = \frac{\pi \beta_n}{\omega_{0n}} \quad \beta_n = \frac{\delta_n \omega_{0n}}{\pi} = \bar{\beta} \quad (19)$$

Fung et al.⁶ found that β_n was approximately a constant for various modes of aluminum cylindrical shells. Thus at resonance, where $\omega_{0n} = \omega$

$$Y_n(\omega) = M_n \beta_n \omega_{0n} = M_n \beta \omega_{0n} \quad (20)$$

In a structure subjected to random loading the most severe response in the spectrum will occur at the frequencies corresponding to resonances of the modes. Therefore at these frequencies

$$(p.s.)_d \sim \frac{1}{\bar{\beta}^2} = \bar{\delta}^2 \quad (21)$$

So for a model of different material than the full scale

$$\frac{(p.s.)_{d'}}{(p.s.)_d} \sim \frac{1/\bar{\beta}'^2}{1/\bar{\beta}^2} = \frac{\delta'^2}{\delta^2} \quad (22)$$

Thus in order to scale material damping properly at the critical frequencies

Displacement:

$$\frac{(p.s.)_{d'}}{(p.s.)_d} \approx \frac{\bar{\delta}'^2}{\bar{\delta}^2} \lambda^3 \sqrt{\frac{\sigma_p}{\sigma_E}}$$

Velocity:

$$\frac{(p.s.)_{v'}}{(p.s.)_v} \approx \frac{\bar{\delta}'^2}{\bar{\delta}^2} \lambda \sqrt{\frac{\sigma_E}{\sigma_p}} \quad (23)$$

Acceleration:

$$\frac{(p.s.)_{a'}}{(p.s.)_a} \approx \frac{\bar{\delta}'^2}{\bar{\delta}^2} \frac{1}{\lambda} \sqrt{\frac{\sigma_E^3}{\sigma_p^3}} \quad (23)$$

The ratio δ'/δ must be determined experimentally for the different materials. However, a

relatively straight forward experiment might be used to do this; possibly an experiment such as the one performed by Fung, et al.⁶ could be performed on cylindrical shell models made of the prototype and model materials.

A more basic assessment of material damping and its dependency on level of excitation is offered by Crandall.⁷ In Crandall's analysis of damping for a nonlinear relaxation process the value of $\bar{\beta}$ (defined by (19)) would be:

$$\bar{\beta} = 2g_s \omega_n = 2 \left[R \left(\frac{\rho E}{S_0} a \right)^{\frac{\bar{n}}{\bar{t}}} \frac{1}{\bar{t}} \right]^{\frac{1}{\bar{n}+1}} \quad (19a)$$

where g_s is the damping coefficient (equal to $\frac{\delta}{2\pi}$, δ being the logarithmic decrement). R is a dimensionless number which depends on the configuration of the body and the mode under consideration, ρ is the mass density of the material, E is the modulus of elasticity of the material, S_0 , \bar{n} and \bar{t} are material constants, and a is the acceleration or a measure of the same. If it is assumed that R is independent of the mode and only dependent on the configuration of the body, then

$$\frac{\bar{\delta}'^2}{\bar{\delta}^2} \approx \frac{\left[\left(\frac{\rho E}{S_0} a \right)^{\frac{\bar{n}}{\bar{t}}} \frac{1}{\bar{t}} \right]^{\frac{2}{\bar{n}+1}}}{\left[\left(\frac{\rho' E'}{S_0'} a' \right)^{\frac{\bar{n}'}{\bar{t}'}} \frac{1}{\bar{t}'} \right]^{\frac{2}{\bar{n}'+1}}} \quad (22a)$$

If the material constants ρ , E , S_0 , \bar{t} and \bar{n} for the model and prototype can be found then the ratio (22a) will be dependent on a power of the ratio of excitation accelerations. If a and a' are interpreted as the total accelerations then a can be measured in the model, scaled to the full scale by eq. (12) to obtain a' , and then (22a) will give the desired ratio to substitute into eq. (23).

DISCUSSION

The use of very small scale models of different material than the prototype offers a very efficient and inexpensive way to obtain order of magnitude answers which would ordinarily be very difficult or impossible to obtain. This type of scaling also offers an opportunity to use much lower pressure excitation levels than the full scale. There might also be a possibility of using light weight easy to construct plastic models to perform initial tests.

Invariably damping offers a big problem in using scale model results to extrapolate full scale response. If we limit our discussion to the resonance response then damping can be introduced directly in the scaling as shown in the previous section. This resonant limitation is not serious in random loading since the major response at each frequency of interest is usually composed of primarily resonant contributions in modes which are close to this frequency.

REFERENCES

1. M. WEBER, "The Universal Principle of Similitude in Physics and its Relation to the Dimensional Theory and the Science of Models," DTMB Translation 200, Sept. 1949.
2. A. E. H. LOVE, "The Mathematical Theory of Elasticity," Dover Publications, p. 293.
3. M. C. JUNGER, "The Scaling Laws Governing the Dynamic Response of a Structure to a Turbulent Boundary Layer," U. S. Navy Journal of Underwater Acoustics, April 1963, p. 439.
4. These scaling laws are implicitly contained in the results of WRIGHT, et al., Westinghouse Research Labs., Buships Contract Nobs 72326, Oct. 1962 and in Mixson, et al, NASA TN D-1593, Jan. 1963. The generality of these results, however, is not apparent from this earlier work.
5. A. POWELL, "On the Response of Structures to Random Pressures and to Jet Noise in Particular," Random Vibration, edited by S. Crandall, John Wiley and Sons, 1958, p. 187.
6. Y. C. FUNG, et al, Jour. Aero. Sci., Sept. 1957, p. 650.
7. S. H. CRANDALL, "On Scaling Laws for Material Damping," Report presented to Goddard Space Flight Center, NASA, Dec. 1961.

ENVIRONMENTAL TEST PROGRAM FOR ARIEL I*

WARNER H. HORD, JR.

The environmental test program for Ariel I, the International Ionosphere Satellite, was accomplished at the Goddard Space Flight Center between April 27, 1961, and March 13, 1962. This program consisted of design qualification tests for the prototype spacecraft and acceptance tests for the two flight spacecraft. The spacecraft separation system and the flight vibration experiment also received design qualification and flight acceptance tests.

Several failures in the prototype spacecraft occurred during the test program and resulted in various modifications and replacements. With these changes incorporated, the two flight spacecraft completed the acceptance tests with only minor difficulties and were shipped to Cape Canaveral in mid-March 1962.

One of the flight spacecraft, Flight Unit 1, was launched into orbit on April 26, 1962. Its flight performance suggests that the test program procedures were effective and that the resulting corrective actions improved spacecraft design and operation.

INTRODUCTION

This document presents a complete description of the environmental test program and system evaluation of the International Ionosphere Satellite, Ariel I (1962 o 1), accomplished at the Goddard Space Flight Center (GSFC). Ariel I is the first of a series of three satellites to be developed as a joint effort of the United States and the United Kingdom. The experiments were designed and built by the United Kingdom, while the United States designed and built the spacecraft and supporting subsystems and provided the launch vehicle.

The purposes of Ariel I are to make a thorough 1-year analysis of the ionosphere—that is, density, temperature, x-ray intensities, and particle mass measurement—and, in general, to acquire more knowledge of the ionosphere and its relation to the sun.

Project development efforts culminated in the successful launch of the Ariel I spacecraft from Cape Canaveral, Florida, on April 26, 1962. It weighed 60.6 kg (133.6 lb), of which 17.0 kg (37.5 lb) consisted of experiments. A technical description of the spacecraft and its subsystems is contained in Appendix A.

The spacecraft was placed in elliptical orbit with an apogee of 1214 km (754 miles), a perigee of 390 km (242 miles), and a period of 100.9 min. Figure 1 shows Ariel I in an orbital configuration.

The Ariel I environmental test program was designed to produce a high degree of confidence in the ability of the spacecraft to withstand the environments expected during handling, shipment, launch, and orbital flight. The program was divided into four phases: structural model tests, functional tests (such as spinup and despin), design qualification, and flight acceptance. The *functional* tests for despin and appendage erection were conducted to 90 percent and then 110 percent of anticipated flight spin rates. The *structural model and design qualification* tests were conducted at levels considerably more severe than those expected during handling, shipment, launch, and orbital flight. In the case of mechanical tests, the environments are 1.5 times those expected from flight while, in the case of thermal tests, a temperature 10°C in excess of the expected environment is used. The *acceptance* phase consisted of tests conducted at the levels of the expected environments. These tests were conducted on the flight model spacecraft to demonstrate that the design-qualified unit had been successfully duplicated and that there were no defects in materials or workmanship. A complete

*Published as NASA Technical Note D-2099, February 1964.

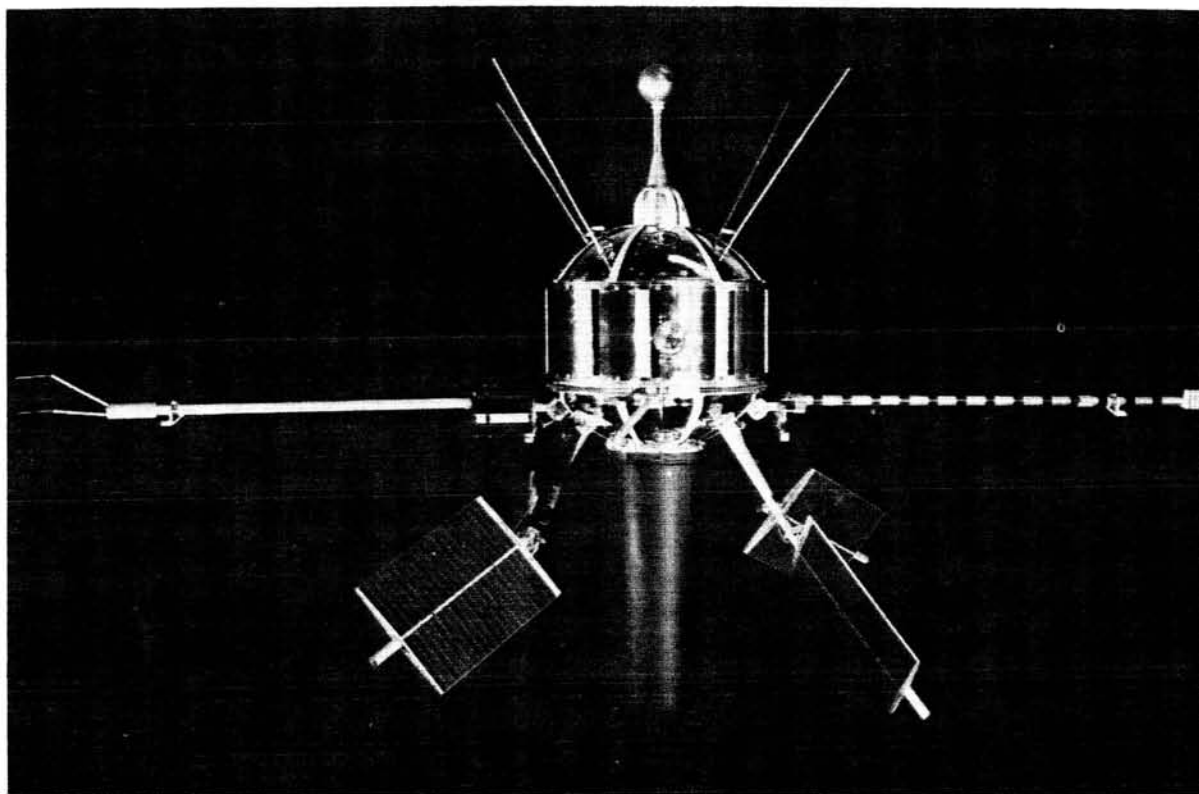


FIGURE 1.—Ariel I in orbital configuration.

test chronology of the environmental test program is given in Appendix B.

Pertinent background for the Ariel I test program is presented in the next section. The sections following cover the objectives of the test program; the proposed test plan; a description of the tests, and modifications and replacements based on test results; and, finally, an evaluation of the test results and the spacecraft system.

BACKGROUND

The original Ariel I plans called for the spacecraft to be launched by a Scout vehicle from the Wallops Island, Virginia, range in February 1962. As a result, the structural model tests were based on "Environmental Test Specification and General Test Procedures for Design Qualification and Flight Acceptance Testing of Scout Launched Satellites," August 11, 1961 (prepared by Test and Evaluation Division, GSFC); and the Scout configuration was used in the mounting arrange-

ment. These tests were conducted between April 26 and May 9, 1961.

On September 12, 1961, NASA Headquarters assigned the Ariel I spacecraft to a Delta launch vehicle with launch to be from Cape Canaveral in March 1962. It was determined unnecessary to repeat the structural model tests because they had been successful for Scout, and test levels for a Delta launch would not be significantly different. The change in launch vehicle did require that an adapter section, approximately 30.5 cm (12 in.) high, be inserted between the final-stage rocket motor and the spacecraft separation system; this adapter is referred to throughout this report as the Dutchman. Furthermore, the responsibility for procuring and testing the spacecraft separation system was transferred from Langley Research Center to GSFC.

Because of the added space provided by the Dutchman and the increase in spacecraft weight permitted by the greater thrust of the Delta, a separate telemetry system that made possible an

experiment for measuring vibration during the period from launch to the separation of the spacecraft from the third stage was installed in the Dutchman. This telemetry system had three subcarrier oscillators for vibrational data and a fourth subcarrier oscillator that provided data on nose-fairing contamination, closure of the third-stage pressure switch, spin rate, and aspect angle. Appendix C contains a description of the vibration telemetry system. As a result of these changes, tests of the separation system and vibration experiment had to be included in the environmental test program.

The design qualification test program of the prototype Ariel I spacecraft was conducted during the period September 16, 1961, through January 8, 1962. Because of modifications and changes determined necessary during the qualification tests, the vibration and thermal-vacuum tests were repeated during the period February 12 through March 12, 1962. Design qualification tests of the vibration experiment and separation system were conducted during the period January 15 through January 18, 1962, and were repeated during the period March 10 through March 15, 1962.

Acceptance tests of Flight Unit 1 were conducted during the period January 11 through February 27, 1962; and Flight Unit 2 acceptance tests were conducted February 5 through March 13, 1962. Acceptance tests of the flight separation system and vibration experiment were conducted during the period January through March 1962.

The Prototype Unit was shipped to the Atlantic Missile Range (AMR) on March 13 for coordination and interference checks with the launch vehicle. Subsequently, Flight Units 1 and 2 were shipped.

After checkout, Flight Unit 1 was selected for launch. Operations at AMR were relatively smooth with the only significant problems being that (1) the electron density experiment boom was damaged, requiring substitution of the Flight Unit 2 boom; and (2) a gear shaft of the boom escapement mechanism failed, and required replacement. (A chronology of AMR operations is given in Appendix D.) The first launch attempt on April 10, 1962, was cancelled because of problems with the second stage. A 2-week slip

was necessary so that the second stage of the booster could be replaced. The second launch attempt on April 26, 1962, was successful; and at 1300 hours EST, after a hold of 1 hour, the Ariel I Flight Unit 1 was launched into orbit.

TEST OBJECTIVES

General

The four test programs in which the Ariel I project was involved—structural model, functional, design qualification, and flight acceptance—had distinct purposes as described below. These programs had the common purpose of providing a spacecraft of maximum reliability at a reasonable cost by using test levels and procedures reasonably related to predicted environments of the Ariel I spacecraft.

Further, the series of environmental exposures provided an unparalleled opportunity for the operating crew to gain experience with the spacecraft system prior to operations at the launch site—thereby serving as an excellent training program.

Structural Model

The purpose of this program was to subject a structural model of the spacecraft with dummy components to environmental exposures up to the design qualification level in order to “prove out” the structural design prior to building a prototype spacecraft with operating subsystems. Thus, considerable confidence could be established prior to design qualification tests that the prototype spacecraft structure would withstand required levels of exposure.

Also, data from these tests were needed to establish levels in test specifications for subsystems to be mounted in the spacecraft.

Functional

This test program had the objective of determining the ability of a dynamic mockup of the spacecraft to perform functional requirements such as spinup, despin, boom erection, paddle deployment, and separation from the third stage.

Design Qualification

The design qualification tests for the prototype-spacecraft system and subsystems had the purpose

of establishing a high degree of confidence that the spacecraft system and subsystems as designed would not be impaired by predicted environments, and thus would be suitable for required operations. Since this confidence had to be established by testing just one prototype, normal statistical sampling was not possible. Instead, test specimens were subjected to considerably greater rigors of environment than expected during ground-handling, launch, and flight.

Flight Acceptance

The acceptance tests for the flight spacecraft system and subsystems were designed to locate latent defects in material and workmanship, thereby providing assurance that none of the essential characteristics of design had been degraded during manufacturing and the accompanying inspection and handling. An additional objective was to demonstrate the compatibility of the subsystems and other elements of the spacecraft system under simulated launch and orbital environments. The levels of the flight acceptance tests approximated predicted environmental conditions. This choice of test levels is based on the philosophy that this degree of exposure allows detection of latent defects in material and workmanship without the unnecessary risk of damage from exposure above predicted levels. In other words, the spacecraft is tested through its early operating lifetime or infancy when its parts mortality is greatest according to past experience.

TEST PLANS

General

The Ariel I environmental test program was mainly accomplished at GSFC. A description and evaluation of the test results constitute the basic purposes of this report.

The test program consisted of mechanical tests at the design qualification level applied to the structural model of the spacecraft; design qualification tests for one spacecraft (the Prototype) as well as prototype models of the vibration experiment, separation system, and one prototype solar paddle; and flight acceptance tests for two spacecraft (Flight Units 1 and 2), the vibration experiment, and spacecraft separation systems.

Structural Model

Engineering Test Unit 1 (ETU 1), a model of the Ariel I spacecraft structure, was to be statically balanced and subjected to exposures of spin, acceleration, and vibration at design qualification levels to determine the balance and response of ETU 1 to these exposures.

ETU 1 was to contain no electronic equipment, but the intended flight assemblies were to be simulated in respect to weight and size by dummy weights placed in appropriate positions in the model.

Functional

Inertially correct mockups, including all essential hardware, of the Ariel I spacecraft and final-stage X-248 booster were to be sequentially subjected to spinup by PET rockets, yo-yo despin, experiment boom erection, solar paddle and mass boom deployment, and separation of spacecraft from the X-248.

Design Qualification

Balance

Static and dynamic balancing of the spacecraft is necessary to insure spin stability of the spacecraft during launch and orbital flight. It was specified for the prototype spacecraft so that this unit would be dynamically similar to the flight unit during acceleration and vibration testing. Balancing of the prototype was mainly for determining suitable locations for balance weights as well as determining that methods planned for balance of the flight units were suitable.

Spin

The prototype spacecraft was to be spun at 225 rpm for 1 min and at 150 rpm for 30 min to exceed the predicted 150 rpm experienced by the spacecraft third-stage combination from third-stage ignition. Operation of spacecraft was to be checked during spin.

Acceleration Test

The maximum acceleration (18g), imparted to the Ariel I spacecraft by the Delta launch vehicle, occurs just prior to third-stage burnout. The orientation of the spacecraft on the centrifuge was selected so as to simulate the sustained loading of this thrust-induced acceleration. In addition,



FIGURE 2.—Prototype setup prior to installation for acceleration.

transverse acceleration tests were specified based on expected handling loads of 2g. Figures 2 and 3 show the prototype before and after installation on the centrifuge.

Shock Test

A shock environment is produced in several ways—handling, shipment, stage ignition, and stage separation being the most common. The Ariel I shock test parameters were dictated by handling and transportation considerations, since the shock pulses generated by the Delta launch vehicle were expected to be less severe.

Temperature and Humidity Tests

This series of tests had the purpose of simulating the temperature and humidity conditions of

storage, transportation, and pre-launch periods to determine the effect of those conditions on the Ariel I spacecraft. The operational temperature tests are based on expected flight temperature and are conducted prior to thermal-vacuum tests.

The program was planned in five parts: (1) -30°C soak, (2) $+60^{\circ}\text{C}$ soak, (3) -10°C operational, (4) $+50^{\circ}\text{C}$ operational, and (5) humidity exposure. Parts 1 and 2 were to consist of 6-hour periods of exposure at each temperature with performance checks at ambient temperature (25°C) after each exposure. Parts 3 and 4 were to consist of monitoring spacecraft performance at the stabilized temperature of -10°C and $+50^{\circ}\text{C}$. Performance checks also were to be performed at ambient temperature after each exposure. Part 5 was to consist of a 24-hour exposure to 30°C and a relative humidity of 95 percent. At the end of the exposure period, spacecraft performance was to be checked. Spacecraft performance was to be monitored during the drying-out period. Figure 4 shows the setup for these tests.

Vibration

Vibratory excitation of the spacecraft arises from shipment and rocket motor burning, as well as from acoustic and aerodynamic sources. Random excitation and combustion resonance tests along the thrust axis and two lateral axes were to be conducted to simulate conditions anticipated in a launch by the Thor-Delta vehicle.

Vibration testing of the Ariel I spacecraft was complicated by the fact that there are several



Figure 3.—Prototype spacecraft installed on centrifuge for acceleration.

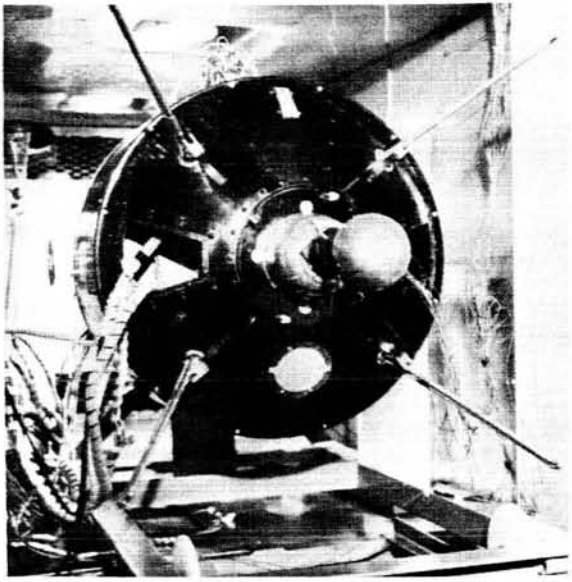


FIGURE 4.—Setup for prototype temperature and humidity test.

appendages (electron density boom, solar paddles, etc.) which, prior to separation, are folded down along the case of the last-stage X-248 rocket motor.

To simulate this situation, it was necessary to fabricate a cylindrical structure of the same nominal diameter as the rocket motor and of adequate length to accommodate the appendages, yet retaining the desirable stiffness properties of a good vibration fixture. The ideal vibration fixture is a massless, infinitely stiff device designed to adapt the test specimen to the shaker table. It was recognized that building such a fixture with no resonances within the applicable frequency band would be impossible. The approach taken was to build the fixture as stiff as practical, and then to check it out in vibration with a dummy load simulating the Ariel I system. When this was done, it was found that the setup was usable to about 1000 cps in the thrust direction and about 200 cps in the lateral direction.

Because of the frequency limitations imposed by this setup, each of the tests was divided into two parts. The lower frequency tests were to be performed with the complete system, including the appendages, in place of the cylindrical fixture (which was called the "tall fixture"). Higher frequency tests were to be performed without the

tall fixture and without appendages, using a simple flat plate (short fixture) to adapt the shaker table to the Dutchman interface. Figures 5 and 6 present the tall and short fixtures, respectively. It will be noted that in following this procedure the various booms and paddles were not to be tested above 1000 cps in the thrust axis or 200 cps in the lateral axes.

Additional subsystem tests of these appendages were planned, but were waived on the basis that the appendages were essentially decoupled from the vibration input at the high frequencies.

Thermal-Vacuum Tests

Background: These tests were to be conducted to simulate the extremes of temperature in a vacuum environment that the spacecraft was expected to encounter in orbit. Satisfactory operation of the spacecraft system at -8°C and $+35^{\circ}\text{C}$ with a vacuum of 1×10^{-5} mm Hg was to be verified.

During its orbital life of 1 year, the spacecraft will have several aspect positions relative to the

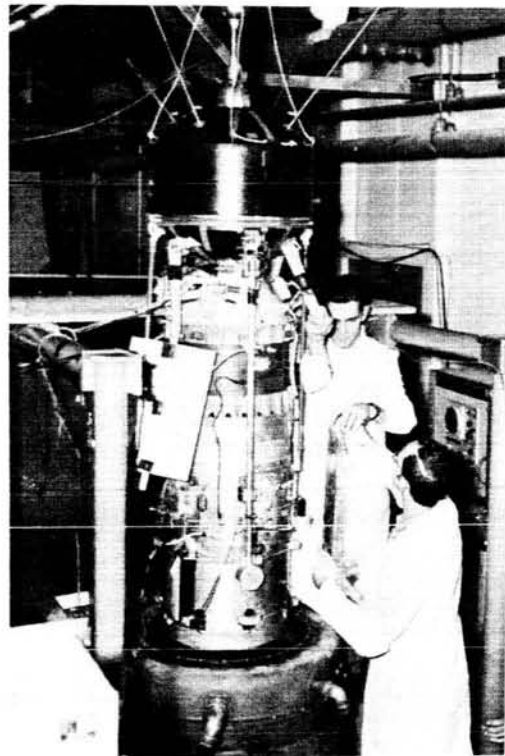


FIGURE 5.—Installing solar paddle prior to thrust axis test of prototype unit on tall fixture.

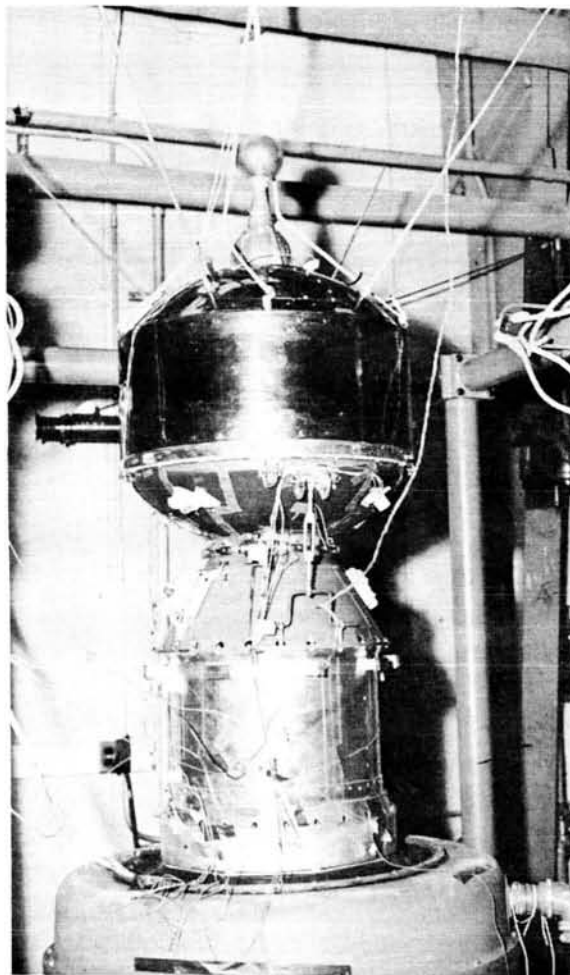


FIGURE 6.—Thrust axis setup using short fixture.

sun. This means that heating of one side occurs while the opposite side is being cooled; and, later in orbital life, the heated and cooled surfaces will be reversed. The test procedure was to cover the simultaneous heating-cooling type of environment as well as the conventional uniform hot or cold environment.

Description: The test was to consist of four parts: (1) cold, (2) hot, (3) 30° aspect, and (4) 135° aspect. Test 3 was to simulate the space environment where the sun is at an angle of 30 degrees with respect to the longitudinal spin axis of the spacecraft and illuminates the top portion of the system. Test 4 was to simulate the space environment where the sun is at an angle of 135 degrees with respect to the longitudinal axis of the spacecraft illuminating the bottom portion of

the system. (The projected area is a maximum at the 30° aspect and a minimum at 135°.) Tests 1 and 2 were to be conventional tests in which the entire spacecraft was to be soaked at a uniform temperature.

The external power supply was to be such that:

1. Spacecraft could be operated solely on external power.
2. Spacecraft could be operated solely on internal power.
3. Batteries could be charged during the test using the shunt regulator.

Flight Acceptance

Balance

Both static and dynamic balancing of the flight spacecraft were specified to insure spin stability of the spacecraft during launch and orbital flight. Figure 7 shows the Ariel I spacecraft on the balance machine.

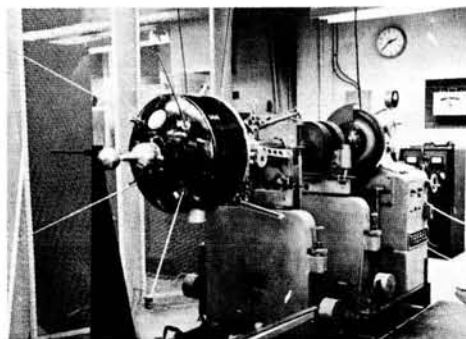


FIGURE 7.—Ariel I balance configuration.

Spin

The flight spacecraft were to be spun at 150 rpm to verify the spacecraft system operation at the rate of spin expected prior to third-stage ignition.

Vibration

Vibration levels no greater than those expected during launch and injection into orbit were to be applied in each of three orthogonal directions.

The flight spacecraft were to be operated in a duty cycle typical of that to be employed in actual launch and monitored for malfunctions in telemetering and other systems that operate during boost.

Thermal-Vacuum

Flight spacecraft were to be exposed to predicted vacuum and temperature. Predicted solar conditions were to be simulated. The operation of the flight spacecraft was to be monitored during above exposures.

TEST RESULTS

Structural Model

The structural model of the spacecraft was tested from April 7 through May 9, 1961. The test environments included static balance, spin, acceleration, and vibration. These tests were conducted at increasing levels of severity up to design qualification levels. Vibration equipment consisted of two 2250 kg (5000 lb) exciters; one was used for thrust axis excitation, and the other for lateral axis excitation.

The static unbalance of this unit was reduced from 1.5×10^7 dyne-cm (218 oz-in.) to 4.4×10^6 dyne-cm (61.8 oz-in.) by the addition of 0.68 kg (1.5 lb) of lead to the main shelf.

The main difficulties arose during vibration. During vibration in thrust axis at acceptance level, several screws holding the support tube to the main shelf sheared; this problem was solved by increasing the size and number of screws. The second failure of screws occurred during lateral vibration at acceptance level when several that fastened the struts to the main shelf failed; this was tentatively solved by increasing the number and size of screws at this location and assuring that the countersink for the screw heads was concentric with the through-hole. The third failure of screws occurred during vibration in the thrust axis of design qualification levels; in this case the screws fastening the struts to the support tube failed. The solution adopted was to add shear pins between the support tube and struts.

The design of subsequent units was changed to include a machined ring on the support tube and slots on the struts which, when assembled, would act as a shear pin. In addition, shear pins were used at the joint between the struts and the main shelf.

The structure successfully passed the design qualification level exposures after the above-mentioned modifications.

Response data for use in subsystem testing were taken during these vibration tests. However,

because of the change from the Scout to the Delta launch vehicle—requiring the addition of an adapter section (the Dutchman), the data were no longer valid.

Functional

The functional tests were conducted on inertially correct mockups of the Ariel I spacecraft and final-stage X-248 booster between October 16 and November 16, 1961. All tests were run under vacuum conditions in the 18.3 m (60 ft) sphere at Langley Research Center (LRC).

Results of the functional test program were generally satisfactory, and details are presented in Appendix E.

Prototype

Balance

Initial balance operations resulted in addition of a 0.454 kg (1 lb) balance weight, bringing the total weight of the spacecraft to 48.4 kg (106.6 lb). Because of difficulty in removing protective devices for spacecraft experiment sensors, balancing was conducted without their removal; but the method of attachment was changed so that covers could be removed for later balancing operations.

Spin

The spacecraft was spun at 225 rpm for 1 min and at 150 rpm for 30 min. During initial turn-on operations, the spacecraft went into under-voltage on insertion of the turn-on plug because of low battery voltage. This condition was corrected, and a satisfactory test was conducted.

Temperature and Humidity

General: The operational temperature test revealed problems in the electron density experiment, tape recorder, and cosmic ray experiment. However, it was decided not to repeat this exposure, since an operational temperature test under vacuum was to be conducted in the thermal-vacuum test program to follow. Furthermore, required replacement parts for the tape recorder and cosmic ray experiment could not be procured in a short time. The x-ray and mass spectrometer experiments showed some effect from high humidity.

Cosmic Ray Experiment: During the systems checkout and on completion of the $+60^\circ\text{C}$ exposure test, excessive counts were observed for the

cosmic ray experiment. As the temperature test progressed, degradation continued. This condition was due to a defective Geiger tube. The malfunction could not be corrected because replacement tubes were not available at that time, but the faulty tube was later replaced.

X-Ray Experiment: The x-ray experiment showed some effect on completion of the 24-hr exposure at 30°C and 95 percent relative humidity: The experiment was drawing excessive current. However, during the first 5 minutes of the drying period (RH=80 percent), the current returned to normal and its operation continued satisfactorily.

Mass Spectrometer Experiment: The effect of humidity was also observed in the output of the mass spectrometer experiment. Four test points were below minimum tolerance. However, during the first 10 minutes of the drying period (RH=42 percent), these readings returned to within the minimum tolerance.

Electron Density Experiment: During the initial -10°C operational test, improper operation of the electron density experiment was observed: The experiment was found to be gating at random. The output indicated maximum (6 to 7 volts) and could not be varied when the capacitance of the sensor was varied. At that time it was believed that the malfunction was a result of a temperature gradient within the electronics (this being due to the short stabilization period of 2 hr). Two additional thermocouples were installed within the electronics pack on completion of the initial test. During the final -10°C test, the experiment again failed to operate at low temperature. Operation was checked at 5-degree intervals from +5°C to -10°C. Improper operation appeared at +5°C and continued as the temperature was reduced to -10°C; no further attempt was made to correct the condition at that time. Operation was satisfactory at +50°C and room temperature. Since the experiment was not designed to operate under 0°C and the experimenter did not wish to redesign it, thermal coatings designed to prevent temperatures below 0°C were developed for the boom electronics can.

Tape Recorder: Also, during the initial -10°C test it was observed that the tape recorder was drawing excessive current and was operating at a slower than normal speed. This condition was

also experienced during the final low temperature exposure; at that time, improper operation was observed at -5°C and continued as the temperature was reduced to -10°C. No further action was taken to correct the condition at that time. Operation was satisfactory at +50°C and at room temperature.

The prototype tape recorder—which would not operate satisfactorily below -5°C—was replaced subsequent to this test by the tape recorder planned for Flight Unit 1, which had been modified since delivery of the prototype recorder. The later unit had been operated by the designer satisfactorily at -20°C.

Vibration

General: For the first prototype test, the spacecraft was installed atop a fixture that was hinged at the floor and driven at the fixture—Dutchman interface. By so doing, the input was angular motion rather than the desired true translational input. Consequently, components near the top of the spacecraft would receive a somewhat higher input than desired.

For the prototype retest and flight unit tests, all thrust axis tests were done on the 4500 kg (10,000 lb) vibration exciter. The lateral axes tests were performed on a hydrostatic bearing table driven by the same shaker. The hydrostatic bearing table represents an innovation in vibration testing—being essentially similar to the slippery tables or oil-film tables commonly used for this purpose, but much more effective in constraining lateral and vertical cross axis motion, as well as being able to withstand much higher bending moments without deleterious effects. The result is the closest approach to true unilateral motion believed to be attainable thus far.

Because of the large mass of the system it was not possible to get the required level of $\pm 86g$ for the prototype combustion resonance test with the shaker. Therefore, the shaker was driven to its maximum output, which was about $\pm 60g$ —slightly above the flight level. This represented the only deviation from the specified test program.

Although no structural failures occurred in the initial series of tests, conducted in October 1961, there were several discrepancies that were attributed in part to the severity of the fundamental thrust axis resonance.

Post Test 1: To reduce the severity of the thrust axis resonance, the Dutchman (extension section) was redesigned to provide greater stiffness and damping. All tests of the second series, conducted in February 1962, were completed satisfactorily with one exception: When main thrust axis resonance (85 cps) was reached, the yo-yo despin weights shifted considerably, shearing some of the lead from the weights. As a result, the despin weights on all units were changed from lead to brass.

Cosmic Ray Experiment:

Test 1: The photomultiplier tube failed.

Post Test 1: After exhaustive testing, the photomultiplier tube was replaced by a redesigned unit.

Electron Density Experiment:

Test 1: Operation was intermittent.

Post Test 1: The electron density grids were found to have been previously damaged and were replaced.

Thermal-Vacuum

General: For this test program, a 2.4×2.4 m (8×8 ft) thermal-vacuum chamber was used. It has a temperature range from -65°C to +100°C and can achieve an ultimate vacuum of 6×10^{-8} mm Hg.

The spacecraft was mounted on the test fixture, in the thermal-vacuum chamber, with the spin axis horizontal. Three solar paddle arms were attached (paddles not available) and extended, and the fourth arm was removed. The electron temperature and the electron density booms were attached to the structure in the folded position. The spacecraft was supported by a test fixture attached to one of the spacecraft struts and to the forward end of the spacecraft mid-skin structure. The protective strip coat was removed from the structure after completion of all instrumentation. Figure 8 indicates the method of mounting and the location of the two heat lamp rings that were utilized during the 30 degree and 135 degree solar aspect tests. Figure 9 shows the general arrangement of the spacecraft being prepared for thermal-vacuum testing.

Numerous difficulties extended this test program from a normal 2½ weeks to 9 weeks and caused the spacecraft to experience three cycles each of low and high temperature instead of the

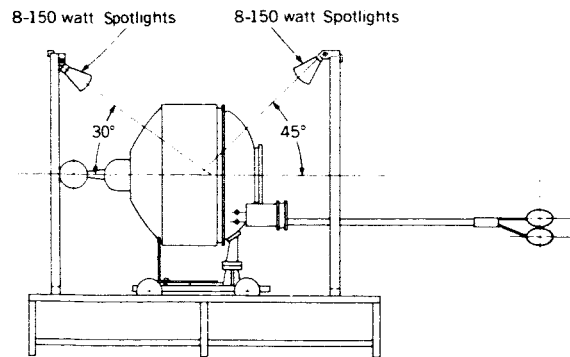


FIGURE 8.—Thermal-vacuum test setup.

planned program of one test of each type. Figure 10 shows the spacecraft checkout equipment used.

Major problems required the redesign of the x-ray and cosmic ray experiments, the tape recorder, and the shunt regulator. In addition, it was determined that the accuracy of the electron density experiment was reduced at low temperatures and that the U.K. converter required heat sinks. Also, there were a number of component problems in the spacecraft circuitry that were routinely resolved.

X-Ray Experiment:

Test 1: During initial chamber evacuation, the high voltage developed by the x-ray extra high tension (EHT) card failed at 70 mm Hg. Additional insulation applied to the high voltage lead connectors and sensors did not resolve the difficulty. Prior to the third cold test, the EHT card was replaced by a dummy.

Post Test 1: A transformer in the EHT card was replaced by a vacuum-impregnated high voltage unit; and one capacitor, operating above its rated voltage, was replaced with a new type.

Test 2: The x-ray experiment was successfully operated through the corona region. Arcing of the high voltage occurred at approximately 87 mm Hg, but operation returned to normal at approximately 5×10^{-2} mm Hg (50 microns).

Cosmic Ray Experiment:

Test 1: A large Geiger tube failed during the +55°C vacuum test. Since a suitable replacement could not be obtained in a reasonable length of time, this tube was permanently removed from the experiment. During the -10°C vacuum test one binary, in a chain of 15, was observed to be

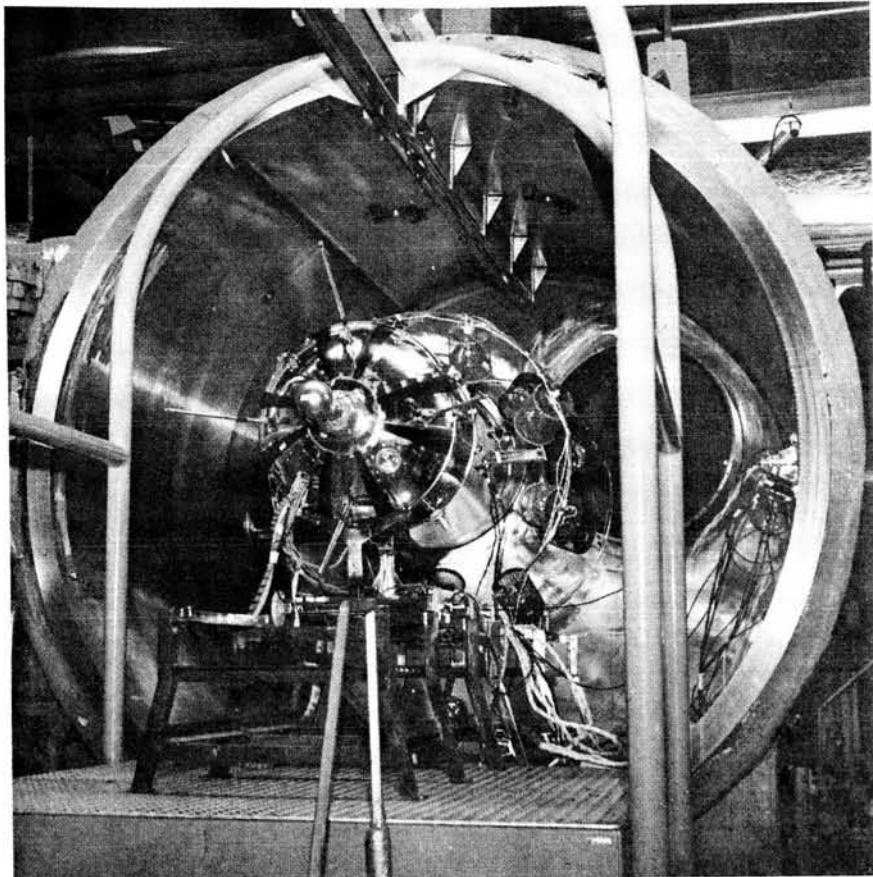


FIGURE 9.—Ariel I prototype system prepared for thermal-vacuum testing.

sensitive to low temperatures, resulting in excessive counts. Because of its location within the circuit no attempt was made to correct the problem; it did not recur on the flight units in the acceptance test program.

Electron Density Experiment:

Test 1: During the 30 degree solar aspect test, this experiment operated improperly — there being a single period of 138 μ sec in place of the normal two sawtooth pulses. This was caused by the low temperature to which the boom was exposed for this test.

Post Test 1: A thermal coating of evaporated gold and black paint was applied to the boom. This coating was designed to maintain the temperature of the boom electronics between 11° and 51°C during orbital flight, and thus avoid low temperature problems in this experiment.

Tape Recorder:

Test 1: The tape recorder failed during the first -10°C test because of a broken magnetic tape splice, shunted ground wire, and failure to preload motor bearings. Later, on return to -10°C , the tape recorder did not turn off when system was in undervoltage condition and drew excessive current during the $+47^{\circ}\text{C}$ test.

Post Test 1: The broken tape splice was mended, the short repaired, and the bearings preload adjusted. In addition a fail-safe power control circuit was designed to turn the tape recorder off when current exceeded 150 ma for 70 sec. This circuit also was to respond to the playback command to the tape recorder by switching power on to tape recorder. If the current continued to be excessive, it would again turn the tape recorder off.

Test 2: During the low temperature test, the tape recorder began to draw excessive current and

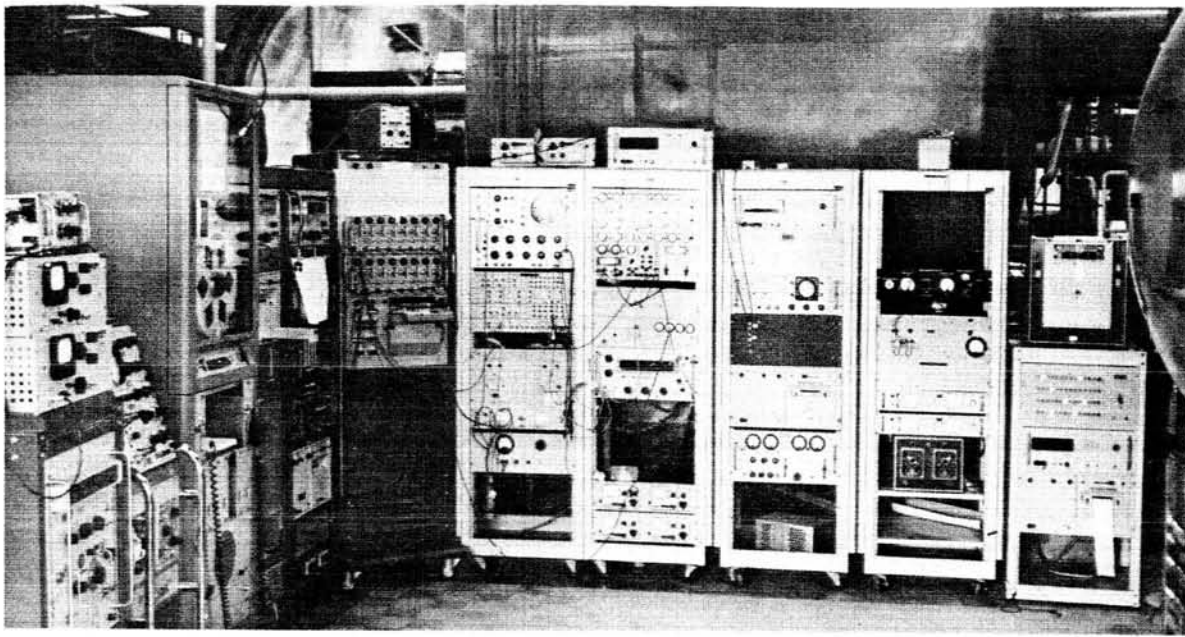


FIGURE 10.—Spacecraft checkout equipment used during thermal-vacuum testing.

was turned off by the fail-safe circuit installed after Test 1.

Post Test 2: Investigation revealed that a rib within the housing was causing the tape recorder to bind and thus draw excessive current. The housing was modified.

Test 3: The tape recorder operated satisfactorily after corrections of the following mechanical problems: loose drive belt, misaligned drift shaft, and broken drive belt.

U.K. Converter:

Test 1: A high-temperature condition occurred during the initial $+55^{\circ}\text{C}$ and was corrected by addition of a suitable heat-flow path from the card to the instrumentation shelf.

Shunt Regulator:

Test 1: Redesign was accomplished to electrically isolate the regulator from the structure when it was discovered in initial setup that the regulator was grounded to the structure. The regulator was also redesigned to prevent batteries from being charged at an excessive rate after it was found during the $+47^{\circ}\text{C}$ test that the regulator had no current-limiting capability, resulting in a high-temperature condition ($+72^{\circ}\text{C}$) in the batteries.

Post Test 1: Redesign resulted in the limitation of current and voltage at $+70^{\circ}\text{C}$ to 0.3 to 0.4 amp with 13.6 volts and at -10°C to 1.2 amp at 14.2 volts. Also, the battery pack was modified to improve the surface contact between it and the instrumentation shelf. These modifications were to reduce the temperature of the batteries to a tolerable level.

Special Test Programs

Solar Paddle: The prototype solar paddle, serial no. 1, was subjected to the following environments:

1. Temperature storage, -30°C and $+60^{\circ}\text{C}$, 6 hr each
2. Humidity, 95 percent relative humidity at 30°C for 24 hr
3. Vacuum, 5×10^{-4} mm Hg, ambient temperature
4. Temperature shock, from $+26^{\circ}\text{C}$ to -65°C in 20 min, to $+50^{\circ}\text{C}$ in 1 hr, to $+26^{\circ}\text{C}$ in 20 min; this cycle was repeated three times
5. Temperature shock, from -26° to 100°C , back to 26°C
6. Thermal-vacuum, -10°C and $+55^{\circ}\text{C}$ at 1×10^{-5} mm Hg

The paddle was checked for insulation resistance and diode leakage, and visually checked for changes in the surface. No problems were encountered.

Vibration Experiment:

Temperature and Humidity: The low- and high-temperature storage consisted of 6-hr exposures at -30°C and $+60^{\circ}\text{C}$, with a complete checkout of the system before and after each exposure. The humidity exposure consisted of 24 hr at 30°C and 95 percent relative humidity; the system was checked out before and after the test; no apparent humidity effects were observed. During the low-temperature test (0°C), it was found that the external power supply had to be reduced from 28 to 26 volts for the system to calibrate properly. Operation was normal during the high-temperature test, which was performed at $+40^{\circ}\text{C}$.

Vibration: The simulated battery box separated from its mounting brackets because the six spot welds were not strong enough. Subsequently, the angle brackets supporting the battery pack for the vibration experiment were improved by changing from spot welding to continuous welding.

Thermal-Vacuum: Operation of the system was normal during both the low- and high-temperature thermal-vacuum tests. The system was monitored through the corona region and showed no adverse effects. The low-temperature thermal-vacuum test was conducted at 0°C with the system operated on internal power for 15 min. The high-temperature vacuum test was conducted at $+50^{\circ}\text{C}$. The system was again operated for 15 min on internal power.

Separation System: This system was mounted in the normal configuration with the spacecraft and Dutchman for the second series of prototype spacecraft vibration tests. For the remainder of the second round of the prototype environmental test program, the separation system was separately tested.

No problems arose from the prototype test program except that, during vibration, the operation of the mechanical timers in two redundant subsystems was not satisfactory. No pair could be selected from the eight available that could match each other's time within the required 2 percent in

performing the necessary operational sequence. Therefore, electronic timers were designed and fabricated in time for the flight acceptance environmental test program.

Flight Unit 1

Balance

Balance operations added 0.844 kg (1.86 lb), bringing the total weight of the spacecraft to 49.56 kg (109.26 lb).

Spin

The spacecraft was spun at 150 rpm for 30 min. During this time unbalance readings were taken, and telemetry was transmitted and recorded on tape.

Vibration

No problems were encountered.

Thermal-Vacuum

Irregularities were noted, and action was taken as described below.

Cosmic Ray Experiment: This experiment would not turn on at -8°C ; but operation became normal at $+10^{\circ}\text{C}$, and it would operate at -8°C if the temperature were reduced with the experiment in operation. No action was taken.

X-Ray Experiment: The experiment was not gating properly because of a faulty diode that subsequently was replaced.

Electron Energy Experiments 1 and 2: Experiment No. 1 malfunctioned because of an open capacitor that was replaced. Memory cards for both experiments were modified to eliminate the problem of digitizing low-speed data.

Shunt Regulator: Transistors operated at excessive temperatures because of an inadequate heat sink. The method of mounting was modified to improve the conductive path to the solar paddle arms.

Electron Density Experiment: The operation of this experiment was questionable during the hot test, so a modified waveform generator card was installed.

Special Test Programs

Vibration Experiment: No problems were encountered.

Separation Assembly: There were no problems except that, during vibration, the two mechanical

timers required replacement by GSFC electronic timers, developed as a result of the prototype testing.

Flight Unit 2

Balance

Balancing operations added 0.69 kg (1.52 lb) to bring the final spacecraft weight to 49.4 kg (109 lb).

Spin

The spacecraft was spun for 30 min at 150 rpm. Operation was satisfactory.

Vibration

Cosmic Ray Experiment: There was a malfunction during the thrust axis random test because of failure in weld of the photomultiplier tube. To correct this situation, a spare cosmic ray unit was qualified and installed as a replacement for the defective unit.

Antenna: The antenna lost power because of a loosened micro-dot connector, which was tightened.

Mass Spectrometer: Faulty operation of this unit disclosed that one of the sphere connector pins had sheared, apparently because the mounting nut had not been sufficiently tightened.

Thermal-Vacuum

General: The following actions were taken prior to thermal-vacuum tests as a result of the behavior of Flight Unit 1 and the Prototype Unit in thermal-vacuum exposures:

1. The method of mounting the shunt regulator power transistors was modified to improve the conductive path to the solar paddle arms.
2. A modified waveform generator card was installed in the electron density experiment.
3. The tape recorder housing was modified, and the terminals were insulated.

Mass Spectrometer: This equipment malfunctioned because of a broken pin in the sphere. The sphere was replaced.

Tape Recorder: Operation of the dc control unit was questionable in checkout because of a non-qualified circuit design. The problem was resolved by installing a unit modified in accordance with prototype test results.

The tape recorder malfunctioned in the cold test because of a faulty idler bearing, which was replaced. Malfunction reoccurred after 10 hrs of hot test. After tests, the idler bearing and drive motor were replaced.

Electron Density Experiment: This experiment malfunctioned, but the malfunction could not be reproduced on a subsystem basis. Voltage spikes on several supply lines were determined to be the cause. (This problem did not occur in Flight Unit 1, which was flown.)

Special Test Programs

Solar Paddles: No problems were encountered.

Vibration Experiment: There was no Flight Unit 2 vibration experiment; no backup unit was available.

Separation System: It was again verified during vibration that the mechanical timers could not meet operational requirements and thus installation of the GSFC electronic timers was necessary. This was the only problem.

SUMMARY AND EVALUATION

Structural Model

The structural model test program revealed one major weakness in the structural design: Namely, the use of screws to support the shear load as well as the tension load was unacceptable. The addition of shear pins relieved the screws from support load and thus solved the problem.

It is interesting to note that the basic structure was subjected to approximately 1½ hr of vibration testing without sign of structural fatigue.

The required functional operations of the spacecraft were satisfactory.

Functional Tests

Satisfactory performance of the functional tests demonstrated that the functional and structural objectives of the spacecraft had, in general, been achieved.

The structural integrity of the spacecraft design was verified under functionally satisfactory tests of spinup, despin, experiment boom erection, solar paddle and mass boom deployment, and separation of the spacecraft from a simulated third stage.

A Lyman-alpha detector, which was in place during the spinup, showed no evidence of con-

tamination from burning of the spinup motors.

Spacecraft separation from the third stage was satisfactory, with no indication of tipoff when one or both explosive release devices were fired.

After additional tests at GSFC, but without vacuum simulation, the escapement mechanism for the experiment booms was modified so that full deployment of booms was achieved at the minimum anticipated flight spin rate.

After twelve tests and with use of heat-treated springs, the stretch yo-yo system's performance satisfied flight requirements as verified during the final series of three tests. For the final series the springs had been treated for a high yield point. Improved performance resulted, but a slight yield occurred at 25 percent over nominal spin rate, constituting a slight deviation from design objectives. This system was used for flight.

Prototype and Flight Spacecraft

Vibration

The first vibration test of the prototype spacecraft revealed that the resonant frequency of the structure was 80 cps. The resulting acceleration levels up to 83g caused the photomultiplier tube of the cosmic ray experiment to fail.

It was eventually determined that the Dutchman was adversely affecting the structure's frequency, and stiffening of the Dutchman flanges raised the natural frequency to 95 cps.

At the same time the U.K. and the U.S. developed improved versions of the photomultiplier tube. Retest showed both these types of tubes to be satisfactory and demonstrated that the entire spacecraft could withstand design qualification levels of vibration.

The general behavior of the structure in vibration was relatively simple, as shown in Figures 11 through 16. In the thrust axis, one significant mode was established in the vicinity of 85 to 95 cps, beyond which the response declined to a 1-to-1 ratio at about 180 cps, with the responses above 180 cps being generally below the input. The main exception to the rule of isolation above 180 cps was in the case of the main equipment shelf (Figure 5), which experienced two sharply tuned resonances at 280 and 960 cps. Figure 17 depicts the Ariel I system with accelerometer locations.

The degree of attenuation at the 550 to 650 cps resonant-burning region was sufficiently great to limit the levels to reasonable values.

In the lateral direction two significant resonances were encountered: one at 35 cps, and the

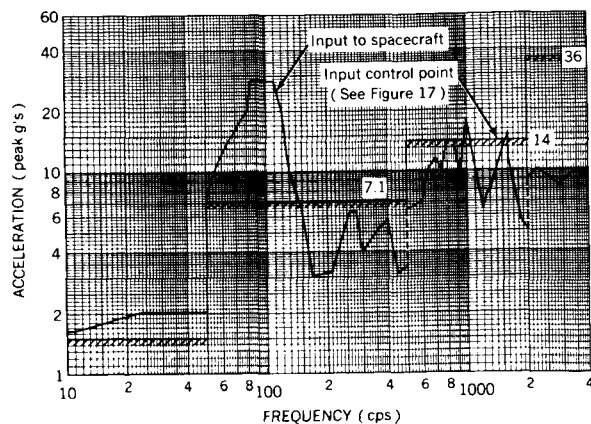


FIGURE 11.—Flight unit 1 thrust axis, input to spacecraft.

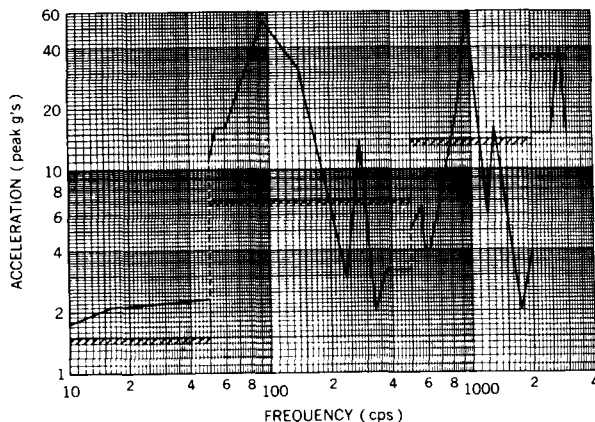


FIGURE 12.—Flight unit 1 thrust axis, equipment shelf.

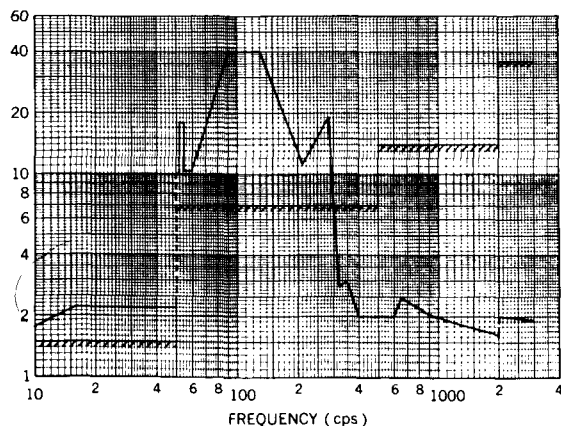


FIGURE 13.—Thrust axis, input to cosmic ray experiment.

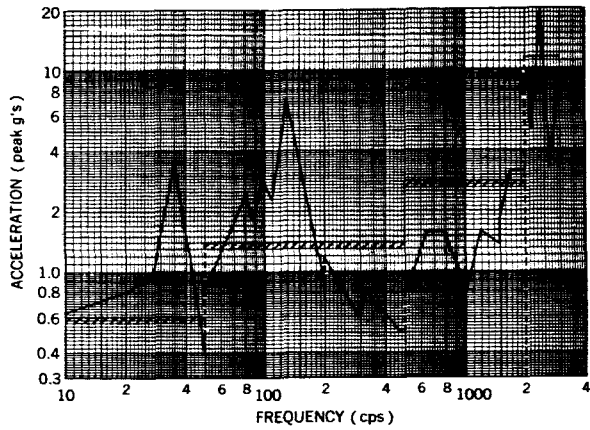


FIGURE 14.—Flight unit 1 lateral axis, input to spacecraft.

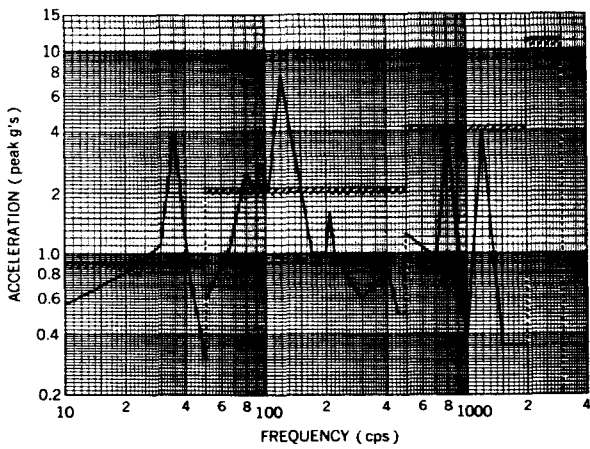


FIGURE 15.—Flight unit 1 lateral axis, equipment shelf.

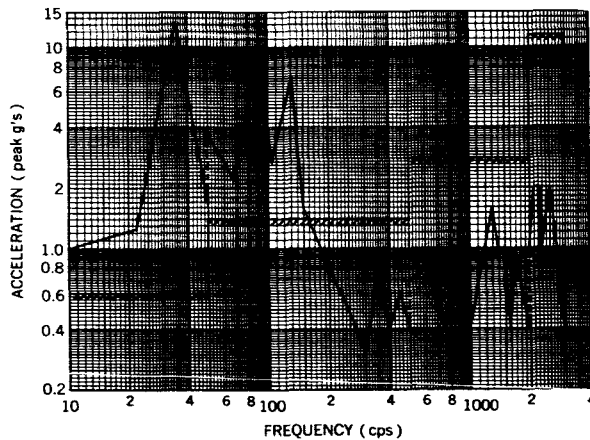


FIGURE 16.—Flight unit 1 lateral axis, input to cosmic ray experiment.

other at 120 cps. The former was so narrow in frequency as to be minor from a fatigue stand-

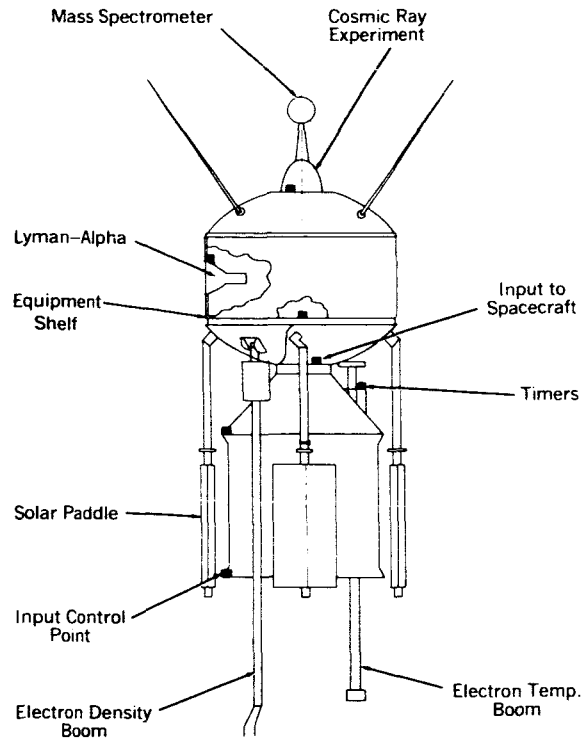


FIGURE 17.—Typical Ariel I test configuration showing accelerometer locations.

point. The latter, which may or may not have been a manifestation of coupling with the fundamental thrust axis mode, was more severe but, because of the lower input level, probably represented less damage potential than the thrust axis test for all except equipment with particular susceptibility to lateral vibration.

Shunt Regulator

While the prototype spacecraft was being prepared for the first thermal-vacuum test, it became evident that the shunt regulator case needed to be electrically isolated from the basic structure. Redesign solved this problem.

During the first thermal-vacuum test the power dump transistors burned out because the overall shunt regulator circuit could not dissipate the energy required for spacecraft operation. The circuit was redesigned.

During the second thermal-vacuum test, the batteries overcharged because the shunt regulator had no provision for current control. The redesigned circuit mentioned above also took care of this difficulty.

It was also observed that the power dump transistors were overheating during spacecraft operation. The trouble was traced to poor conductivity through the beryllium oxide washers into the solar paddle arms. Improved assembly technique solved this problem.

After solution of these design and assembly problems, the shunt regulators in the prototype and flight spacecraft performed satisfactorily in the required thermal-vacuum environments.

Tape Recorder

Problems and Solutions: Shorting of internal electrical leads to recorder case was solved by providing insulation between leads and case. A tape splice broke during testing; its weakness was caused by contact with tape lubricant at time of splice. A loose belt, causing failure in recording, was likewise caused by improper assembly techniques. Excess current drawn by the tape recorder resulted from seized bearings; the lubricant in the bearings was found to be unsuitable.

Modification: As a result of the numerous tape recorder failures, causing excess current demands on the spacecraft, a current-sensing relay that would shut off the recorder when excessive current was drawn was placed in the powerline. Thus the overall operation of the satellite would be protected from excessive power drain in case of improper tape recorder operation.

Experiments

Electron Density: This experiment would not operate in a -10°C environment because it had not been designed for this condition. The problem was solved by application of thermal coatings to electronic cans on flight spacecraft. This modification changed the environmental requirement to 0°C , where operation was satisfactory.

Cosmic Ray Experiment: The problem posed by the photomultiplier tube and its solution were discussed above under Vibration. During prototype thermal-vacuum, a Geiger counter tube failed; it was removed and was not replaced on flight spacecraft because it did not substantially affect experimental measurements.

X-Ray Experiment: The high voltage generator card failed during corona checks because of insufficient insulation. Additional insulation around the x-ray counters and high voltage leads at the counters and high voltage card was neces-

sary. Another failure of the high voltage generator card was traced to leaks in the transformer on the card. Replacement of all transformers with a vacuum-impregnated type resulted in satisfactory operation during flight spacecraft testing.

Wiring Harness

The wiring harness had numerous operational failures because of incorrect assembly of connectors and numerous unshielded monitoring leads; this resulted in false indications of experiment malfunctions during prototype testing.

Redesign resulted in an improved harness with fewer monitoring leads, and improved assembly techniques also helped produce a superior harness for the flight spacecraft.

System and Test Evaluations

System

The results of the Ariel I environmental test program indicate that the spacecraft system and experiments were generally well-designed and reasonably reliable with the exception of the shunt regulator and the tape recorder.

After redesign and application of improved assembly techniques, the reliability of the shunt regulator was demonstrated by satisfactory operation during the flight acceptance tests and the retest of the prototype.

The operation of the tape recorder was erratic during both prototype and flight acceptance tests. Therefore, a circuit was designed to remove the tape recorder load from the spacecraft electrical system if a failure occurred. This fail-safe circuit resulted in an overall increase in the spacecraft's reliability.

Several other difficulties were attributable to improper assembly rather than to design shortcomings, and were discovered and corrected as a result of the test program.

Test

The test results indicate a minimum of difficulty traceable to human error in testing and therefore strongly imply that the test procedures were effective in avoiding damage to the spacecraft from improper handling and operation. In a broader sense, the effectiveness of the test program is supported by the successful operation in space of the Ariel I satellite.

APPENDIX A

TECHNICAL DESCRIPTION OF ARIEL I

General

The basic configuration of Ariel I (Figures A1 and A2), which must fit the shroud of the Delta rocket's spacecraft compartment, is that of a short fat cylinder 10-11/16 in. long and 23 in. in diameter.

Each end of the cylinder has a spherical section with an inboard terminator circle 23 in. in diameter and a smaller outboard terminator circle

8-7/16 in. in diameter. These spherical sections are 5-7/16 in. in diameter and 5 1/4 in. long, with an outer surface radius of 13 1/2 in. To this basic configuration are attached the various appendages necessary to support and conduct the spacecraft experiments.

The spin axis of the satellite is the central axis of the cylinder; this is also considered as the vertical axis. At the bottom of the spacecraft is

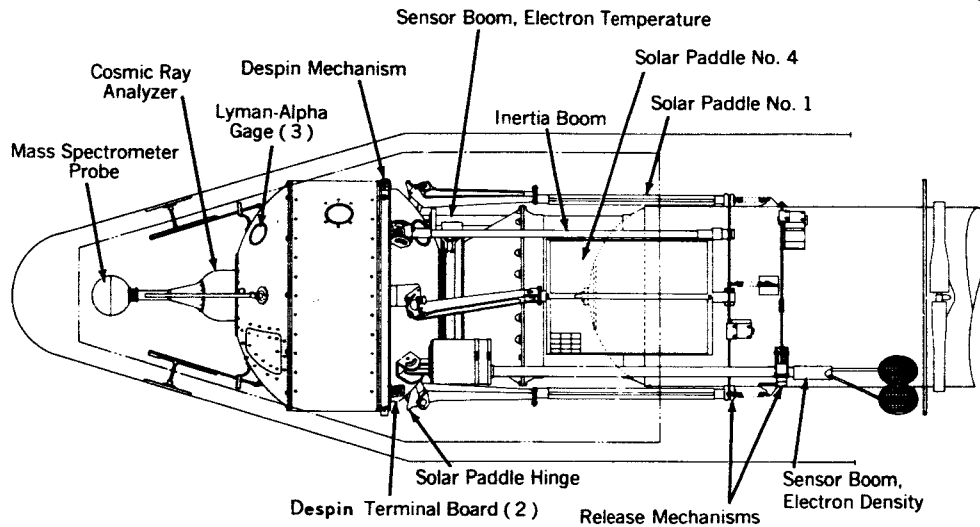


FIGURE A1.—Ariel I in launch configuration.

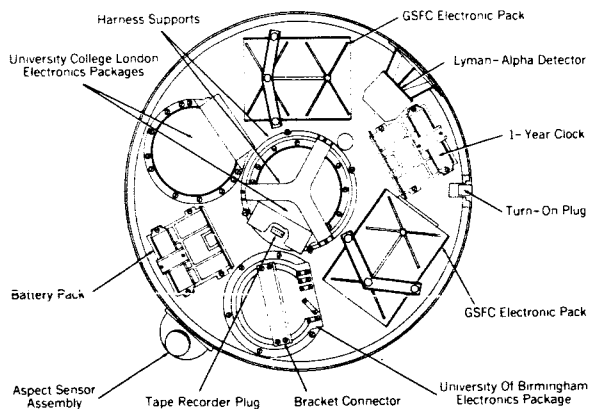


FIGURE A2.—Top view of inside positioning for Ariel I electronic modules.

a 9 3/8-in.-diameter third-stage separation flange, where an electron temperature gage and the tape recorder are installed.

Extending out horizontally from about midway up the lower spherical section, at 90-degree intervals around the circumference of the satellite, are four solar paddles.

Two 4-ft. booms are offset 45 degrees circumferentially from the solar paddles. They are opposite to, and exactly counterbalance, each other and extend radially in the same horizontal plane. The end of one boom holds two circular condenser plates of the electron density sensor. Electronics associated with this experiment are housed in a

4¼-in. diameter by 6¼-in. long cylinder mounted on the boom close to the spacecraft body. The end of the other boom holds a second electron temperature gage, whose electronics are located inside the spacecraft. A 3½-in. diameter hemispherical solar aspect sensor is located on the central cylinder section.

On top of the spacecraft in line with the spin axis is a 5-in. diameter cylinder containing the cosmic ray Cerenkov detector. Above this, on a 4-in. long conical section tapering from a 3-in. to a 1-in. diameter is a 4-in. diameter ion mass sphere.

Four antennas spaced circumferentially at 90 degrees and angling up at 45 degrees are mounted on the top spherical section. Three flush-mounted solar radiation (Lyman-alpha) gages are mounted on the satellite skin. Two proportional x-ray counters are located opposite the Lyman-alpha gages.

Structure and Mechanical Design

Basic structural materials in Ariel I are plastic-bonded fiber glass. The main or central body section is an epoxy-bonded monofilament-wound cylinder structure with a density of 0.071 lb/cu. in. The upper and lower spherical sections (domes) are molded of the same materials and have the same density. The upper dome is 1/16 in. thick; the lower dome, 1/32 in. thick. The upper dome is bonded to an aluminum ring that, in turn, mates with an aluminum ring bonded to the mid-skin. The lower dome is assembled in sections.

Internally at the top of the upper dome is a 0.2-in. thick aluminum disk, on which is centered an integrally machined aluminum cylinder 4 in. deep with a 7-in. inside diameter. Bonded to this central structure and the inside diameter of the upper dome are eight 1/16-in. thick mat-molded epoxybonded fiber glass stiffening ribs. These ribs have holes in them for weight reduction.

At the top of the lower dome and at the bottom of the mid-skin is the instrument shelf, 0.080 in. thick and machined from 6061-T6 aluminum; on the underside of this shelf are eight integral stiffen-ribs. A 7-in. diameter base ring enclosing the tape recorder and providing a structure for bolting the spacecraft to the separation flange is mounted below the shelf. Bolted to the base ring and to the instrument shelf are six aluminum

struts, also drilled out for weight reduction. Two struts support boom-mounted experiments, and the other four support the solar paddles.

The battery compartment is a simple mechanical container fitted with Teflon spacers. It is not pressurized but provides structural support to the sealed and pressurized nickel-cadmium batteries.

The flush-mounted x-ray sensors, light in weight, are bolted to the skin. Aluminum doublers mounted inside the skin at these points give additional stiffness. Lyman-alpha gages in the two domes are bolted to the skin and to brackets bolted to the central cylinder and the lower portion of the top shelf.

The Lyman-alpha sensor located on the cylindrical section is supported only by the skin, as is the aspect sensor. The top structure, including the Cerenkov sensor, is bolted to a plate on top of the central cylinder.

In the base of the spacecraft just below the tape recorder is an escapement device by which cabling is attached to the sensor booms to control their erection rate and timing. The "stretch yo-yo" despinn system consists of two steel springs wound ½ turn each around the bottom of the cylinder just above the instrument shelf. At the end of each spring is a relatively heavy weight.

Approximate weight of the structure (without instrumentation, batteries, etc.) is 35 lb. Weight of the four antennas is 0.8 lb.

Thermal Design

A variation of 35°C is expected within the satellite structure over a period of 1 year because of variations in the amount of time spent in sunlight, and shadowing effects. Skin temperature may vary from about 20° to 60°C, depending on sun-spin-axis angle. The variation in temperatures of boom-mounted components due to sun-spin-axis angle changes may be somewhat greater than the temperature for the main structure because of nonspherical geometry and greater shading effects. Solar paddle temperatures should remain between +33°C and -63°C, the temperatures corresponding to the hottest and coldest combinations of spin axis, orbital plane, and orbital position locations with respect to the sun.

Ariel I uses the thermal coatings of evaporated gold with about 25 percent of the total surface

area covered with a combination of black and white paints to achieve the desired absorptivity/emissivity ratio. This is a compromise between the experimenters' requirement for a conducting surface with a preference for a gold or rhodium over other metals, and a thermal requirement for maximizing the painted areas to minimize the tolerances of the absorptivity/emissivity ratio.

A special process was developed for applying the thermal coatings. Mirrorlike gold surfaces that should withstand heating to 250°F were obtained. First the surfaces were sanded, cleaned, and baked at 310°F for 1 hr. Then layers of varnish, lacquer, paint, and metals were applied in the following sequence and baked at the temperatures and for the time intervals indicated: (1) sealing varnish, 300°F for 20 min; (2) lacquer, 290°F for 30 min; (3) conducting silver paint, 280°F for 13 hr; (4) electroplated copper, 1.5 mils thickness; (5) lacquer, 275°F for 30 min; (6) evaporated gold; and (7) four coats of black paint in longitudinal stripes, 250°F for 30 min.

Telemetry

General

The radiating system on Ariel I is a slightly modified crossed dipole or turnstile array mounted on the upper part of the spacecraft. The antennas are mounted on the upper part of the spacecraft, on the upper dome, directly over the solar paddles in a canted or V-shaped turnstile configuration. The telemetry transmitted operates at a frequency of 136.410 Mc and will be used for both data transmission and as a signal source for tracking. The output power to the antenna system is 250 mw with an overall transmitter efficiency of 35 percent. The transmitter is designed for use with a phase-lock-type receiving system in the ground stations.

The command receiver is a single-channel type operating on the standard NASA command frequency, which is amplitude-modulated by an assigned subcarrier tone. The standby power consumption is approximately 50 mw, and the sensitivity is -100 dbm. Upon interrogation of the command receiver from a ground station, information stored in the flight tape recorder is transmitted back to the ground station.

Telemetry Encoder

The telemetry system is a Pulsed Frequency Modulation (PFM) system. This is a particular form of time-division multiplexing in which the information being telemetered is contained in the frequency of a sequential series of 10-millisecond pulses separated by 10-millisecond intervals. The pulse frequency is derived from a set of pulsed subcarrier oscillators, each having a frequency range of from 4 to 15 kc.

Two encoders, termed the high-speed and low-speed encoders, are used. The encoders will accept transducer outputs from various experiments, commutate them, and produce a frequency proportional to the value of the parameter being measured in each experiment. The output from the high-speed encoder will modulate the transmitter directly (real-time data), while the output from the low-speed encoder will be recorded on a tape recorder for a complete orbit at 1/48 the information rate of the high-speed system. On command, the tape output will be played back at 48 times the recorded speed. Thus the output from both systems, when received at a ground station, will have the same bandwidth.

Some 66 separate parameters will be telemetered from the experiments. In addition, there will be three sync outputs generated in the encoder for a total of 69 parameters being telemetered.

The high-speed encoder output consists of 256 channels arranged in 16 frames with each frame in turn consisting of 16 channels. Since the blank-burst interval for each "channel" takes 20 millisecond, a complete high-speed telemetry sequence of 256 channels will take 5.12 sec.

The low-speed encoder output into the tape recorder consists of 32 channels arranged in two frames with each frame in turn consisting of 16 channels. Since the burst-blank interval for each channel is 0.96 sec (e.g., 1/48 that of the high-speed system), a complete low-speed telemetry sequence of 32 channels will take 30.72 sec.

Recorder

General

The spacecraft recorder is designed to store encoded data for time periods up to 100 min. Input to the recorder is from a low-speed encoder

that furnishes data signals at 1/48th of the rate at which data are transmitted in real time. Recorder playback, initiated on receipt of ground-station command, will be at 48 times record speed — thus, the transmitted information rate will be identical for recorded data and real-time data.

The recorder contains 150 ft of special lubricated tape in an endless-loop configuration. Upon command, a 2-sec burst of 321 cps will be made to provide a distinctive time-reckoning mark (this burst will also be transmitted to signify successful command initiation). After the 2-sec interval, the recording ceases and the recorder will play back for 2.1 min, and then commence recording again.

Weighing $2\frac{1}{2}$ lb, the recorder is 7 in. in diameter and 3 in. high. Total power input required is 0.7 watt.

Tape Recorder Programmer

This programmer provides command means by which the tape recorder may be placed in the playback mode to allow extraction of orbital data. A precision reference frequency source is also an integral part of the programmer.

One-Year Timer

Two long-term electrolytic timers are provided in a redundant system to permit radio silencing of the satellite, nominally 1 year after launch.

The time interval is set by the selection of current allowed to flow in each timer. The switch on each timer is normally closed and connected in parallel so that the last timer to actuate determines the timing period and turns off the transmitter.

Power

Power System

The main elements of the power system are:

1. Four solar paddles
2. Shunt regulator and battery charge current limiter
3. Battery switching network
4. Batteries
5. Undervoltage system
 - (a) Undervoltage sensing circuit
 - (b) Timer for shutdown of satellite for 18 hr for battery recharging

The solar paddle outputs are combined and connected to a voltage regulator that is in parallel with the solar cells and the remaining power system.

The solar cells charge the batteries and supply all operating power to the spacecraft during the time the spacecraft is in the sunlight.

The regulator limits the system voltage and the battery charge current from the solar cells to values below which battery internal pressures rise as a result of generation of hydrogen.

The battery switching network is in series with both batteries, and the combination is in parallel with the solar cell paddles and shunt regulator.

The battery switching network connects either Battery A or Battery B to the load.

The hold-off relay is connected in the main powerline in series with the spacecraft subsystems. It will provide remote control of satellite turn-on.

The undervoltage system is connected in the main powerline in series with the spacecraft subsystems. It will turn off the satellite if the load voltage of Battery A and the load voltage of Battery B go below a predetermined value. During the shutdown period, all power output of the solar cells is connected to the battery switching network that charges the battery on line and trickle-charges the other battery.

Power Supply

The power supply for Ariel I consists of photoelectric converters; these generate electrical energy used to charge electrochemical storage batteries.

The solar converters are four fixed-position paddles, measuring approximately 1 ft by $1\frac{1}{4}$ ft, having solar cells covering both sides. The planned output of the total array is slightly in excess of 10 watts at 15 volts. In order to prevent overcharging of the storage batteries, a zener shunt regulator and battery charging current limiter was designed by GSFC to limit the voltage to 14 ± 2 volts.

The nickel-cadmium storage battery is a 6-amp hr size. Two 10-cell batteries are employed, one on standby service; the nominal voltage is 12 volts. GSFC designed an electronic circuit that should automatically keep the load on the battery in the highest state of charge and best operating condition.

Because of the different voltages required by the various experiments within the spacecraft, four dc to dc transistorized electrical converters are used. Three of these converters were designed and manufactured at GSFC.

Separation System

For reliability, the separation system is composed of two completely independent and isolated systems with either one having the capability for performing all required functions; see Figure A3.

The major components of both systems are mounted on a hollow frustum. Circuitry extends through the flyaway connector at the spacecraft separation system interface into the spacecraft for release of the despin mechanism. In addition, circuitry extends down along the Dutchman and the third-stage rocket motor to the experimental boom and initial boom release mechanisms.

Each independent system consists of the fol-

lowing elements: (1) battery, eight silver-zinc cells, $\frac{1}{2}$ amp hr; (2) timer, electronic; (3) guillotine; (4) pin puller; and (5) separation nuts, powder cartridge.

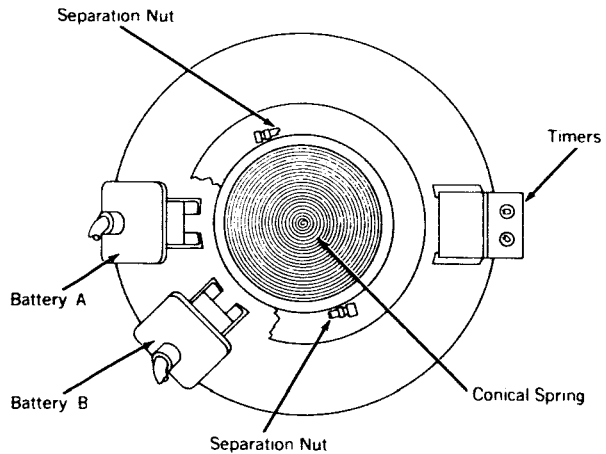


FIGURE A3.—Top view of separation system.

APPENDIX B

TEST CHRONOLOGY

Structural Model

Balance Operations

Time: April 27, 1961.

Weights: Initial unbalance— 1.5×10^7 dyne-cm (218 oz-in.)
Addition—0.68 kg (1.5 lb)
Residual unbalance— 4.4×10^6 dyne-cm (61.8 oz-in.)

Vibration

Time: May 2, 1961.

Discrepancy: Screws sheared in several locations.

Acceleration

No problems.

Final Vibration

No problems.

Spin

No problems.

Prototype Spacecraft

Balance Operations

Time: September 15–19, 1961.

Weights: Initial spacecraft weight—105.6 lb*
Balance weight added—1.0 lb
Final spacecraft weight—106.6 lb
Spacecraft Initial Unbalance: Static—176 oz-in.
Dynamic—448 oz-in.²

Residual Unbalance: Static—5 oz-in.
Dynamic—25 oz-in.²

The battery switching network was not installed until the completion of the balance operations. The spacecraft unbalance on completion of the balance operations was calculated to be:

Static—82 oz-in.
Dynamic—273 oz-in.²

Results: The planes selected for balance weight additions were found to be suitable.

Spin Test

Time: September 19, 1961.

Test Conditions: 225 rpm for 1 min.
150 rpm for 30 min.

The specification was not followed exactly because the balancing machine could not be operated below 150 rpm.

*Not including the weight of the booms and paddles (weight as balanced).

Difficulty:

Problem—Spacecraft would go into undervoltage upon insertion of turn-on plug.
Cause—Low battery voltage.

Temperature and Humidity Tests

Instrumentation and Test Setup: September 20–22, 1961.

Transmitter and receiver leads were interchanged. Investigation showed that the two leads were mislabeled in the spacecraft harness (Sept. 22, 1961).

Satisfactory checkout at +25°C (Sept. 22, 1961).

Cold-Storage Temperature Test: September 22, 1961.

Spacecraft subjected to -30°C storage test (nonoperating) for 6 hr.

Ambient Temperature Checkout: September 23, 1961.

Satisfactory spacecraft checkout at +25°C after -30°C storage test.

Ground Station Problem: September 23, 1961.

Sanborn recorder was functioning unsatisfactorily. Also, the oscilloscope used as additional aid in evaluation was not operating properly.

Hot-Storage Temperature Test: September 23, 1961.

Spacecraft subjected to +60°C storage test (nonoperating) for 6 hr.

Ambient Temperature Checkout Following +60°C test:

Connector problem on the prime converter. Smaller connector standoffs were required for proper connection. Also, additional thermocouples were installed within the electronic modules. (Sept. 24, 1961).

Faulty connector for transmitter lead in spacecraft harness was replaced (Sept. 25, 1961).

Satisfactory checkout at +25°C completed September 25, 1961.

Cold-Operation Temperature Tests:

The spacecraft was stabilized at +5°C nonoperating. The spacecraft was then operated and a satisfactory checkout completed (Sept. 25, 1961).

The spacecraft, while nonoperative, was then stabilized at -10°C. The spacecraft was operated following stabilization, and the following problems were observed:

Electron density experiment was gating at random (improper). The output indicated maximum (6–7 volts) and could not be varied when the capacitance of the exciter was varied.

Optical aspect readout not correct. Later investigation showed this to be caused by insufficient closing time of microswitches in the exciter electronics (readjusted at a later date).

X-ray experiment showed erratic counting of the reference number. No final solution as of October 1, 1961.

Cosmic ray experiment count was improper.

Tape recorder was operating but at a slower rate. Also, it was drawing excessive current.

Ambient checkout at conclusion of the cold-operation temperature tests, September 25, 1961.

Satisfactory checkout at ambient temperature. All experiments recovered. Therefore, no repairs were made.

Hot-Operation Temperature Test:

The spacecraft, while nonoperative, was stabilized at +50°C. The spacecraft was operated following stabilization, and the following problem was observed:

Cosmic ray count was not proper. A defective Geiger tube was suspected.

Ambient Checkout Following 50°C Test: September 28, 1961.

Cosmic ray problem as noted in hot-operation test.

The decision was made to continue with the repeat of the cold-operation temperature test with this problem present.

Cold-Operation Temperature Test Return: September 28, 1961.

Spacecraft temperature stabilized at +5°C prior to checkout. The following problems were noted:

The tape recorder was drawing excessive current. Electron density operation was not normal. Cosmic ray degradation continued. Temperature chamber temperature was set at 0°C, and the temperature of the monitored points of the satellite was allowed to stabilize with the satellite operating.

Bulb in optical aspect exciter burned out. Tape recorder and electron density problems continued. Cosmic ray degradation continued.

The chamber temperature was lowered to -5°C with the spacecraft operating during the temperature change. All problems in cold-operation temperature test rerun continued. No new problems developed.

The chamber temperature was lowered to -10°C with the spacecraft operating. No new problems developed.

Ambient Checkout Following -10°C test: September 28, 1961.

Satisfactory checkout prior to humidity. All experiments recovered from -10°C except cosmic ray (defective Geiger tube).

Humidity Test:

Spacecraft exposed to 95 percent RH at 30°C for 24 hrs and then checked out (Sept. 29, 1961).

X-ray showed excessive current at 95 percent RH. Recovered when chamber reached 80 percent RH (5 min).

Several test points were out of tolerance until 42 percent RH was reached by chamber (10 min). Then, these parameters reached the values of minimum acceptable tolerance.

Ambient Temperature Checkout Following Humidity Test: September 29, 1961.

Satisfactory checkout except for cosmic ray, which had a defective Geiger tube.

Vibration Tests

Time: October 10-13, 1961.

Thrust Axis Test (Tall Fixture): October 10, 1961.

Sinusoidal Sweep Frequency Test (10-250 cps) -

1. Electron density disk broke off during the 83 cps resonance:

Cause: Previously damaged during handling.

Solution: An epoxy cement was used to attach disk for remainder of test (immediate).

New disks will be fabricated and attached prior to re-test (final).

2. Tape recorder would not operate during checkout at end of test:

Cause: Blown fuse (in spacecraft).

Solution: Fuse replaced by jumper wire. (Tape recorder was not operating during tall fixture axis test.)

3. Cosmic ray count rate shifted at 83 cps structural resonance:

Cause: Failure in photomultiplier tube.

Solution: (a) Reduce amplification in system that occurs during resonance.

(b) Redesign tube to withstand higher acceleration levels.

(c) Final solution has not been proven as of January 16, 1962.

Random Test (psd = 0.07 g²/cps; frequency band, 20-250 cps) -

No problems noted.

Thrust Axis Test (Short Fixture): October 11, 1961.

Sinusoidal Sweep Frequency Test (250-3000 cps) -

No problems encountered.

Random Test (psd = 0.07 g²/cps; frequency band, 250-2000 cps) -

No problems encountered.

Transverse Test, X'-X' Axis (Tall Fixture): October 12, 1961.

Sinusoidal Sweep Frequency (10-150 cps) -

Intermittent operation of the electron density experiment.

Random Test (psd = 0.07 g²/cps; frequency band, 20-150 cps) -

Electron density problem continued.

Resonance Test (sinusoidal sweep, 550-650 cps) -

Electron density problem continued.

Transverse Test, Y'-Y' Axis (Tall Fixture): October 12, 1961.

Sinusoidal Sweep Frequency (10-150 cps) -

Electron density problem continued.

Random Test (psd = 0.07 g²/cps; frequency band, 20-150 cps) -

Electron density problem continued.

Vibration experiment simulated battery box separated from its mounting brackets.

Cause: Failure of six spot welds used to attach battery box to brackets.

Solution: (a) A continuous weld was used for attaching box to brackets.

(b) Additional bolts were employed to attach the battery box to the Dutchman.

Antenna failure.

Cause: Antenna was previously damaged.

Solution: Antenna replaced.

Transverse Test, Y'-Y' Axis (Short Fixture): October 13, 1961.

Sinusoidal Sweep Frequency (150-2000 cps) -

No problems encountered.

Random Test (psd = 0.07 g²/cps; frequency band, 150-200 cps) -

No problems encountered.

Resonance Test (sinusoidal sweep, 550-650 cps) -

No problems encountered.

Transverse Test, X'-X' Axis (Short Fixture): October 13, 1961.

Sinusoidal Sweep Frequency (150-2000 cps) -

Electron density problem continued.

Resonance Test (sinusoidal sweep, 550-650 cps) -

Electron density problem continued.

Random Test (psd = 0.07 g²/cps; frequency band, 150-2000 cps) -

Tape recorder played back without initiation command during 600 cps resonance test.

Cause: Test repeated twice to determine whether any interactions had inducted playback - none were found.

(b) Command frequency and tone for playback of Ariel I (S-51) and the Atmospheric Structures Satellite (S-6) were found to be the same (playback could not be traced to S-6 operation).

Solution: Ariel I command playback tone changed.

Post Vibration Test Checkout: October 14, 1961.

Mass spectrometer inner sphere unscrewed, causing several electrical contacts to shear.

Cause: Lock-tite had not been applied.

Solution: Closer pre-test inspection required.

Blown fuses in x-ray and aspect - 6.5 volt lines and mass spectrometer 24 volt line.

Cause: Suspect switching transients when system is turned off and on.

Solution: Remove all fuses from spacecraft.

Acceleration

On October 17, 1961, thrust axis acceleration test of 25g was conducted at NRL's Chesapeake Bay Annex.

Problem at Test Site:

Tape recorder would not play back on RF command.

Cause: Defective switch on command control panel (general support equipment)

Solution: Replaced switch.

Checkout After Test:

Spacecraft performance data were magnetically recorded before, during, and after the test and were returned to GSFC for evaluation on October 18, 1961. Since the spacecraft had not been repaired after the vibration test, the same problems noted were observed to have been present before and after the test. The following additional problems were noted:

Wiring harness—excess loading on -2.7 volt line.

Cause: Two loads on one supply line.

Solution: Separation lines installed in harness to supply each load.

Undervoltage detector circuit—initiation of both oscillators did not always occur (redundant system).

Cause: Design problem.

Solution: The wiring circuit was changed.

Shunt regulator box was installed but not connected to circuit, since power transistors were not available.

Command receiver—length of modulation burst became excessive (5 sec)—normal, 2 sec.

Cause: Defective diode.

Solution: Replaced.

Cosmic ray operation not proper.

Cause: Defective Geiger tube (large tube); first noted at +55°C temperature test.

Solution: Removed tube from circuit.

Leaky tape recorder.

Cause: Surface damage produced poor seal.

Solution: Replaced after shock test.

Electron density operation improper.

Cause: Malfunction of sawtooth sweep due to simultaneous gating-on of three power supplies and broken lead.

Solution: Design was modified so that -6.5 volt supply would gate-on before the other two supplies, and lead was repaired.

Shock

On October 25, 1961, the shock test ($\frac{1}{2}$ sine pulse of 30g peak amplitude and 11 millisecond duration) was conducted at the Naval Research Laboratory.

Checkout After Test:

Spacecraft performance data were magnetically recorded before, during, and after the test; and the tape was returned to GSFC for evaluation.

Electron temperature experiment not operating properly.

Cause: Defective twinax connector.

Solution: Connector replaced.

Cosmic ray output not correct.

Cause: Defective coaxial connector between experiment and encoder.

Solution: Connector replaced.

Electron density experiment not functioning properly.

Cause: Twinax connector defective (probably quality control problem).

Solution: Connector replaced.

Thermal-Vacuum

Setup: November 1-5, 1961.

Ground Station (Nov. 1, 1961).

Cause: Sanborn recorder had no center reference to ground.

Solution: Corrected

Shunt regulator circuit not operating properly (Nov. 1, 1961).

Cause: Regulator was mounted to the structure (system ground). Structure was being used as a heat sink for the regulator.

Solution: New regulator temporarily electrically installed. New design will be electrically isolated from the system.

Lyman-alpha sensor would not respond when exciter lamp was utilized (Nov. 2, 1961).

Cause: Sensor shorted to ground.

Solution: Repaired.

Electron temperature probe No. 1 would not sweep (Nov. 2, 1961).

Cause: Faulty micro-dot cable at connector from electronics to sensor.

Solution: Replaced cable.

Encoder clock would not give period pulse on low speed (Nov. 2, 1961).

Cause: Sensitive to position.

Solution: Replaced with spare clock. Also, low-speed and high-speed encoders updated at this time.

Ground Station (Nov. 2, 1961).

Cause: Decoder not proper.

Solution: Repaired.

Spare encoder clock not operating properly (Nov. 3, 1961).

Cause: Capacitor missing from the card.

Solution: Replaced encoder clock with a pre-prototype clock.

Chamber Evacuation: November 5, 1961.

Spacecraft was operated and monitored until chamber pressure of 1×10^{-4} mm Hg was achieved.

X-ray showed rapid count at 70 mm Hg (Nov. 5, 1961).

Cause: Arcing of -1600 volts to sensors.

Solution: Test continued. Experiment operated properly at 2×10^{-4} mm Hg. (Special tests performed at later date.) EHT Card No. 503.

Downtime: November 5-9, 1961; for trouble-shooting of x-ray.

Setup: November 9, 1961.

Faulty thermistor in UCL-2.

Cause: Open lead in monitor board.

Solution: Corrected November 11, 1961.

Chamber Evacuation: November 9, 1961.

Operation of the spacecraft appeared normal except as noted in "Setup" section. Also, the x-ray experiment was not operated during the chamber evacuation to 1×10^{-4} mm Hg.

Stabilization of the Spacecraft (not operating) at -10°C : November 9, 1961.

Low-speed oscillator malfunctioned.

Cause: Shorted gate transistor. This condition occurred during the initial chamber evacuation but was not noted until the low-speed data were reduced at a later date.

Solution: Corrected on November 11, 1961.

Downtime: November 9-11, 1961; for correction of problem noted above.

Setup: November 12, 1961.

Low-speed encoder malfunctioned (Nov. 12, 1961).

Cause: Open lead in printed circuit board of the low-speed matrix.

Solution: Repaired.

Cosmic ray would not gate properly (Nov. 13, 1961).

Cause: Three broken coaxial (miniature) leads.

Solution: Repaired.

Chamber Evacuation: November 13, 1961.

Spacecraft operated during pumpdown until chamber pressure of 1×10^{-4} mm Hg was achieved. Operation appeared normal; however, x-ray experiment was not operated during evacuation.

-10°C *Vacuum Soak:* November 13-21, 1961.

X-ray experiment would not turn on (Nov. 15, 1961).

Cause: A temperature sensitive transistor within the EHT generator card was suspected.

Solution: Experiment operated after 1 hr, 20 min warmup period.

Cosmic ray not operating properly (Nov. 15, 1961).

Cause: Temperature-sensitive binary counter.

Solution: No action; cleared at -2.0°C .

X-ray experiment malfunctioned (Nov. 16, 1961).

Cause: Failure of -1600 volts to sensors.

Solution: Test continued. Additional potting applied at connector of -1600 volt line and sensor potted. Also, a vacuum-impregnated high voltage transformer was installed.

Shunt regulator would not operate (Nov. 17, 1961).

Cause: Dumping transistors burned out.

Solution: Redesign of circuit required.

$+55^{\circ}\text{C}$ *Vacuum Soak:* November 21-25, 1961.

Problems as indicated in preceding three paragraphs continued.

Tape recorder malfunctioned (Nov. 23, 1961).

Cause: (a) Broken splice in tape.

(b) Burned ground wire.

(c) Motor bearing not pre-cooled.

(e) Pressure switch wired backward.

Solution: Repaired.

Cosmic ray not operating properly (Nov. 23, 1961).

Cause: Faulty Geiger tube (big).

Solution: Cut out of circuit, as no replacement was available.

Lyman-alpha not operating properly (Nov. 23, 1961).

Cause: Diode open.

Solution: Repaired.

Optical aspect drawing excessive current (Nov. 23, 1961).

Cause: Shorted transistor.

Solution: Repaired.

U.K. converter—excessive temperature (Nov. 24, 1961).

Cause: Inadequate heat sink.

Solution: Improve heat sink.

Downtime: November 25 to December 6, 1961; for correction of problems November 12 on.

Setup: December 6, 1961.

The following modifications were accomplished:

1. Undervoltage changed to 18 hr at 11.1 volts from 24 hr at 10.5 volts.
2. Shunt regulator redesigned.
3. One-year timers installed.

X-ray drawing excessive current, no high voltage at the sensors (Dec. 7, 1961).

Cause: Faulty card.

Solution: Replaced.

Replaced pre-prototype encoder clock by first encoder clock, which had been repaired (Dec. 8, 1961).

Chamber Evacuation: December 9, 1961.

X-ray experiment malfunctioned (Dec. 9, 1961).

Cause: Arcing of -1600 volts at 7 mm Hg.

Solution: Removed EHT card from system and installed dummy card: New card available approximately February 1, 1962; x-ray experiment will be qualified at this time.

Cosmic ray not operating properly (Dec. 9, 1961).

Cause: Ground station.

Solution: Repaired.

+55°C Vacuum Soak: December 10-13, 1961.

Cosmic ray not operating properly (Dec. 10, 1961).

Cause: Ground station.

Solution: Repaired.

Optical aspect drawing excessive current (Dec. 10, 1961).

Cause: Faulty capacitor in card.

Solution: Repaired.

U.K. converter temperature excessive -65°C (Dec. 10, 1961).

Cause: Heat sink not properly installed.

Solution: Installation corrected on December 21, 1961.

Tape recorder pressure indicated low (Dec. 10, 1961).

Cause: Faulty pressure switch.

Solution: Use of pressure switch abandoned.

Undervoltage detector turned spacecraft off at 11.8 volts when undervoltage was set at 11.1 volts (Dec. 11, 1961).

Cause: Suspect that undervoltage detector was not properly temperature-compensated.

Solution: Power system and undervoltage detector presently being investigated for possible design change.

Lyman-alpha operation not proper (Dec. 11, 1961).

Cause: Potting found in right-angle connector at the Lyman-alpha sensor.

Solution: Replaced connector.

Battery life questionable (Dec. 11, 1961).

Cause: Charging rate excessive.

Solution: Possible redesign of power system may be required.

Shunt regular allowing charging of batteries at excessive rate (Dec. 13, 1961).

Cause: Suspected design problem in shunt regulator.

Solution: Replaced shunt regulator by a redesigned unit.

Electron density (SEL) data store showed no erasure (Dec. 13, 1961).

Cause: Unknown.

Solution: Special temperature test showed operation normal.

Downtime: December 14-22, 1961; for correction of items occurring in +55°C vacuum soak.

Setup: December 22, 1961.

Electron temperature not operating properly (Dec. 22, 1961).

Cause: Faulty connector at dummy load.

Solution: Replaced.

-10°C Vacuum Soak: December 26-29, 1961.

Electron density not operating properly. Large shift in null.

Cause: Sensitive to low temperature.

Solution: (a) Heat lamps on boom electronics were used to increase temperature. Operation became normal at approximately 0°C.

(b) Study is being conducted to determine whether this item should have a thermal coating applied.

Tape recorder would not turn off during undervoltage (Dec. 27, 1961).

Cause: Test panel wiring.

Solution: No action.

Downtime: December 29 to January 2, 1962; test suspended because of holiday.

+47°C Vacuum Soak: January 2-5, 1962.

Battery temperature excessive (Jan. 4, 1962).

Cause: Shunt regulator is designed to limit the voltage to the batteries at 13.5 volts. When the batteries are being charged, their temperature increases to cause a decrease in impedance and voltage and an increase in the current.

Solution: Shunt regulator was redesigned to incorporate current-limiting features as well as voltage regulation. In addition, the bottom of the battery pack was machined flat to form a better thermal contact with the satellite instrumentation shelf.

Tape recorder drawing excessive current (Jan. 4, 1962).

Cause: Lubricant crystallized.

Solution: During the remainder of the +47°C test and the subsequent 30 and 135 degree solar aspect tests, the tape recorder was turned off.

30° Solar Aspect:

Shunt regulator produced a 1-Mc oscillation across the batteries when dumping circuit was operating (Jan. 6, 1962).

Electron density was not operating properly. Single period of 138 μ sec was present instead of normal two sawtooth pulses (Jan. 6, 1962).

Cause: Suspect temperature gradient from sensor to electronics on boom.

Solution: Operation became normal as temperature of the electronics was increased.

Electron temperature not operating properly (Jan. 7, 1962).

Cause: Open condenser in sensor located in boom.

Solution: Repaired.

135° Solar Aspect: January 7-8, 1962.

No new problems developed. Previous problems continued.

Downtime: January 10 to February 10, 1962; all satellite deficiencies were corrected during this time.

Vibration Retest

Time: February 15, 1962.

Results: Satisfactory.

Thermal-Vacuum Retest

Time: March 7-12, 1962.

Results: Satisfactory.

Prototype Vibration Experiment and Separation System

Vibration

Time: March 10, 1962.

Results: Mechanical timers in separation system not accurate enough.

Temperature

Time: March 10, 1962.

Results: Satisfactory.

Humidity

Time: March 11-12, 1962.

Results: Satisfactory.

Thermal-Vacuum

Time: March 17-18, 1962.

Results: Satisfactory.

Flight Unit 1**Balance**

Time: January 11-15, 1962.

Simulated Items: Cosmic ray experiment.

Shunt regulator circuit.

Tape recorder current limiter.

Vibration

Time: February 7-11, 1962.

Results: Flight versions of above simulated items installed prior to vibration. No problems.

Thermal-Vacuum

Setup and Checkout: February 16, 1961.

Cold Test: February 17-21, 1962.

Cosmic ray experiment would not turn on at -8°C . Operation became normal at $+10^{\circ}\text{C}$.

X-ray experiment was not gating properly.

Cause: Faulty diode.

Solution: Replaced (action taken subsequent to thermal-vacuum test).

Electron energy No. 2 malfunctioned.

Cause: Open capacitor.

Solution: Replaced (action taken subsequent to thermal-vacuum test).

Shunt regulator power transistors operated at excessive temperatures.

Cause: Inadequate heat sink.

Solution: Modified method of mounting.

Hot Test: February 23-25, 1962.

Last three items in cold test (above) continued during this test. Electron density questionable.

Cold Test: February 26-27, 1962.

Items in paragraph above continued.

Corona Check: February 27, 1962.

No new problems.

Flight Unit 2**Balance**

Time: February 5-8, 1962.

Simulated Items: X-ray experiment stack (UCL-2).

Cosmic ray experiment.

Tape recorder.

Vibration

Time: February 21-24, 1962.

Setup: Flight versions of above simulated items installed prior to vibration.

Results:

Cosmic ray experiment failed during random vibration—tall-fixture thrust direction.

Cause: Failure of weld in photomultiplier tube.

Solution: Spare cosmic ray unit was vibration-qualified and installed in the system prior to thermal-vacuum.

Mass spectrometer operation was intermittent.

Antenna problem.

Cause: Loose micro-dot connector.

Solution: Tightened.

Thermal-Vacuum

Setup and Checkout: March 1-3, 1962.

Mass spectrometer malfunctioned.

Cause: Broken pin in sphere; probably occurred during vibration testing.

Solution: Replaced sphere.

Tape recorder dc control unit questionable.

Cause: Circuit design problem.

Solution: Unmodified unit installed. All units are being modified to correct design problem and reduce sensitivity.

Corona Test: March 3, 1962.

No problems encountered.

Cold Test: March 5-8, 1962.

Tape recorder malfunctioned.

Cause: Faulty idler bearing.

Solution: Replaced bearing.

Electron density questionable.

Cause: Under investigation.

Solution: Experiment was being operated at an unrealistically low temperature during test.

Hot Test: March 10-13, 1962.

No new problems. Tape recorder operated satisfactorily for 10 hr during this test, at which time malfunction reoccurred.

Flight Unit 1 Vibration Experiment and Separation System

Temperature and Humidity

Time: March 11-12, 1962.

Results: Satisfactory.

Vibration

Time: March 17, 1962.

Results: Mechanical timers in separation system not accurate enough.

Thermal-Vacuum

Time: March 17-18, 1962.

Results: Satisfactory.

Flight Unit 2 Separation System

Temperature and Humidity

Time: March 11-12, 1962.

Results: Satisfactory.

Thermal-Vacuum

Time: March 17-18, 1962.

Results: Satisfactory.

Vibration

Time: March 21-24, 1962.

Results: Mechanical timers in separation system inaccurate.

APPENDIX C

DESCRIPTION OF FLIGHT VIBRATION EXPERIMENT

The vibration telemeter installed in the Ariel I Dutchman is intended to provide measurements of flight vibration levels imposed on the spacecraft during burning phases of the Thor-Delta vehicle. Three mutually orthogonal accelerometers located on the Dutchman motor attachment ring will provide these data for the thrust, pitch, and yaw axes. An additional channel on this telemeter has been allocated for a combined measurement of solar aspect, optical contamination, and third-stage pressure switch closure.

The vibration telemeter is a four-channel FM/FM system employing a solid-state transmitter with 1.8 watts output on 240.2 Mc. An RF band-pass filter suppresses spurious radiation to IRIG specifications. Antennas are a pair of quadrupoles, diametrically positioned and bonded to the third-stage casing near the after end.

The vibration pickups are small piezoelectric accelerometers operating into separate charge amplifiers. System frequency response is limited by a low-pass filter with a cutoff frequency of 600 cps and a rolloff of about 9 db/octave. Thrust, pitch, and yaw vibration data modulate voltage-controlled oscillators on IRIG bands E, C, and A respectively. Solar aspect, contamination, and third-stage pressure switch closure data signals are mixed and modulate a voltage-controlled oscillator on IRIG band 13. A three-point ground-controlled voltage calibrator in the package provides 0, 2.5, and 5 volt levels to all voltage-controlled oscillators in parallel prior to launch. Power for the vibration telemeter package is provided by a pressurized pack of 20 Ag cells. Current drain is approximately 800 ma. Power for the aspect and contamination sensors is supplied by two 2-volt Hg cells.

APPENDIX D

CHRONOLOGY OF ATLANTIC MISSILE RANGE OPERATIONS, 1962

March 27:

It was found that the separation system and solar paddle arms (in folded position) interfered with each other. Removal of spacers on battery box and slight re-routing of separation system wiring harness near electronic timers resolved the problem.

This incident shows the need for mechanical interface checks between all major assembled units prior to shipment to Cape Canaveral.

March 29:

Fitting of Flight Unit 2 to third stage was checked and found to be satisfactory.

Final assembly of Flight Unit 1 on third stage was completed.

March 30:

During checkout a bad contact in one of the dipole motor flyaway pins (between the spacecraft and separation system) was found. Spacecraft was removed to clean contacts.

The escapement mechanism for the electronic booms was adjusted.

The solar paddles were installed but were found to be incorrectly located, and were re-positioned.

March 31:

Third stage and spacecraft were placed on balance machine to start balancing. In the afternoon the electron density boom grids were inadvertently damaged.

April 2:

The electron density boom from Flight Unit 2 was substituted for the damaged one.

The spacecraft was removed from the balance machine and taken to the antenna range for boom calibration. During reassembly for balance the

escapement mechanism for boom erection was found to be defective and was replaced.

April 3:

The spacecraft weighed 135.50 lb, and the separation ring weighed 0.46 lb; this weight plus the Dutchman and separation system gives a total of 177.9 lb.

April 4:

Balancing was completed. Spacecraft, third-stage combination was placed in handling container in preparation for installation on pad. Installation was accomplished at 1500 EST.

April 5 and 6:

Integrated system checkout verified satisfactory performance of spacecraft.

April 7:

Spacecraft was on gantry with launch vehicle. Strip coat was removed in afternoon after rain stopped.

April 8:

Thermal coatings were touched up, and batteries were charged.

April 9:

Complete spacecraft checkout was satisfactory.

April 10:

Countdown proceeded until T-6 minutes when 2nd stage failed to pressurize. Launch was scrubbed after hold of 4½ hr.

April 11:

Nitrogen pressurization tank of second stage was found deficient. Launch was rescheduled for

April 25, 1962, with a new second stage. The spacecraft was to be removed from third stage for recalibration. The vibration experiment also was to be recalibrated.

Reassembly and balance was scheduled for April 18-20, and mating with third stage was to be at 0800, April 21.

April 12:

Spacecraft and spacecraft separation system were removed from third stage. The vibration experiment was left in place to avoid possibility of damaging third stage in removal of studs.

A complete system check of Flight Unit 1 revealed satisfactory operation.

April 16:

An RF check of the vibration experiment was conducted.

April 17:

The vibration experiment was checked and found to be operating satisfactorily.

April 18:

Spacecraft and separation system were reassembled on the third stage, and the resulting combination was placed on the balance machine.

April 19:

Balance was completed with addition of 164.5 grams in weights—less than needed in balance for first launch attempt.

April 21:

Third stage was mated with second stage.

April 23:

T-1 day checkout of spacecraft and vibration experiment were completed on T-3 day.

April 24:

Strip coating was removed from spacecraft. Spacecraft thermal coatings were touched up, and its batteries were charged.

April 25:

Thermal-coating touchup was completed.

April 26:

After completion of countdown at 1200, launch vehicle failed to fire.

Launch vehicle problem was resolved, and countdown reverted to T-15 minutes. Thereupon, countdown proceeded to lift off at 1800:1703Z.

APPENDIX E

FUNCTIONAL TEST PROGRAM

Introduction

General

The functional tests were conducted on inertially correct mockups of the Ariel I spacecraft and final stage X-248 booster between October 16 and November 16, 1961. All tests were run under vacuum conditions in the 18.3m (60 ft) sphere at Langley Research Center (LRC). The operations required considerable inertia estimation, measurement, and compensation to provide correct spin-axis inertia for each test.

Equipment

The spin-drive system consisted of a vertically mounted 10 hp motor with a magnetic clutch driving an extension shaft having slip rings and a disk with 32 radially protruding studs, which generated pulses in an adjacent magnetic pickup to yield a Visicorder record of speed changes.

Two types of spin table assembly were used with the drive. For all testing involving the tall (12+ft) Ariel I plus X-248 test item, it was necessary to couple the drive to the diaphragm of a Scout spin bearing and skirt assembly, although the frictional speed decay of this assembly was excessive. It was possible to effect some improvement in the "coasting" by using a small suspension bearing at the top of the test item. The suspension cable tension was adjusted to support most of the weight.

For the stretch yo-yo test series, where better coasting characteristics were essential, a small support bearing was used, and a 58-cm (23-in.)-diameter magnesium table was made. The spacecraft structure shells were attached to inertia disks that were attached to the magnesium table. The small [approx. 10.2 cm (4 in.) outside diam.] support bearing and 2.86-cm (1 $\frac{1}{8}$ in.)-diameter extension shaft were adequate and suitable for the compact stretch yo-yo structure models, but were structurally inadequate to safely spin up the tall

top-heavy Ariel I composite assemblies. By suspending the composite assembly by an overhead cable and spinning it by hand, it appeared that substantial unbalance existed; and, even if accurate balancing could have been done, precise alignment could not have been insured.

Spinup

General

Two tests were run, using an inert-loaded X-248 furnished by LRC, with the Ariel I spacecraft shell with all appendages folded, installed on the Scout spin bearing and skirt, with drive disconnected. Inertia was estimated at 6.844 slug ft², comparable to flight condition.

First Spinup and Results

The first spinup test was conducted at 0.35 mm Hg. Objectives were to determine final speed and to check Lyman-alpha sensors for contamination. Results were a final speed between 120 and 130 rpm and no visible contamination of the sensors, which were returned to GSFC for further examination. The indicated final speed is not considered conclusive. The impulse rating of the PET rockets was below that required by the system, and the appearance of the PET rockets after firing indicated incomplete burning. Also the breakaway torque and running friction of the bearing was excessive (breakaway torque, 4 lb-ft+), and misalignment was suspected. The contamination check has value, but it is limited by the incomplete environmental simulation—that is, vacuum and spin environment was simulated, but zero-gravity environment was not.

Second Spinup and Results

The second spinup test was conducted at 10 mm Hg, using higher impulse PET rockets and a different skirt bearing, with part of the weight supported by an overhead bearing. Friction of

the system was substantially lower. A final spin rate of approximately 180 rpm was attained. This is comparable to the flight objective.

Despin

General

These tests were performed with the spacecraft structural shell inverted on the spin table with all appendages folded up and secured to the expended X-248 by the tiedown hardware (Figure E1). The assembly was slugged up to match the flight inertia, and appendages were restrained by rubber cords adjusted to compensate for gravity effect. It became evident that the low-friction table assembly with the small support bearing would not accommodate the unbalance of this assembly. So the Scout spin bearing and skirt were used, with a suspension cable and swivel-mounted bearing adjusted (with a turnbuckle and

spring scale) to take most of the weight of the assembly. The spin axis of the assembly was carefully aligned, and some effort was made to adjust the cable tension for minimum running friction. This adjustment was complicated by problems of dynamic resonance of the support cable in the operating speed range and the need to anticipate dimensional changes due to evacuation of the chamber and/or temperature changes. Temperature changes had the most effect on cable tension. After the tests it was apparent that running friction was undesirably high, but marginally acceptable.

Before each operational test, it was necessary to calibrate the running speed and speed decay so as to insure planned speed at the initiation of the operation. This required reduction of recorder tape data before each step. Sequence tests were conducted at 10 mm Hg vacuum.

90 Percent Flight Spin

All operations were initiated at 90 percent of nominal flight spin rate and recorder data on speed taken before, during, and after each operation. Photo coverage at 1000 frames/sec from three angles was also obtained at each operation. Operations were triggered manually, via slip rings and hard lines to control room. As a preliminary test, the spacecraft sequence timers were caused to cycle, while spinning at nominal flight speed (160 rpm) in vacuum; but the timer contacts ignited flash bulbs mounted on the payload structure instead of actuating release mechanisms. The time sequence of the flashes was noted, and met the performance specification for the timers. Sequential operation of yo-yo despin, experimental boom erection, and solar paddle deployment was then accomplished as planned except that the experiment booms failed to lock into extended position and had to be locked in by overspinning.

110 Percent Flight Spin

This was intended to be a repeat of the first sequence test, but with operations conducted at 110 percent of nominal speed and sequence timers operated 30 rpm after despin and erection events. However, initiation of the first event (yo-yo despin) fired all release squibs (except separation squibs, which had not been installed), and yo-yo

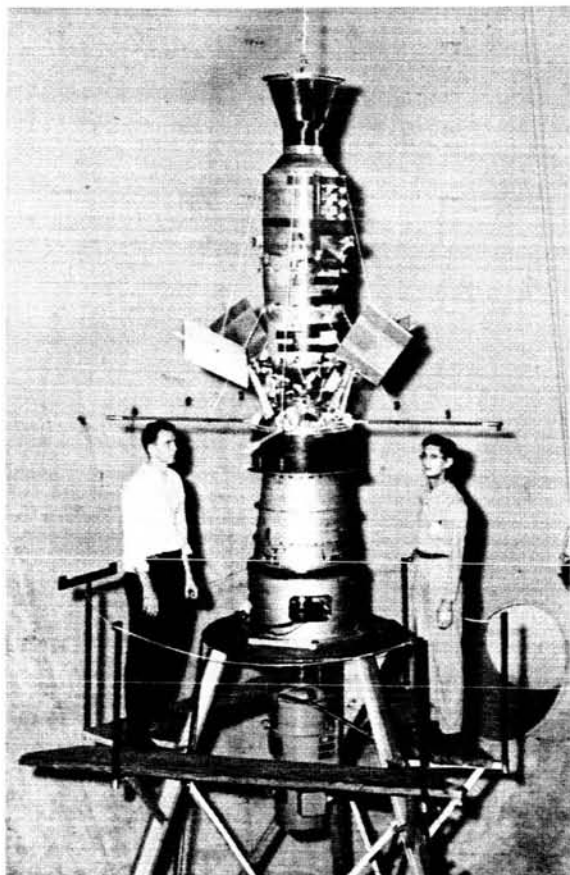


FIGURE E1.—Setup for boom and paddle erection test.

despin, experiment boom erection, and mass boom and solar paddle deployment all occurred simultaneously while spinning at 176 rpm. The reason for this misfire was discovered to be faulty wiring in the firing box—the three firing switches were inadvertently wired in parallel. The error did not show in the preceding test because firing leads were connected only as needed for each individual operation.

The results of this accidental overtest, of some value in showing structural weak points, were: speed change was violent; yo-yo operation was normal except that one wire snagged a broken paddle but snapped loose; experiment booms snapped nylon string of escapement mechanism, but booms survived extreme deflection without structural failure; both inertia booms snapped into detent, bounced back, and on rebound tore off both attachment brackets by shearing the screws—the inertia booms were restrained by the rubber cords, but one boom swung around and struck the side of the assembly; both single-hinged paddles sheared off the attachment screws at the hinge, broke the bungee cords, and fell to the floor of the chamber; both double-hinged paddles stayed on, but the detent pins in the secondary hinges were sheared off; all appendage attachment screws were loosened and elongated; the single-hinged paddles opened a little ahead of the double-hinged paddles, and the inertia booms slightly ahead of the experiment booms; all appendages opened within about $\frac{1}{4}$ second; and the yo-yo disengaged from hook about $\frac{1}{2}$ second after release. Final spin rate was about 30 rpm. The sequence timers were operated after the spacecraft assembly coasted to a stop. One timer functioned normally; the other failed to operate but was later found to be operable. The failure was found to be due to a leaky battery case, which is not considered significant because the overnight vacuum soak was not representative of flight conditions—neither was the extended location and operation of the batteries in an inverted position. The spacecraft assembly was reconstructed for another try—all damaged parts were replaced, and the firing box wiring was corrected.

Final Despin

This was a repeat of the first successful sequence test except that all operations were run at 110

percent of nominal speed and the sequence timer check was omitted—no replacement for the defective battery was available, and the reliability of the timers before and after exposure to despin was considered to have been established. The operational sequence was again accomplished as planned except that the experiment booms again failed to lock in extended position, despite intentional initial overspin estimated to compensate for friction-induced decay in speed during erection time of approximately 4.5 sec.

Test Results

The Visicorder records plotted speed vs. time curves for the despin operations; and, after due allowance for friction-induced speed decay, the results showed good correlation with theoretical predictions. The only significant deficiency was the failure of the experiment booms to lock in erected position.

GSFC Tests

Additional tests were run at GSFC using a test rig substantially the same as at LRC without vacuum simulation. As a result, the escapement mechanism for the experiment boom was successively modified until full deployment was achieved at minimum anticipated flight-spin rate.

Separation Tests

General

These tests were performed with the expended X-248 installed with nozzle down in the Scout skirt. The Dutchman separation device and the spacecraft shell with all appendages extended were installed on top of the X-248. The rubber cords and the batteries were removed from the spacecraft shell. The separable configuration weighed 23.1 kg (51 lb), closely approximating the requisite mass to achieve the in-flight separation velocity. A device to compensate for gravity effects during separation was installed. This consisted of an overhead bearing and support cable, with the cable wound around a conical grooved pulley and connected to a long calibrated rubber spring under initial tension. This device operated as a constant force spring, over a limited range, to counteract the weight of the separable configuration. A ratchet device was used to prevent slipback. Separation was effected by firing

the release squibs of the separation device through slip rings and a hard line to the control room. Separation tests were conducted at 10 mm Hg vacuum.

Test No. 1

Separation was effected by firing both diametrically opposed release squibs, while the assembly was spinning at approximately 30 rpm (lower than the intended 39 rpm). Results were satisfactory, by visual observation and later review of slow motion movies. The spacecraft separated straight and true, with no apparent tipoff, and was smoothly arrested by the counterweight device after rising approximately 3 ft. The spacecraft continued rotating for over 13 min on the small suspension bearing, with very little wobble.

Test No. 2

This was a repeat of the first separation except that only one release squib was fixed and spin rate at separation was about 20 rpm. The spacecraft again separated with no apparent tipoff. For both separation tests the spin rates were lower than planned—perceptible slowing down was noted in the few seconds preceding separation. Visicorder data for the preceding tests indicated the same phenomenon, to a lesser degree; and it was realized that the change in voltage caused by the photoflood electrical load was responsible. The remedy was to calibrate the speed with floodlights on, and this was done for the later tests.

Results

The separation device operated satisfactorily under the test conditions—that is, 10 mm Hg vacuum, spin rates as noted, spacecraft mass adjusted for flight condition relative to separation velocity, gravity effect on the spacecraft effectively nullified, spacecraft balanced, and spin axis at separation constrained to be substantially in alignment with geometric centerline of spacecraft, X-248, and separation spring. The spin-axis constraint condition is a limitation of simulation technique, comparable to flight situation only if adequate balance of spacecraft and of expended X-248 is assumed. This is not a valid assumption, as expended boost units characteristically have some unbalance. Therefore the tests verified alignment of the separation spring thrust

vector with the geometric centerline of the assembly in that no tipoff was apparent, but are inconclusive concerning possible tipoff due to dynamic unbalance at separation.

Stretch Yo-Yo System

First Series

Procedures: The Ariel I stretch yo-yo system with preloaded springs was despun from spin rates 80 and 120 percent of nominal. Designed resultant spin rate is 73.9 rpm. All tests were conducted at 10 mm Hg. Initial spin was adjusted to allow for speed decay after declutching. Despin was initiated manually, via slip rings and a hard line to the control room. Each test configuration was slugged up to the inertia of the flight system. Visicorder and photographic records were made of each run.

Results: System inertia was 2.774 slug ft.² Results were:

1. Spun at 160 rpm. Direction of rotation was incorrect, and weights and springs fell off on release.
2. Despun from 162.6 rpm to approx. 73.6 rpm.
3. Despun from 129.4 rpm to approx. 73.8 rpm.

Second Series

Results: The stretch yo-yo system was operated nine times in vacuum conditions, and the despin effect was noted for various combinations. For nominal initial spin and inertia the final spin rate was approximately as predicted, the effect of inertia changes was as expected, and the stretch yo-yo compensated satisfactorily for initial under-spin. However initial overspin, especially when combined with higher inertia, causes yielding on the yo-yo springs and overcompensation. As the overspin-correction discrepancy is attributed to yield of the spring material, it was planned to repeat pertinent tests using the higher yield point springs.

Post-Test: Springs were heat-treated for higher yield points.

Third Series

Procedure: The same as for previous tests.

Results: Three tests were run, using yo-yo springs that had been heat-treated for high yield point. Springs used in earlier tests had yielded

excessively. Improved performance was noted, although slight yield did occur at 25 percent over nominal spin rate. Observed results satisfy flight requirements but show slight deviation from design objectives.

The system was designed to give a final spin rate of 73.9 rpm, with nominal initial speed considered 160 rpm and nominal inertia considered 2.885 slug ft.² The low final spin obtained in series No. 3, where slight yield occurred, is consistent with previous findings that spring yield

will cause overcompensation of high initial spin. The high final spin obtained in series No. 1, followed by almost exactly the design spin rate for series No. 2, cannot be explained. Neither error is sufficient to compromise the Ariel I mission; therefore, as a flight item qualification test, the results are satisfactory. From the point of view of design theory verification it must be noted that the deviations from designed final spin, when no yield occurred, exceed the likely experimental error.

CREATION OF SPACE VACUUM UTILIZING CRYOGENIC TECHNIQUES*

J. E. A. JOHN AND W. F. HARDGROVE

Existing space simulators operate in the range of 10^{-6} to 10^{-10} Torr. To investigate phenomena that occur in the space vacuum, the design of an ultrahigh vacuum facility is being considered. The techniques of cryogenic pumping and cryosorption are evaluated with respect to providing a capability for simulating the space environment in the 10^{-15} Torr and lower range. High pumping speeds at ultrahigh vacuum are possible both for the condensables by cryogenic pumping at 15°K to 20°K , and for the noncondensables by cryosorption. Utilization of the combined advantages of cryosorption and cryopumping is described. Design considerations are discussed.

INTRODUCTION

Space simulation chambers currently in use for conditioning of spacecraft are capable of pressures down to 10^{-8} to 10^{-9} Torr. However, the pressures to which the external surfaces of the spacecraft are exposed in space may be many orders of magnitude below this. Effects such as cold welding, seizure of certain clean metals in contact, changes in friction coefficient, and other unforeseen problems could arise in a spatial environment, but not be revealed in any of the space simulation chambers now in operation. Mechanical devices such as the pantograph for the surface sampler on the Surveyor spacecraft, the unfolding boom extensions, the active-mechanical temperature control system of spacecraft, such as OGO, with a planned apogee exceeding 50,000 miles, require environmental testing under conditions considerably more stringent than can be provided by chambers operating at 10^{-8} Torr. Spacecraft test criteria must be developed for the ultrahigh vacuum range of exposure. For these reasons the design of a system is being considered which will more closely simulate the spatial environment. It is believed that this system represents an advancement in the art of spacecraft testing and also in the application of cryogenics to vacuum technology.

The ultimate steady-state pressure in a vacuum system results from a balance between pumping speed and the summation of the real and virtual

leak rates. Thus, $P_{\text{ultimate}} = Q/S$ where Q = sum of real and virtual leak rates in Torr liters per second and S = pumping speed in liters per second. The design of an ultrahigh vacuum system must incorporate a high pumping speed with a minimal leak rate.

The method proposed for reducing the real gas leak rate into the test chamber is to provide a guard vacuum around the chamber. This reduces the pressure difference across any leak and thus the leak rate into the test volume by several orders of magnitude. High pumping speeds at ultrahigh vacuum in the test chamber for condensable gases are provided by cryogenic pumping (Figure 1). For the main component of the noncondensable gas load, hydrogen, a cryosorption system is to be utilized which will selectively capture this molecule. A high pumping speed for hydrogen at ultrahigh vacua is a necessity for the system since, presumably, hydrogen is one of the constituents of the virtual leak rate due to outgassing of spacecraft materials.

DESCRIPTION OF SYSTEM AND OPERATION

The proposed system consists of three concentric chambers, the outer three feet in diameter by three feet high and the inner containing a test volume of size $20'' \times 20''$ (see Figure 2).

The space between the outer two chambers is to be evacuated by a conventional pumping system consisting of a mechanical pump, liquid nitrogen cold trap and diffusion pump. Seals across the outer wall are to consist of double elastomer

*Proceedings of the Institute of Environmental Sciences, April 1964.

O-rings. The space between the O-rings is evacuated to provide a guard vacuum. The outer chamber will be baked out to minimize outgassing, and thus allow a lower pressure in this outer volume. Further, provisions have been made for cooling the O-rings to minimize gas permeation; considerable evidence (1) indicates that this technique is successful in improving the ultimate vacuum.

The intermediate chamber walls are to be cooled with liquid nitrogen to reduce the heat load on the walls of the inner chamber which are cooled to 15°K to 20°K by gaseous helium. The vacuum obtainable in the intermediate space should be at least an order of magnitude better than that existent in the outer space, since it is isolated from the possibility of in-leakage from atmospheric pressure. A separate pumping system, again consisting of a diffusion pump, liquid nitrogen cold trap, and mechanical pump, is to be used for initial evacuation of the intermediate space and test volume (Figure 3). The pressure, after pumpdown, should be approximately 10^{-10} to 10^{-11} Torr in the test volume and surrounding guard vacuum.

After pumpdown, the valve connecting the test volume to the surrounding guard vacuum is closed, isolating the test volume from the mechanical and diffusion pumps. Gaseous helium is circulated in the walls of the test volume, cooling it to 15°K to 20°K. Thus, oxygen, nitrogen and all other condensable molecules are cryogenically pumped, leaving only the noncondensables in the test chamber. Following this, gaseous helium at or below 20°K is circulated in the chevron baffles and cryosorption panel for pumping of residual hydrogen and helium. The baffles prevent molecules other than hydrogen or helium from reaching the cryosorbent panel. Thus, a system is provided that is capable of attaining ultrahigh vacuum and, further, has the capacity, at low pressures, for handling large gas loads (both of condensable and noncondensable gases) which would be outgassed from the spacecraft equipment under test.

DESIGN CONSIDERATIONS

Heat Loads on Shrouds

One of the fundamental considerations involved in the design of a cryogenic pumping system is the

heat loads on the cryogenic shrouds which determine either the amount of cryogenic fluids required or, in this case, the size of refrigerator required to supply gaseous helium at 15°K to 20°K. Before cryogenic fluids are introduced, the system will have been pumped to a pressure level where the only heat transfer is due to radiation. The heat transfer, via radiation between a surface and another enclosing it, is given by:

$$q = \frac{(.1714 \times 10^{-8})(T_1^4 - T_2^4)A_1}{\frac{1 - \epsilon_1}{\epsilon_1} + \frac{(1 - \epsilon_2)A_1}{\epsilon_2 A_2} + 1}$$

where:

q = heat transfer rate (Btu/hr)

ϵ_1, ϵ_2 = emissivities of surfaces, assumed equal to absorptivities

A_1 = enclosed area (ft²)

A_2 = enclosing area (ft²).

The values of ϵ for condensate-coated surfaces can be assumed as approximately .10, since the thickness of the coating will be quite small for the pressures involved (3). Under steady-state conditions, the heat exchange between the 77°K surface, of surface area 29.4 ft², and the 20°K surface, of area 18.8 ft², will then amount to:

$$q = .76^\circ\text{Btu/hr or } .22 \text{ watts.}$$

If there is a test object in the chamber maintained at room temperature, the gaseous helium refrigeration requirements greatly increase. If it is assumed that the test object is black ($\epsilon=1$), in the form of a one-foot diameter sphere (surface area = 3.14 ft²) and maintained at 273°K, the heat load increases to:

$$q = 127 \text{ Btu/hr on } 37.3 \text{ watts.}$$

This estimate of heat load is optimistic in that no account has been taken of the heat transfer by conduction through support rods, cryogenic piping, etc., or of the refrigeration requirements during initial cooldown. However, the heat load due to radiation from the surrounding shroud and from the test object certainly represents the largest part of the refrigeration required. In any case, the load is within the capacity of a 250-watt helium cryostat.

Cryogenic Pumping Speeds

Cryogenic pumping speeds at ultrahigh vacua are dependent on the rate at which the molecules of condensable gases strike the cryogenically-cooled surfaces and also on the sticking coefficient of these molecules. The number of molecular collisions per unit area per unit time can be found from kinetic theory (4):

$$\sqrt{v} = \frac{1}{4} n \bar{v}$$

where:

- \sqrt{v} = number of collisions per unit area per unit time
- n = number of molecules per unit volume
- \bar{v} = mean molecular velocity

$$= 14,511 \sqrt{\frac{T}{M}} \text{ cms per second.}$$

Further, the volume flow striking the surface is given by:

$$\dot{V} = \frac{1}{4} \bar{v} \text{ or } = 3,638 \sqrt{\frac{T}{M}} \text{ cm}^3 \text{ per second per cm}^2.$$

where:

M = molecular weight of gas striking surface.

The actual pumping speed will be given by:

$$S = f \dot{V}$$

where f is the sticking coefficient, dependent on the surface temperature, vapor pressure of the gas at this temperature, and kinetic energy of the molecules striking the surface.

The main constituents of the cryopumped gas load will be oxygen and nitrogen. Data available suggest that the value of f for these gases on a 20°K surface will be close to 1.0 (5). Thus, for nitrogen, at a mean temperature of 100°K:

$$\begin{aligned} S &= 6.69 \text{ liters per second cm}^2 \text{ of surface area} \\ &= 117,000 \text{ liters per second for the actual} \\ &\quad \text{surface area of the inner chamber.} \end{aligned}$$

For oxygen,

$$\begin{aligned} S &= 6.27 \text{ liters per second cm}^2 \text{ of surface area} \\ &= 110,000 \text{ liters per second.} \end{aligned}$$

Pumping of Noncondensable Gases

It has been shown that cryogenic pumping provides high pumping speeds for the condensable gases. However, it is anticipated that hydrogen, with vapor pressure of 760 mm of mercury at 20°K, will be a constituent of the gas load from

spacecraft materials. Therefore, provisions must be made for pumping hydrogen gas. Attempts, meeting with limited success, have been tried using cryotrapping (6) and gettering using evaporated molybdenum films (7). However, the method which seems to provide the greatest possibility for success is cryosorption, the adsorption of gas on an adsorbent material maintained at cryogenic temperatures. The potentialities of cryosorption have long been recognized; however, its use in pumping hydrogen at ultrahigh vacua has met with two major engineering difficulties, that of maintaining the adsorbent bed at cryogenic temperatures with minimal temperature gradients across the adsorbent, and that of containing the adsorbent in the chamber.

Both problems appear to have been solved by the successful bonding of an adsorbent (Molecular Sieve 5A) to stainless steel or aluminum plate (2). The bond is reported to be unaffected by temperature cycling from room temperature to 20°K. The adsorbent thickness (from $\frac{1}{64}$ " to $\frac{1}{8}$ ") is such as to minimize temperature differences across the adsorbent layer. The panel, coated with adsorbent and cooled to 20°K, has been shown to provide pumping speeds for hydrogen of 6,600 liters per second ft² of adsorbent surface area at 10⁻⁸ Torr, corresponding to an adsorption coefficient of 0.16 (2). This pumping speed remains fairly constant until an appreciable fraction of the surface has been utilized. It becomes of primary importance, then, to obtain information concerning the volume of hydrogen that can be adsorbed before saturation of the panel is approached, this influencing the pumping speed and the length of time before desaturation of the adsorbent is necessary.

The volume of a gas adsorbed by a surface is primarily a function of temperature and pressure. The Dubinin-Radushkevitch equation (9) has met with success in predicting the adsorption of nitrogen on glass at 77°K (10) and that of hydrogen on charcoal at 20°K (11). An application of this equation to estimating the adsorption of hydrogen on Molecular Sieve 5A would thus seem to be of considerable merit.

According to the Dubinin-Radushkevitch equation:

$$\log_{10} V = \log_{10} V_{\text{sat}} - AT^2 \left[\log \frac{P_{H_2}}{P_{\text{sat}}} \right]$$

where

V = volume adsorbed at hydrogen pressure

P_{H_2} (std. cc)

P_{sat} = vapor pressure of hydrogen at 20°K

A = experimental constant.

Utilizing data obtained at 10^{-6} to 10^{-9} Torr (2) for the adsorption of hydrogen on Molecular Sieve 5A, the numerical constants in the equation were obtained. Thus:

$$\log_{10} V = 3.35 - 5 \times 10^{-5} T^2 (\log_{10} P / P_{sat})^2$$

at $P = 10^{-15}$ Torr, $V = .00089$ std. cc/gram of adsorbent, or $V = 6.8 \times 10^{11}$ liters of hydrogen per gram of adsorbent at a pressure of 10^{-15} Torr. Since 4 square feet of panel area corresponds to 370 grams of adsorbent, the total capacity of the panel amounts to $2.5 = 10^{14}$ liters of hydrogen at a chamber pressure of 10^{-15} Torr. If there were to be a gas load of 10^{-8} Torr liters of hydrogen per second from the test object, the panel would be saturated after:

$$\begin{aligned} t &= \frac{2.5 \times 10^{14}}{10^{-8} \times 10^{15}} \\ &= 2.5 \times 10^7 \text{ seconds} \\ &= 7.0 \times 10^3 \text{ hours.} \end{aligned}$$

This example demonstrates that the adsorbent can be used for a long term test with no need for desaturation and regeneration. Further, such a small fraction of the adsorbents' capacity will be utilized at ultrahigh vacua, that the adsorption coefficient will remain constant. Since, from kinetic theory, the volume of a gas striking unit area of a surface per unit time is independent of the gas pressure, the pumping speed for hydrogen will remain invariant during the test, 6,600 liters per second ft² of panel area.

The cryosorbent panel must be baffled in order to prevent gases other than hydrogen from reaching the adsorbent. Data are available (2,8) which indicate that small volumes of condensable gases, e.g. nitrogen, on the adsorbent surface cause a marked decrease in the hydrogen adsorption capacity of the cryosorbent.

It is of interest to compare, at this point, the capabilities for handling hydrogen of the cryopumping-cryosorption system operating at 15° to 20°K with those of a cryopump system operating

with liquid helium at 4.2°K, as has been constructed at the National Research Corporation and at Lewis Laboratories (NASA). Whereas the quoted pumping speed for cryosorption of hydrogen on Molecular Sieve 5A corresponds to an adsorption coefficient of 0.16, the sticking coefficient for hydrogen on a 4.2°K surface is only about .01 (12). Hence, cryosorption at 20°K provides higher pumping speeds than cryopumping at 4.2°K. Further the cryosorption system has obvious economic advantages.

Leak Rates

The performance of a vacuum system depends, to a great extent, on the magnitude of the real and virtual gas leak rates. The rate of real gas in-leakage has been minimized by providing a guard vacuum of 10^{-10} to 10^{-11} Torr. If it is assumed that a hole exists in the test chamber of area .01 cm², the molecular flow through the hole would be given by:

$$Q = C \Delta P$$

where

C = conductance (liters per second)

Q = rate of in-leakage (Torr liters per second)

P = pressure differential across opening

where

$$C = 3.64A \sqrt{\frac{T}{M}} \text{ liters per second (4)}$$

A = orifice area (cm²).

The worst possible case of in-leakage would be that of hydrogen, since hydrogen has the lowest molecular weight. Assuming the hydrogen originates from a volume at 77°K,

$$\begin{aligned} C_{H_2} &= 3.64(.01) \sqrt{\frac{77}{2}} \\ &= 226 \text{ liters per second} \end{aligned}$$

Thus, the rate of real gas in-leakage for this case is equal to:

$$Q = .226 \times 10^{-10} \text{ Torr liters per second.}$$

The virtual gas load is due to outgassing of the test object and the chamber walls. Since the walls of the test chamber are maintained at 20°K, and Dayton (13) has shown that the outgassing rate of metals decreases logarithmically with temperature, it is assumed that the gas load emanating

from the chamber walls will be negligible. The test object, however, may be operating at room temperature, providing a significant contribution to the gas load in the chamber.

In estimating the performance of the system, as indicated by the ultimate pressure in the chamber, two cases will be considered; first, a clean and empty chamber and second, a chamber with a test object having a prescribed outgassing rate. The pumping speeds provided for the condensable gases by cryogenic pumping and the pumping speed for hydrogen by cryosorption have already been indicated. It will be assumed that the surface area of the cryosorption panels is 4.0 square feet, providing a pumping speed for hydrogen of 26,400 liters per second.

For the clean and empty chamber, with negligible virtual leak rate, the real gas in-leakage was found to be 2.26×10^{-11} Torr liters per second, assuming the gas to be hydrogen and an opening of area .01 square centimeters. Thus, the ultimate pressure attainable, at steady state conditions, would be:

$$\begin{aligned} P_{\text{ultimate}} &= \frac{Q}{S} \\ &= \frac{2.26 \times 10^{-11}}{2.64 \times 10^4} \\ &= 8.7 \times 10^{-16} \text{ Torr.} \end{aligned}$$

With a test object in the chamber, it will be assumed as an example, that the gas load will be 2×10^{-9} Torr liters per second, with 1 percent of this being hydrogen, the rest nitrogen. Now, the ultimate pressure is determined by the pumping speed for nitrogen, 117,000 liters per second. The partial pressure of nitrogen is given by:

$$\begin{aligned} &= \frac{2 \times 10^{-9} (.99)}{1.17 \times 10^5} \\ &= 1.69 \times 10^{-14} \text{ Torr.} \end{aligned}$$

The partial pressure of hydrogen, for this case, is:

$$\begin{aligned} P_{H_2} &= \frac{2 \times 10^{-11}}{2.6 \times 10^4} \\ &= 7.7 \times 10^{-16} \text{ Torr.} \end{aligned}$$

Thus, for the gas load specified in this example, the limit on the performance is determined by the

pumping speed for the condensable gas, not the noncondensable gas.

INSTRUMENTATION

In order to specify the environment to which a test object has been subjected, it is necessary to be able to measure the pressure or molecular flux at several points in the test chamber and also, to be able to determine the predominant molecular species, and their partial pressures.

Figure 4 shows the approximate pressure ranges of several different types of gauges. Those of interest in this application, where the ultimate pressure will be of the order of 10^{-15} Torr, include the cold cathode (Redhead) gauge, hot cathode gauge, and the mass spectrometers. It is significant that Alpert (14) asserts that the lowest known direct measurement of total pressure was 4×10^{-13} Torr, made by Lafferty (15) using a hot cathode, magnetron gauge. However, with any gauge operating at these pressure levels, especially the hot cathode gauges, there is the question of the contamination introduced into the chamber by surface effects (out-gassing, adsorption, etc.) taking place on the gauge itself. Obviously, it is a requirement of any measuring instrument not to change the environment in which it is inserted. It seems, at this time, that methods of achieving ultrahigh vacua have surpassed the means for measuring vacua.

Instrumentation for the chamber will have to be provided for measuring the principal constituents of the residual gas in the chamber at any time and their percentages of the total gas present. Alpert (4) indicates that the optimum instrument at the present time is a high sensitivity, mass spectrometer developed by Davis (16) and capable of measuring partial pressures as low as 10^{-16} Torr. Again, the accuracy of the instrument depends on a knowledge of the surface effects introduced by the instrument itself. Initial measurements will be made with a cold cathode, Redhead gauge and electrometer, capable of indicating pressures down to 10^{-13} Torr. It is envisioned that, when the research chamber becomes available, improvements will have been made in gauging that will extend the measurable range down below the present level. It is anticipated that research will have to be done in order to amend mass spectrometers currently available

to allow measurements of partial pressures of hydrogen, helium and other gases present in the chamber at 10^{-15} Torr.

CONCLUSIONS

In conclusion, a system designed to approach the pressure environment of space, utilizing cryogenic pumping and cryosorption at 15°K to 20°K has been described. Calculations have been shown which demonstrate the feasibility of such a system. High pumping speeds are provided for the condensable gases by cryopumping, and for hydrogen, the principle noncondensable gas, by cryosorption. Ultimate pressures in the test volume are calculated to be beyond the limit of present instrumentation.

REFERENCES

1. FARKASS, I., GOULD, P., and HORN, G., " 1×10^{-11} Torr in Large Metal Chambers," Transactions of Ninth Vacuum Symposium, American Vacuum Society (1962).
2. HEMSTREET, R., et al., "The Cryosorption Pumping of Hydrogen at 20°K," Research Laboratory Linde Company, Division of Union Carbide Corporation, Tonawanda, New York, Final Report, AEDC-TDR-62-200, to be published under AF40(600)-944.
3. MOORE, B., "Effect of Gas Condensate on Cryopumping," Transactions of Ninth Vacuum Symposium, American Vacuum Society (1962).
4. DUSHMAN, S., "Scientific Foundations of Vacuum Technique," Second Edition, John Wiley and Sons, Inc. (1962).
5. MULLEN, L., and JACOBS, R., "Some Characteristics of a Simple Cryopump," Transactions of Ninth Vacuum Symposium, American Vacuum Society (1962).
6. HEMSTREET, R., et al., "Research Study of the Cryotrapping of Helium and Hydrogen During 20°K Condensation of Gases," Research Laboratory, Linde Company, Division of Union Carbide, Tonawanda, New York, AEDC-TDR-62-127, May (1963).
7. HUNT, A. L., et al., "Gettering of Residual Gas and the Absorption of Hydrogen on Evaporated Molybdenum Films at Liquid Nitrogen Temperatures," Advances in Cryogenic Engineering Vol. 8, Plenum Press, New York (1963).
8. STERN, S. A., et al., "The Cryosorption Pumping of Hydrogen and Helium at 20°K," Laboratory, Linde Company, Division of Union Carbide Corporation, Tonawanda, New York (Interim Report), AEDC-TDR-62-200, October (1962).
9. DUBININ, M., and RADUSHKEVITCH, L., Proceedings of Acad. Sci. USSR, 55 (1947).
10. HOBSON, J., "Physical Absorption of Nitrogen on Pyrex at Very Low Pressures." J. Chem. Phys., 34, 1850 (1961).
11. BEYER, G., "Cryosorption," paper contributed to Final Report of Goddard Summer Workshop Program, NASA Report X-320-63-264 (1963).
12. BACHLER, W., KLIPPING, G., and MASCHER, W., "Cryogenic Pump Systems Down to 2.5°K," Transactions of Ninth Vacuum Symposium, American Vacuum Society (1962).
13. DAYTON, B., "Outgassing Rate of Contaminated Metal Surfaces," Transactions of Eighth Vacuum Symposium, American Vacuum Society (1961).
14. ALPERT, D., "Ultra-High Vacuum: A Survey," Physics Today, August (1963).
15. LAFFERTY, J., "The Hot-Cathode Magnetron Ionization Gauge with an Electron Multiplier Detector," Transactions Ninth Vacuum Symposium, American Vacuum Society (1962).
16. DAVIS, W., "Sputter-Ion Pumping and Partial Pressure Measurements Below 10^{-11} Torr," Transactions Ninth Vacuum Symposium, American Vacuum Society (1962).

A BIFILAR MOMENT OF INERTIA FACILITY*

GEORGE H. JONES

SUMMARY

Most mass measurement facilities perform single functions. However, a facility for performing multiple mass properties measurements of rigid bodies, such as spacecraft components and/or assemblies weighing up to 1,000 pounds, is in operation at the NASA—Goddard Space Flight Center. It provides measurements of weight, center of gravity, moment of inertia, and product of inertia in a completely self contained facility.

Weight measurement is provided by a calibrated load cell mounted in the support system; center of gravity, by load cells arranged to determine the mechanical moment required to balance the weight moment of the test item; moment of inertia, through application of the bifilar pendulum; and product of inertia, by means of dynamic recording of force moments developed in torsional oscillation.

A view into the background of spacecraft mass measurements at Goddard Space Flight Center is discussed. The advent of the second generation or observatory type scientific spacecraft pointed to the need for larger capacity and more accurate mass measurement facilities. This need led to consideration of a multiple mass measuring facility instead of individual facilities for the separate operations.

The development history of the facility is discussed, including problems encountered and their solutions. The historical treatment begins with the design specification and runs through the proposal, contractor's development, installation, modifications, acceptance, and operational experience.

A description of the facility is given with a functional discussion of the major components and assemblies. Built-in handling and maintenance aids are described.

A test sequence for a typical sample spacecraft is described. Use of data sheets for recording test measurements is described, and the derivation of constants used in the data sheets is shown.

BACKGROUND

The experimental measurement of mass characteristics of spacecraft, sounding rockets, and/or components involves individual measurements performed on separate facilities or equipment. This is costly both in time and in the laboratory space needed for the separate facilities. The valuable (often priceless) spacecraft are subjected to a high degree of handling and transportation in the normal performance of routine test measurements. Why, then, continue to make these measurements on separate facilities? One obvious answer, of course, is cost. But if, on the other hand, current or near future needs substantiate a requirement for equipment of a larger

capacity and/or greater accuracy than now exists, why not consider a facility for performing multiple mass properties measurements of rigid bodies? This is precisely what was done at Goddard Space Flight Center.

With the advent of the second generation or observatory type scientific spacecraft, the need for accurate center of gravity location and moment of inertia measurements rather forcefully pointed to a need for new equipment. The Scientific Space Age was no longer in the 100 to 300 pound payload era, but rather in the era of upwards to 1000 pounds for the observatory experiments. The Goddard Space Flight Center simply did not have facilities capable of performing mass measurements of spacecraft of this size and weight. A need existed for the development of new facilities.

*Presented at the 23rd National Conference of the Society of Aeronautical Weight Engineers, Dallas, Texas, May 18-21, 1964.

DEVELOPMENT HISTORY

Experimental measurement of mass moments of inertia can be made in a number of different ways. Some of the more common methods employ principles of the simple torsional pendulum, the compound pendulum, the bifilar pendulum, the inclined plane and the rocker. These are shown in simple form in Figure 1. Consideration of the advantages of: (1) relatively uncluttered space above the support platform; (2) ease of immobilizing a support platform for loading and unloading; (3) relatively safe measurement operation; and (4) potentially high degree of accuracy led to the selection of a bifilar pendulum facility for moment of inertia measurements. A specification was subsequently prepared for the design and construction of a moment of inertia facility capable of measuring weight, center of gravity, moment of inertia, and product of inertia of rigid bodies such as spacecraft components and/or assemblies weighing up to 1000-pounds.

To provide a higher degree of accuracy without undue complexity, it was specified that the facility have two measuring ranges defined as low range and high range. In the low range configuration, the facility was required to provide measurement capability of test items having the following characteristics: (1) maximum size—30 inches diameter \times 48 inches long; (2) weight 30 pounds through 180 pounds; and (3) moment of inertia—may be as low as 0.3 slug-ft² about any principal axis. In the high range configuration the facility was required to provide measurement capability of test items having characteristics as follows: (1) maximum size—4 feet diameter \times 10 feet long; (2) weight—160 pounds through 1000 pounds; and (3) moment of inertia—may be as low as 3.0 slug-ft² about any principal axis.

Accuracy requirements were specified for low range as follows: (1) weight—within ± 0.15 percent or ± 0.2 pound, whichever is greater; (2) center of gravity—within 0.030 inch of actual position; (3) moment of inertia—within ± 0.5 percent or ± 0.006 slug-ft², whichever is greater; and (4) product of inertia—minimum sensitivity of 0.05 slug-ft² required. Similarly, the high range accuracies were specified as follows: (1) weight—within ± 0.15 percent or ± 1.0 pound, whichever is greater; (2) center of gravity—within 0.030

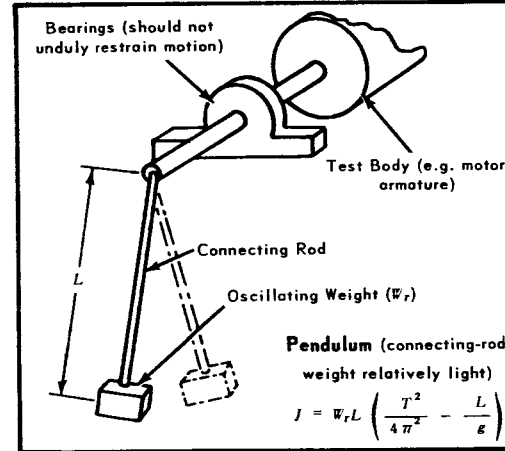
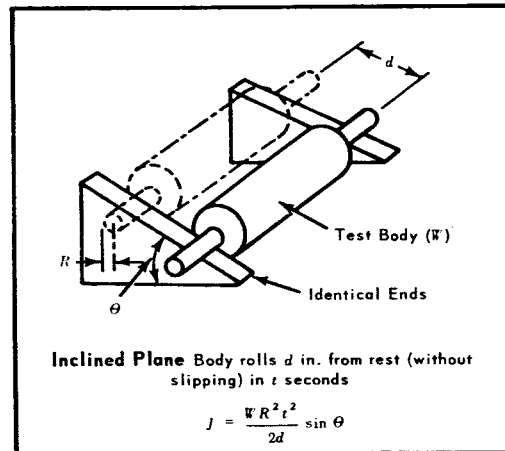
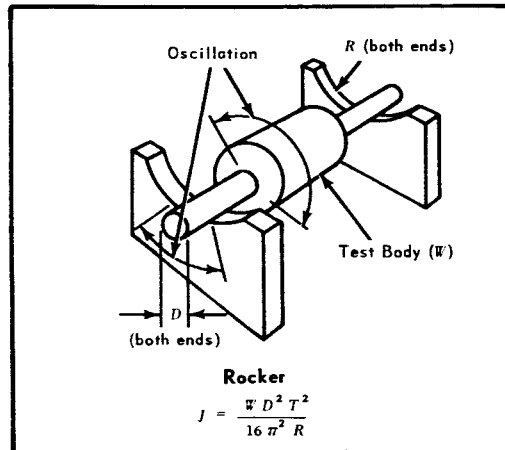
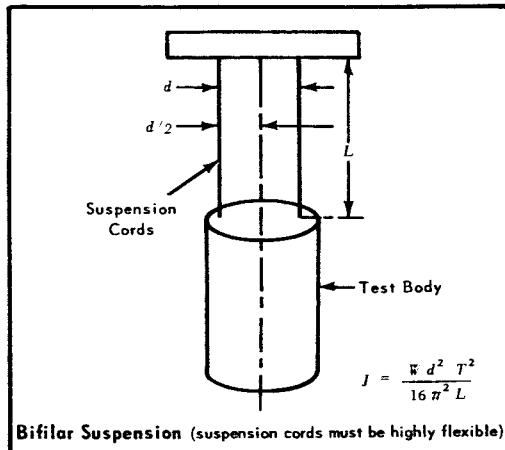
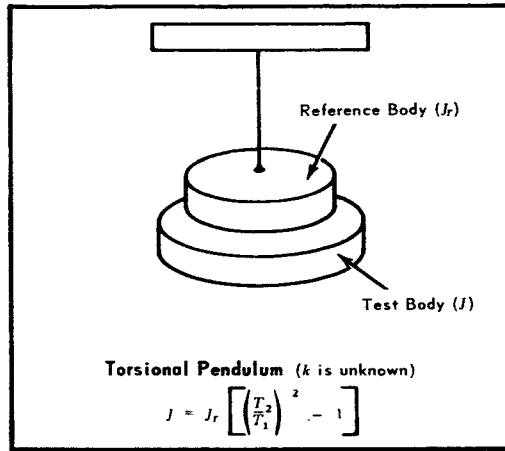
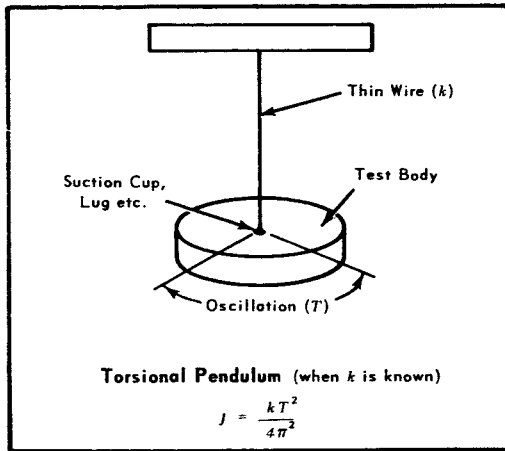
inch of actual position; (3) moment of inertia—within ± 0.5 percent or ± 0.04 slug-ft², whichever is greater; and (4) product of inertia—minimum sensitivity of 0.50 slug-ft² required.

A Request for Proposal was submitted to a number of qualified contractors in routine manner and the normal procedures were followed in evaluating the technical proposals. A contract was ultimately awarded to Miller Research Laboratories (now Miller Research Corporation) of Baltimore, Maryland for the design, development, manufacture, installation, calibration and acceptance testing of a multipurpose mass measurement facility in accordance with Goddard's specification.

The measurement principles in the proposed design had been tried and proven and represented no radical departures from familiar techniques. Weight is measured simply by the use of a calibrated electronic load cell mounted in the support system of the base structure. Load cells of two capacities are supplied to provide a higher degree of accuracy in the low range.

Center of gravity displacement measurement utilizes a mechanical moment principles. The unbalanced moment created by a displacement of the center of gravity of the test item from the neutral axis of the facility is balanced by a mechanical moment that is easily determined by load cells arranged to directly measure the moment force. This system is significantly more accurate than systems which rely upon load measurements at two or more approximately equal points of support, wherein the center of gravity must be determined by computing the difference between moments calculated from relatively small differences in relatively large weight measurements. The accuracy of the mechanical moment method of center of gravity determination is largely independent of the weight of the item being measured.

Moment of inertia measurement utilizes the bifilar torsional pendulum principles. The torsional pendulum technique has inherent advantages over other pendulum methods in that the axis of oscillation of the torsional pendulum is coincident with (or nearly so) the centroidal axis of the test item about which moment of inertia is to be measured. The error resulting from inaccuracies in the measured parameters asso-



SYMBOLS	
D = dia, in.	k = torsional spring constant, oz-in./radians
d = distance, in.	L = length of pendulum, in.
g = gravity, in./sec ²	R = radius, in.
J = inertia of test body about symmetry axis, oz-in.-sec ²	T = period of oscillation sec
J_r = reference inertia, oz-in.-sec ²	W = weight of test body, oz
	W_r = reference weight, oz

FIGURE 1.—Simple experimental methods of measuring moment of inertia.

ciated with the transfer of axis is thus eliminated or minimized. The bifilar torsional pendulum permits greater freedom and more versatility in positioning the test on a platform or table than is possible with a simple torsional pendulum.

Product of inertia is measured by a dynamic recording of the force moments developed in torsional oscillation. The load cells used to determine the static unbalanced moment for center of gravity location are used in conjunction with a recorder for the dynamic force moment measurement.

DESCRIPTION OF FACILITY

Figure 2 provides an overall view of the facility. The main components are: (1) a supporting A-frame structure; (2) a test item support platform or table; (3) torsion rods; (4) torsion pivot assembly; (5) load cells; and (6) equipment and control console.

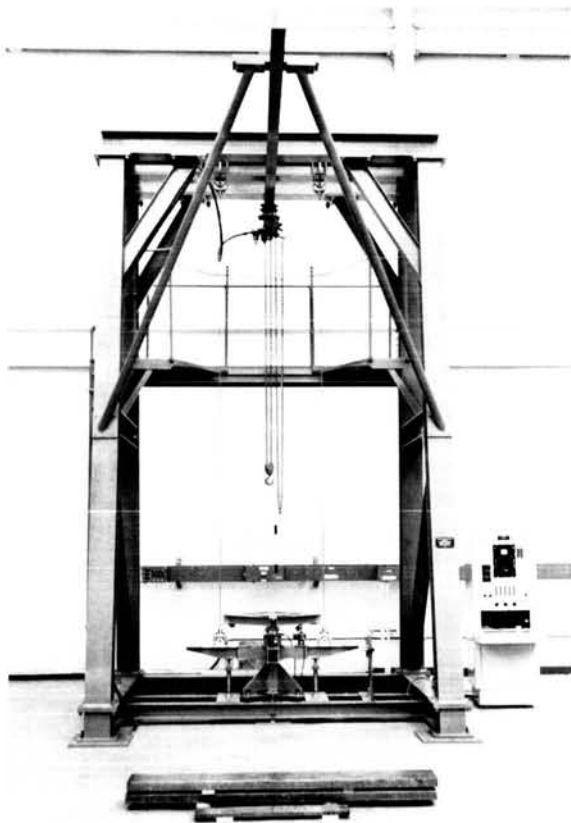


FIGURE 2.—Bifilar moment of inertia facility.

The supporting A-frame structure provides the rigid framework of the facility, and allows it to be a self-contained facility that does not have to be attached to the building structure.

Two support platforms, or tables, are provided for the high and low ranges. They are shown in Figures 3 and 4. The lower tare weight and moment of inertia of the low range table permits

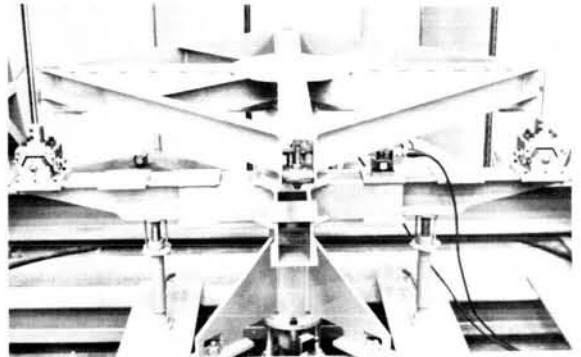


FIGURE 3.—High range table.

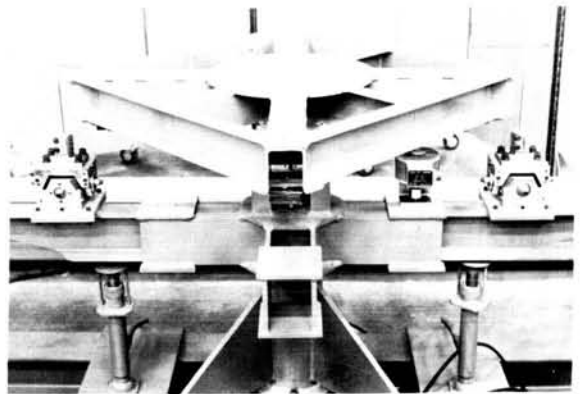


FIGURE 4.—Low range table.

greater accuracies of the measured mass characteristics of small test objects to be obtained. The tables in both ranges are statically and dynamically balanced, within the sensitivity of the measuring equipment of the facility, to minimize inaccuracies and eliminate the need for correction calculations in the characteristics of the swinging gear. Range change can be accomplished rather rapidly (about 30 minutes by two men) and with a minimum of adjustment.

The support platform or table is attached, through moment load cells, to a support beam as

shown in Figure 5. The support beam is, in turn, suspended from the upper structure by two torsion rods or filaments. At either end the bifilar filaments are attached to bearing housings which permit two degrees of freedom during the torsional pendulum oscillations.

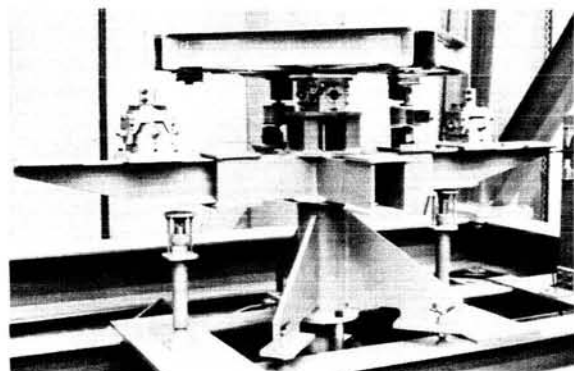


FIGURE 5.—Support beam assembly.

The lower center pivot assembly restrains the oscillation of the support platform against any mode other than torsion when the system is oscillated as a bifilar torsional pendulum. The pivot assembly also permits linear motion to occur during torsional oscillation. The pivot shaft provides the necessary rigidity to support the table assembly and the test object on the weighing load cell when the bifilar filaments are disconnected. This is necessary only when the test object is being weighed.

The hydraulic jacks beneath the support beam are used for two purposes: (1) they immobilize the table when it is being loaded or unloaded; and (2) they raise the support beam to permit the bifilar filaments to be disconnected, and then lower the support beam until the pivot shaft contacts and is supported by the weight load cell. Immobilization of the table is desirable to prevent accidental oscillation, and possible damage therefrom to the test object, during loading and unloading. A measure of protection against over-stressing the load cells is afforded by immobilizing the load beam during loading and unloading.

A monorail extends in both directions from the center of the upper platform to support a 1-ton capacity air hoist that is used to pick up, position,

and set down the test object on the table. Accurately located alignment holes are provided in the tables of both the high and low ranges to facilitate test object positioning. If "G" loading is extremely critical, as it is on some of the spacecraft, a Hydra-set may be used between the air hoist and the spacecraft to permit very gentle vertical positioning of the spacecraft.

A removable ladder and work platform are provided to permit ready access for inspection of the upper bearing housings and to facilitate the repositioning of the bifilar filaments when changing from one range configuration to the other. In normal operation the ladder and work platform remain attached to the facility structure. If an oversized test item, or one with appendages that would otherwise be obstructed by the platform and/or ladder, should require the measurement of mass characteristics, then the ladder and platform can be removed to permit a larger envelope to be tested. The self contained air hoist can be used to lower the work platform to floor level after it is unbolted from the structural frame.

The instrumentation is housed in a desk type console shown in Figure 6. The uppermost piece of rack mounted equipment is a Revere Corporation of America manual balance load indicator with a 3 cell selector switch. The load indicator provides digital readout of the selected load cells. Capacities of the load cells in the high range are: (1) weighing load cell—1200 pounds; and (2) moment load cells (2 load cells)—120 pounds each. The low range capacities are: (1) weighing load cell—240 pounds; and (2) moment load cells (2 load cells)—60 pounds each. The load cells are all Revere Corporation of America universal cells that were calibrated as a system to an accuracy of within ± 0.1 percent of applied load down to 0.1 capacity and ± 0.1 percent of 0.1 capacity load down to zero. Since the capacities of the various load cells are all divisible into 1200 (the capacity of the largest load cell), this was used as the basis for the digital readout of the load indicator. For the 1200 pound capacity load cell, the indicated force is the indicated reading. For the 240 pound capacity load cell, the indicated force is the indicated reading divided by 5. For the 120 pound and 60 pound capacity load cells, the indicated force is obtained by dividing the indicated readings by 10 and 20 respectively.

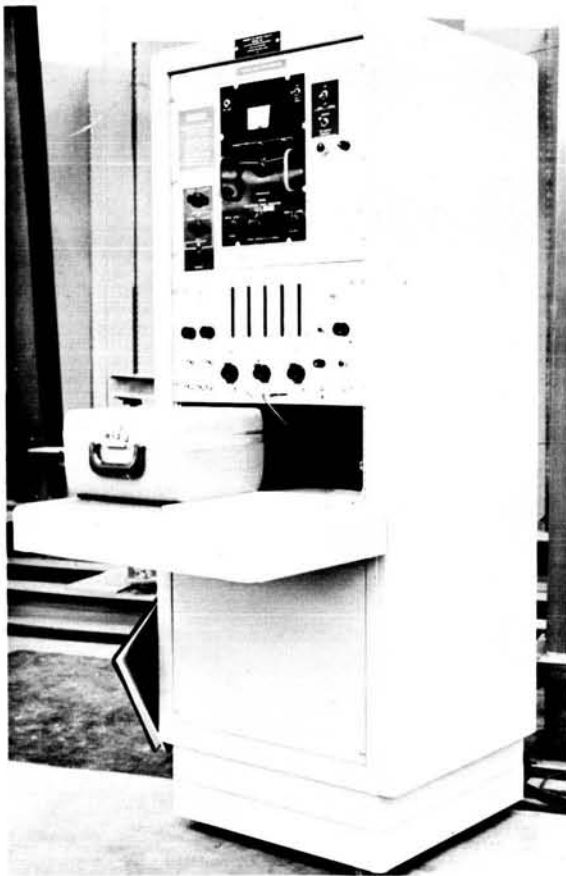


FIGURE 6.—Instrument console.

The other rack mounted instrument is a Hewlett-Packard Model 522BR electronic counter. Its function is to provide an accurate measure of the period of oscillation for a particular test item configuration during moment of inertia operation. During oscillation, a beam of light incident upon a photoelectric cell is periodically interrupted. Signals from the photoelectric cell are sent to the Hewlett-Packard counter. The resulting time lapse between signals represents the half period of oscillation for the test configuration.

The portable unit shown on the desk of the console is a Sanborn Model 301 single channel recorder. It is used in conjunction with the moment load cells to detect the presence of product of inertia. The test configuration is oscillated as in the moment of inertia measurement, except now first one and then the other moment load cell is switched to the recorder.

Any dynamic unbalance forces (product of inertia) on the table will be picked up by either or both of the two moment load cells and registered as a sine wave on the Sanborn recorder.

CALIBRATION AND ACCEPTANCE TESTING

During fabrication and preliminary calibration and checkout of the facility at the contractor's plant, no major problems were encountered. The facility was delivered to Goddard Space Flight Center in September 1962. Erection, alignment and preliminary calibration were completed in early October. Acceptance testing, using steel calibration slugs of simple geometric shapes that had been dimensionally measured and weighed and properly certified by the National Bureau of Standards, was begun. The acceptance testing consisted of demonstrating specification accuracies in weight, center of gravity displacement, moment of inertia, and product of inertia at any three points in each range designated by the Goddard technical representative. Acceptance test points were selected near the bottom of the weight and moment of inertia limit, approximately mid-range, and near the upper limit in both the high and low ranges. This selection provided evidence of performance at both extremes and an intermediate point in each range of operation.

Reduction of the acceptance test data disclosed an inconsistent behavior of the facility in measuring weight of the various test items. Other measurements were within the required accuracies. Additional tests and investigations were made to determine the cause of the irregularities. It was found that a preload existed in the fit of the linear ball bushing spline shaft of the lower pivot bearing assembly. The assembly was dismantled by the contractor and returned to his plant where it was reworked to provide a free fitting assembly with no bearing preload. The facility was then reassembled, realigned, and preliminary tests were conducted to assure consistent measurement behavior. It was mutually agreed by the contractor and the Goddard technical representative that center of gravity accuracy would not have been affected in any way by the rework of the pivot bearing assembly, so retesting of this parameter was waived. Acceptance tests of

weight, moment of inertia and product of inertia were reconducted at all three points of both the high and low ranges. Completion of these tests was accomplished in early December 1962. All measured mass characteristics were found to be within specified accuracies.

Some additional testing was done by Goddard personnel to become fully competent in the use of the facility. During this testing, it was found that the center pivot bearing assembly now appeared to have excessive lateral motion. The contractor was called in to examine the facility and formal acceptance was held up pending the findings of the examination. Actual measurements by the contractor showed lateral motion in excess of 0.004 inch, which was double the amount he deemed necessary for required accuracies on test items with centers of gravity elevated to a maximum height of 6 feet above the support table. The pivot bearing assembly was again removed and this time it was redesigned by the contractor. The splined shaft linear ball bushing was replaced with a ground cylindrical shaft linear ball bushing of precision fit. The diameter of the shaft was increased slightly to provide a greater margin of protection against failure due to bending caused by severe center of gravity offsets.

The modified pivot shaft assembly was returned to Goddard and was reinstalled in the facility. Following the usual realignment and preliminary checkout procedure, acceptance testing was begun anew. Following the now familiar pattern, weight, moment of inertia and product of inertia accuracies were demonstrated at three points in both the high and low ranges. Again, center of gravity testing was waived because it was not deemed to have been affected by the modifications to the center pivot assembly. Again, the accuracies of all measured test configurations were found to be within acceptable limits and the lateral motion of the pivot shaft was also acceptable. The facility was finally declared to be operational in June 1963.

A new obstacle soon loomed in the path of the now smooth sailing operation. A Goddard space shuffle required that the facility be moved. After these many months of striving for an acceptable facility and finally attaining the goal, now this! The only practical solution to retain a valid warranty was to have the move and proof

of accuracy performed by the original contractor. This was accomplished in August 1963 without incident, and the facility has been fully operational since that time.

OPERATIONAL EXPERIENCE

During the period since the facility was first fully accepted in June 1963, a considerable experience log has been compiled as shown in Table 1. The S-48 Flight Unit #1 spacecraft moment of inertia was remeasured about the pitch and yaw axes at the request of the project personnel. The first measurements, performed on August 15, 1963 were: pitch, 1.616 slug-ft²; and yaw, 1.616 slug-ft². When remeasured on November 6, 1963, the values obtained were: pitch, 1.608 slug-ft²; and yaw, 1.614 slug-ft². The measured values repeated within 0.008 slug-ft², which is within the specification accuracy of the facility.

TYPICAL TEST SEQUENCE

A typical test sequence will be described for the measurement of weight, center of gravity, and moment of inertia of a sample spacecraft. Assume that the calculated weight of the spacecraft is 300 pounds and the calculated moment of inertia about the desired axis of measurement (pitch) is 24 slug-ft². The center of gravity is calculated to be 20.000 inches forward of the interface. The spacecraft has previously been statically and dynamically balanced (on a spin balance facility) about the roll axis. The longitudinal center of gravity location, therefore, is the only one required.

The high range configuration is required since the weight of the spacecraft is between 160 pounds and 1000 pounds. This configuration uses a 1200 pound capacity weighing load cell and 2 moment load cells having a capacity of 120 pounds each. The high range table and the appropriate bifilar filament spacing are also required. The table is checked with a precision level to be sure the surface is level to within 0.0005 inch per foot. The load indicator and the period timer are turned on and permitted to warm up for a minimum of 30 minutes prior to starting the test.

Assume, also, that a simple fixture is required to support the spacecraft on the facility. In the case of this spacecraft, the fixture consists of a simple plywood cradle of symmetrical design that

TABLE 1.—Measurements Performed on Bifilar Facility

Date	Spacecraft or Sounding Rocket	Measurements
7-30-63	S-52 Thermal Vacuum Simulator	Wt., c.g., & M.I.
8- 5-63	S-52 Prototype	Wt., c.g., & M.I.
8- 9-63	S-52 Flight Unit	Wt., c.g., & M.I.
8-12-63	S-52 Separation Unit	M.I.
8-15-63	S-48 Flight Unit #1	Wt., c.g., & M.I.
9- 9-63	S-48 Flight Unit #2	Wt., c.g., & M.I.
11- 6-63	S-48 Flight Unit #1	M.I.
1- 4-64	San Marco—Shot Put 8	Wt. & M.I.
1- 6-64	NASA 8.31 Sounding Rocket	c.g. & M.I.
1-28-64	OAO-WEP (Prime Exper. Pkg.)	Wt., c.g., & M.I.
1-31-64	OAO-WEP (Control Elec. Pkg.)	Wt., c.g., & M.I.
2- 5-64	OSO—Balloon Gondola Simulator	Wt. & M.I.
2-26-64	NASA 8.28 Sounding Rocket	Wt., c.g., & M.I.
3-17-64	NASA 14.29 Sounding Rocket	Wt., c.g., & M.I.
3-18-64	AOSO—Balloon Gondola	Wt. & M.I.
4- 7-64	NASA 8.20 Sounding Rocket	Wt. & M.I.

has previously been statically balanced about its geometric axes of symmetry. The fixture weight has been accurately measured and found to be 6.00 pounds.

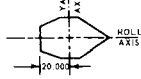
The no load indicated readings of the moment load cells are determined and recorded on line 1 of the Center of Gravity Determination Data Sheet shown as Figure 7. For convenience, the compression readings are set to zero and the tension readings are recorded as 75.6 for Cell No. 2 and 44.2 for Cell No. 3 for no load balanced conditions. A no load zero setting is also made for the weighing load cell (Cell No. 1). The fixture is now placed on the table in the properly oriented position. It is precisely positioned such that its center of gravity coincides exactly with the center of the table. This is easily and quickly accomplished by using the load indicator as a positioning meter. When the fixture is positioned as stated, no unbalanced moment is exerted and the indicated load cell readings are the same for the moment load cells as were the no load indicated readings.

The hydraulic jacks are now raised against the underside of the support beam to immobilize the table and prevent it from being accidentally moved while the spacecraft is being positioned. The air hoist is then used to lift the spacecraft

from its handling dolly, move it along the mono-rail to a position over the fixture, and lower it gently into the fixture with its pitch axis vertical

DATA SHEET
1000 LB. MOMENT OF INERTIA FACILITY
CENTER OF GRAVITY DETERMINATION - HIGH RANGE

PROJECT NAME	DATE
XYZ SPACE EXPLORER	3-16-64
NUMBER	
	PROJECT DESIGNATION ORBITAL SPACECRAFT + FIXTURE



	CELL 2 (Blue)		CELL 3 (Yellow)	
	TEN. (+)	COMP. (-)	TEN. (+)	COMP. (-)
ZERO SET:				
1. NO LOAD INDICATED READING	75.6	0.0	44.2	0.0
2. NO LOAD INDICATED FORCE (Item 1)	7.56	0.0		0.0
3. CALIBRATION CORRECTION (Reverse)	- .01			0.0
4. NO LOAD CORRECTED FORCE (Item 2 + 3)	7.55			0.0
LOADED CONDITIONS:				
5. INDICATED READING	146.4			0.3
6. INDICATED FORCE (Item 5/10)	14.64			0.03
7. CALIBRATION CORRECTION (Reverse)	- .02			0.00
8. CORRECTED FORCE (Item 6 + 7)	14.62			0.03
9. NO LOAD CORRECTED FORCE (Item 4)	7.55			0.0
10. MOMENT FORCE (Item 8 - 9)	7.07			0.03
11. CENTER OF GRAVITY DISPLACEMENT (S - Item 10) (Weight)	0.228			0.001
	S = 9.853		S = 9.903	

FIGURE 7.—Center of gravity data sheet.

and with the interface positioned 20.000 inches from the center of the table. Readings of Load Cell No. 2 and Load Cell No. 3 are made and recorded on line 5 of the Center of Gravity Determination Data Sheet.

Before proceeding further with calculations of center of gravity displacement, it is necessary to determine the weight of the spacecraft. The load indicator selector switch is turned to Cell No. 1. With the support beam supported on hydraulic jacks, the bifilar filaments are disconnected from the support beam by removing the cap screws that attach the lower bearing housings to the support beam as shown in Figure 5. The filaments are then pulled aside and attached by a light chain to the structural leg assembly. While the support beam is still supported on the hydraulic jacks, the load cell balance indicator is set for zero at no load. The hydraulic jacks are slowly retracted to permit the pivot shaft to be fully supported on the compression button of the weight load cell. The load indicator is manually balanced and the indicated reading is entered as Item 1, Column 1 of the Weight Determination Data Sheet as shown in Figure 8. The hydraulic jacks are again extended just enough to remove the load from the load cell. The no load zero setting is checked. The jacks are retracted to permit the load to again be supported on the load cell. The load indicator is again manually balanced and the indicated reading is entered in Column 2, Item 1 of the Weight Determination Data Sheet. The process is repeated twice more and the indicated readings are entered in Columns 3 and 4 of the Weight Determination Data Sheet.

The weight data reduction consists simply of completing Items 2 through 7 on the Data Sheet. These steps involve nothing more than applying the proper calibration correction for the indicated weight, subtracting the tare weight of the system, and taking the average of the calculated values.

Having measured the weights, the center of gravity data reduction can now be completed. Returning to the Center of Gravity Determination Data Sheet (Figure 7), it is noted that the loaded condition indicated reading (Line 5) increased in tension for Cell No. 2, but remained substantially unchanged from the no load reading for Cell No. 3. This indicates the center of gravity is not dis-

DATA SHEET				
1000 LB. MOMENT OF INERTIA FACILITY				
WEIGHT DETERMINATION - HIGH RANGE				
PROJECT NAME	XYZ SPACE EXPLORER			DATE
NUMBER	PIECE DESIGNATION			3-16-64
ORBITAL SPACECRAFT + FIXTURE				
	CELL 1	CELL 1	CELL 1	CELL 1
1. INDICATED READING	462.7	462.8	462.6	462.7
2. INDICATED WEIGHT*	462.7	462.8	462.6	462.7
3. CALIBRATION CORRECTION (MRL Figure 4)	+1.12	+1.12	+1.12	+1.12
4. CORRECTED WEIGHT (Item 2 + 3)	463.82	463.92	463.72	463.82
5. TARE WEIGHT	157.67	157.67	157.67	157.67
6. TEST ITEM WEIGHT (Item 4 - 5)	306.15	306.25	306.05	306.15
7. AVERAGE	306.15			
FIXTURE WEIGHT	6.00			
TEST ITEM ONLY WEIGHT	306.15			

FIGURE 8.—Weight determination data sheet.

placed laterally from the center of the spacecraft as should be expected for a statically balanced spacecraft. The indicated reading of 146.4 in tension for Cell No. 2 shows that Cell No. 2 has been subjected to a tensile force caused by an unbalanced moment acting along the longitudinal axis of the spacecraft. Applying the necessary corrections and following the step-by-step computations outlined on the Center of Gravity Determination Data Sheet leads to the actual longitudinal center of gravity displacement (Line 11, Cell No. 2) of 0.228 inch from the expected position.

The spacecraft can now be repositioned longitudinally to place the center of gravity directly over the center of the table, or an appropriate correction for the transfer of axis can be made for moment of inertia measurement. For the purpose of this discussion, it will be assumed that the spacecraft will be repositioned to place the center of gravity directly over the table center and coincident with the axis of oscillation when oscillated as a bifilar pendulum.

Before moment of inertia measurement is made, the lower ends of the bifilar filament bearing

housings are attached to the support beam with cap screws. This is accomplished with the support beam raised slightly above its normal elevation with the hydraulic jacks. After securely tightening the cap screws to attach the filaments to the support beam, the jacks are retracted, permitting the support beam, table, and spacecraft to be suspended from the bifilar filaments. The cables are disconnected from the moment load cells to eliminate any drag effect they might create. The photoelectric cell light source is turned on. Moment of inertia measurements are ready to be made.

Torsional oscillations are gradually built up by hand until the angular displacement exceeds 20 degrees. When the displacement has decayed to 20 degrees, period measurement is begun. This consists of measuring the time interval of the periodic interruptions of the light beam incident upon a photoelectric cell. For this test, assume a timer reading of 1.3780. It is entered on Line 2 of the Moment of Inertia Determination Data Sheet as shown in Figure 9. This procedure is repeated until three runs have been made, beginning each time with an initial angular

displacement of 20 degrees. Determination of the average period (Line 4) is made from the timer readings. The moment of inertia is then computed by the step-by-step operations outlined in Lines 5 to 11 of the Data Sheet.

The constants used in Lines 7 and 8 are determined from the geometry of the pendulum system for each range. The basic equation for moment of inertia about the axis of oscillation for a bifilar pendulum is given by:

$$I = \frac{WD^2T^2}{16\pi^2L} \quad (1)$$

where D = distance between the bifilar filaments

W = weight of the pendulum system

T = period of oscillation

L = length of the bifilar filaments

Equation 1 does not consider the reduction in period of oscillation due to torsion in the bifilar filaments. A torsional correction must, therefore, be introduced into the equation. It is represented by:

$$I_{\text{Torsion}} = 2 \left[\frac{GJT^2}{4\pi^2L} \right] = \frac{GJT^2}{2\pi^2L} \quad (2)$$

where G = modulus of rigidity of the filaments

J = polar moment of inertia of the filaments

T = period of oscillation

L = length of the filaments

The measured moment of inertia includes both the moment of inertia of the test object and the tare moment of inertia. The final equation for determining moment of inertia of the test object alone is thus reduced to:

$$I = \frac{WD^2T^2}{16\pi^2L} + \frac{GJT^2}{2\pi^2L} - I_{\text{Tare}} \quad (3)$$

Equation 3 assumes a frictionless pivot bearing assembly with no reduction in the speed of oscillation attributed thereto. Realistically, however, friction in the bearing assembly and other unaccountable losses tend to produce a longer period than is produced by the mass properties of the pendulum system, thus tending to degrade system accuracy. A calibration coefficient, K , (Line 8, Figure 9) has been determined for each range of the facility by a plot of actual vs. mea-

DATA SHEET			
1000 LB. MOMENT OF INERTIA FACILITY			
MOMENT OF INERTIA DETERMINATION - HIGH RANGE			
PROJECT NAME XYZ SPACE EXPLORER	PIECE DESIGNATION ORBITAL SPACECRAFT + FIXTURE		DATE 3-16-64
NUMBER PITCH AXIS			
	RUN 1	RUN 2	RUN 3
1. INITIAL ANGULAR DISPLACEMENT	20°	20°	20°
2. TIMER READING	1.3780	1.3779	1.3781
3. PERIOD (2 × Item 2)	2.7560	2.7558	2.7562
4. AVERAGE		2.7560	
5. (Item 4) ²		7.5955	
6. TEST ITEM WEIGHT + 165.00 LBS. 306.15		471.15	
7. UNCORRECTED M OF I (Item 5) [0.009326154 (Item 6) + 0.2754895]		35.4672	
8. CALIBRATION COEFFICIENT		0.9703	
9. CORRECTED M OF I (Item 7 × 8)		34.414	
10. TARE M OF I		10.049	
11. TEST ITEM M OF I (Item 9 - 10)		24.365	
FIXTURE M OF I		.162	
TEST ITEM ONLY M OF I		24.203	

FIGURE 9.—Moment of inertia determination data sheet.

sured moments of inertia. Equation 3 is further reduced to:

$$I = K \left[\frac{WD^2T^2}{16\pi^2L} + \frac{GJT^2}{2\pi^2L} \right] - I_{\text{Tare}} \quad (4)$$

Using the appropriate values for D, L, G, and J from the geometry of the pendulum system, and substituting into Equation 4 gives:

$$I = K (C_1WT^2 + C_2T^2) - I_{\text{Tare}} \quad \text{or}$$

$$I = KT^2 (C_1W + C_2) - I_{\text{Tare}}$$

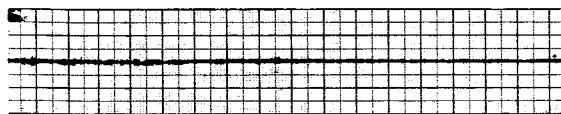
where $C_1 = 0.009326154$ for the high range

$C_1 = 0.004144957$ for the low range

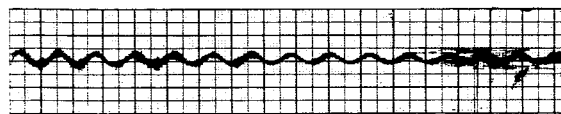
$C_2 = 0.2754895$ for either range

Product of Inertia measurement capability of a bifilar pendulum facility is by no means as sensitive as that of the more conventional spin balance facilities. Yet, it permits a reasonably sensitive determination of the dynamic unbalance of spacecraft that cannot, for various reasons, be balanced by spinning. Copies of typical traces of

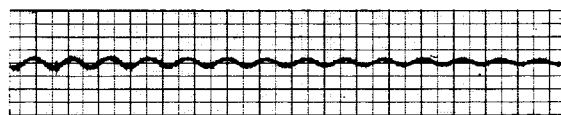
the dynamic forces imposed on the moment load cells are shown in Figure 10. The traces are similar for either range.



Dynamically Balanced Condition Recorded on Load Cell No. 3 (Plane of Roll-Pitch Axes)



Dynamic Unbalance Recorded on Load Cell No. 2 (Plane of Pitch-Yaw Axes)



Dynamic Unbalance Recorded on Load Cell No. 3 (Plane of Roll-Pitch Axes)

FIGURE 10.—Product of inertia traces.

TRIAxIAL BALANCING TECHNIQUES (A STUDY OF SPACECRAFT BALANCE WITH RESPECT TO MULTIPLE AXES)*

WILLIAM E. LANG

Spacecraft unbalance tolerance is likely to be expressed in terms of displacement of the center of gravity from its nominal position, and angular deviation between the principal and reference axes. This study discusses the relations that define the unbalance of a spacecraft, with respect to three mutually perpendicular reference axes, in terms of measurable mass parameters. It was motivated by the need to develop practical methods for balancing the San Marco spacecraft (the San Marco project is a joint effort of Italy and the United States). The theory proved directly and effectively applicable; but complications due to the inaccuracies of measured input data necessitated the development of modified methods for computing corrections.

INTRODUCTION

The meaning and significance of static and dynamic balance of spin-stabilized spacecraft is well defined, and systematic means of measuring and controlling it have been developed.† This kind of balancing, which may be called "single axis balancing," requires that a reference axis (the spin axis) be a principal axis of inertia of the spacecraft and that the spacecraft center of gravity be on this axis.

Some future spacecraft, because of active attitude control systems or other operational considerations, will require extension of this concept to include concurrent static and dynamic balancing with respect to three mutually perpendicular reference axes. This may be called "triaxial balancing."

Triaxial balancing requires that the center of gravity coincides with the intersection of three mutually perpendicular reference axes, and that these reference axes coincide with the principal axes of inertia of the spacecraft. Single axis balancing is necessary but not sufficient for triaxial balance. It is sufficient that static and dynamic balance exists about two of the reference axes—balance about the third axis is then automatic.

Under this concept of triaxial balancing measured values of the magnitude and phase of

static unbalance and of dynamic unbalance will be used to compute the correction needed to balance a spacecraft triaxially. Correction is assumed to be made by adding weight at the surface of a sphere of unit radius with its center at $x = y = z = 0$. The techniques described in this paper are limited to spacecraft which are essentially rigid, solid, and of constant mass.

DEFINITIONS AND COORDINATE SYSTEM CONVENTIONS

Assume a spacecraft of any size and shape with three reference axes, XX , YY , and ZZ , and any point in space, x , y , z , in a rectangular cartesian coordinate system. An imaginary sphere of radius R with its center at $x=0$, $y=0$, $z=0$, would have its surface defined by $x^2 + y^2 + z^2 = R^2$. The unbalance of the spacecraft can be defined in terms of the masses and locations of correction weights theoretically concentrated at points on the sphere surface in order to reduce the product of inertia about all three axes to zero.

This study indicates that either five or three such weights are necessary. More than five could be resolved into five or three; less than three would suffice only in special cases. No rigorous proof is offered for this hypothesis, but it can be supported by considering imaginary situations, and the derivations to be given here that follow from it are logical and consistent.

*Published as *NASA Technical Note D-2144*, March 1964.
†Schaller, N. C., and Lewallen, J. M., "Methods of Expressing Mass Unbalance," *NASA Technical Note D-1446*, May 1963.

One weight would correct static unbalance, i.e., center of gravity displacement. It would be located at the sphere surface on a radial straight line from the actual center of gravity passing through the desired center of gravity, $x=y=z=0$. The other weights would be either one or two diametrically opposed equal pairs. The location of any point can be defined in terms of phase orientation with respect to two of the coordinate axes, plus radial distance from $x=y=z=0$. Static or dynamic unbalance about XX, YY, and ZZ can also be defined by a vector having magnitude in appropriate units and a phase orientation.

To use both rectangular and polar coordinates, we must relate the two systems by appropriate conventions, which are to some extent arbitrary. Suppose the sphere to be enclosed by a cube and the cube, unfolded, to appear as shown in Figure 1. The large dots designate the positive directions of axes passing through the centers of the six faces, and the curved arrows indicate phase angle convention. The tail of each arrow is at 0 degrees. The system follows trigonometric and right-hand vector rule convention; i.e., when one is looking along an axis toward the origin from a positive end

angles increase counterclockwise. The symbols α , β , and γ will be defined later.

The significance of unbalance phase must be discussed. Static unbalance phase means the angle at which correction by weight addition is indicated. Dynamic unbalance phase also means an angle at which correction by weight addition is indicated. However, this correlation has to be made with two weights 180 degrees apart. One system to eliminate ambiguity, is to express dynamic unbalance as a torque vector in a right- or left-hand convention. Instead of this, here the phase of dynamic unbalance will mean the angle at which the correction weight having a positive coordinate parallel to the reference axis should be located. For example, if dynamic unbalance of D_x/a needs weight W at $+x, +y, +z$ plus weight W at $-x, -y, -z$, then a is the phase about the XX axis of the weight at $+x, +y, +z$. (The weight at $-x, -y, -z$ will have a phase angle of $\alpha + 180$ degrees, but the dynamic unbalance phase, by convention, is α). At this point it is necessary to define the following symbols:

$D_x/a, D_y/b, D_z/\gamma$ = magnitude and phase of dynamic unbalances about axes XX, YY, ZZ after correction of all static unbalances about $x=y=z=0$,

I_x, I_y, I_z = spacecraft inertias about XX, YY, and ZZ axes,

I_{xy}, I_{yz}, I_{zx} = products of inertia with respect to indicated subscripts,

M = spacecraft weight,

N and x_N, y_N, z_N = mass and coordinates of static correction weight,

$S_x/a, S_y/b, S_z/c$ = magnitude and phase of static unbalance about axes XX, YY, and ZZ, respectively,

W, x, y, z and $W, -x, -y, -z$ = mass and coordinates of dynamic unbalance correction weights,

x_M, y_M, z_M = coordinates of the spacecraft center of gravity, before correction,

$\delta = \sqrt{x_M^2 + y_M^2 + z_M^2}$, the center of gravity displacement, before correction,

$\theta_x, \theta_y, \theta_z$ = allowable angular deviations between the principal axes and reference axes.

For the mathematical model defined by the foregoing, many relationships are more or less apparent, and the more trivial of these will be stated without explanation or proof.

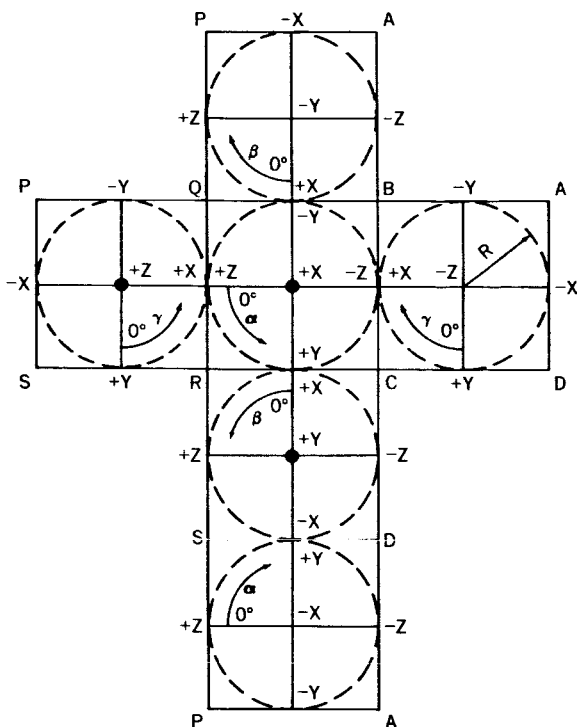


FIGURE 1.—Projection diagram of the coordinate system.

DEVELOPMENT OF SIGNIFICANT RELATIONSHIPS

Static Unbalance

The indicated correction N at x_N, y_N, z_N for center of gravity displacement is defined by $NR = M\delta$, $Nx_N = -Mx_M$, $Ny_N = -My_M$, and $Nz_N = -Mz_M$. From basic concepts or by definition:

$$S_x = N\sqrt{y_N^2 + z_N^2}, \quad (1)$$

$$S_y = N\sqrt{z_N^2 + x_N^2}, \quad (2)$$

$$S_z = N\sqrt{x_N^2 + y_N^2}, \quad (3)$$

$$\tan a = \frac{y_N}{z_N}, \quad (4)$$

$$\tan b = \frac{z_N}{x_N}, \quad (5)$$

$$\tan c = \frac{x_N}{y_N}, \quad (6)$$

$$x_N^2 + y_N^2 + z_N^2 = R^2, \quad (7)$$

$$RN = 0.707\sqrt{S_x^2 + S_y^2 + S_z^2}. \quad (8)$$

RN might be called the triaxial static unbalance. Of course R is an assigned value, which could be unity. From Equations 4, 5, and 6, $\tan a \tan b \tan c = 1$; therefore any two phase angles define the third. Equation 4 implies that $\sin a = y_N/\sqrt{y_N^2 + z_N^2}$; therefore $\sqrt{y_N^2 + z_N^2} = y_N/\sin a$. Similarly $\sqrt{z_N^2 + x_N^2} = z_N/\sin b$ and $\sqrt{x_N^2 + y_N^2} = x_N/\sin c$. Therefore Equations 1, 2, and 3 become:

$$N = \frac{S_x \sin a}{y_N} = \frac{S_y \sin b}{z_N} = \frac{S_z \sin c}{x_N}, \quad (9)$$

and N must be positive. Equations 4, 5, and 7,

or 5, 6, and 7, or 4, 6, and 7 can be solved simultaneously for $x_N, y_N,$ and z_N :

$$\pm x_N = y_N \tan c = \frac{z_N}{\tan b} = \frac{R \cos b \tan c}{\sqrt{\cos^2 b + \tan^2 c}}, \quad (10)$$

$$\pm y_N = z_N \tan a = \frac{x_N}{\tan c} = \frac{R \cos c \tan a}{\sqrt{\cos^2 c + \tan^2 a}}, \quad (11)$$

$$\pm z_N = x_N \tan b = \frac{y_N}{\tan a} = \frac{R \cos a \tan b}{\sqrt{\cos^2 a + \tan^2 b}}. \quad (12)$$

An alternative set of equations, easier to use, is:

$$N = \sqrt{\frac{1}{2}(S_x^2 + S_y^2 + S_z^2)}, \quad (13)$$

$$x_N = \sqrt{1 - \frac{S_x^2}{N^2}}, \quad (14)$$

$$y_N = \sqrt{1 - \frac{S_y^2}{N^2}}, \quad (15)$$

$$z_N = \sqrt{1 - \frac{S_z^2}{N^2}}. \quad (16)$$

Note that, in Equations 10-12 and 14-16, $x_N, y_N,$ and z_N are positive or negative as necessary to make N positive in Equation 9. This means that x_N has to have the same sign as $\sin c$, y_N the same sign as $\sin a$, and z_N the same sign as $\sin b$.

It may be preferable to use three weights, one on each reference axis, rather than N at x_N, y_N, z_N . (This would not introduce dynamic unbalance as a result of static unbalance correction.) The corrections at unit spherical radius would be

$$S_y \cos b = S_z \sin c \text{ at } x = +1, y = z = 0, \quad (17)$$

$$S_x \sin a = S_z \cos c \text{ at } y = +1, x = z = 0, \quad (18)$$

$$S_x \cos a = S_y \sin b \text{ at } z = +1, y = x = 0; \quad (19)$$

negative values imply equal positive values at x , y , or $z = -1$. The utility of Equations 9-19 depends on which of the parameters S_x , S_y , S_z , a , b , and c are known. In general, if enough parameters to physically define the situation are known, the rest can be calculated.

R , N , x_N , y_N , and z_N define a mass moment with respect to $x=y=z=0$ (i.e., RN) and the direction cosines of a straight line radiant from $x=y=z=0$. R is an arbitrary choice, and N varies inversely with R , but the radiant orientation is fixed and independent of R ; i.e., N must lie somewhere along this fixed radiant. Of course two (or more) weights could be used instead of N . Two such weights N_1 and N_2 could be at different radii along the radiant, such that $N_1R_1 + N_2R_2 = NR$. They could also be placed at the same radius R along two secondary radiants, each inclined at angle θ to the primary radiant, with the three radiants coplanar, such that $N_1 = N_2$ and $2N_1 \cos \phi = N$. Addition of vectors radiating from a point in three-dimensional space is inherently more complex than the addition of coplanar vectors, because each vector has three components rather than two, but the general procedures are similar.

Dynamic Unbalance

Once the spacecraft center of gravity is at $x=y=z=0$, (after addition of N) there is no static unbalance. Any remaining unbalance is dynamic, can be expressed as a mass moment couple, and must be corrected by a mass moment couple, or at least the correction must not reintroduce static unbalance. For the "spherical surface correction," this means W at x, y, z plus W at $-x, -y, -z$. W, x, y , and z are to be determined.

Note that D_x/α , D_y/β , and D_z/γ are dynamic unbalances *after* the correction of all static unbalance. This is because the correction N will contribute dynamic unbalance, unless it lies exactly on axis XX , YY , or ZZ . The contributions of dynamic unbalance about these axes due to N are:

$$N_{xx} = N_x \sqrt{y^2 + z^2} \left/ \tan^{-1} \frac{y}{z} + 180^\circ \right. \quad (20)$$

$$N_{yy} = N_y \sqrt{z^2 + x^2} \left/ \tan^{-1} \frac{z}{x} + 180^\circ \right. \quad (21)$$

$$N_{zz} = N_z \sqrt{x^2 + y^2} \left/ \tan^{-1} \frac{x}{y} + 180^\circ \right. \quad (22)$$

Note that the signs of x , y , and z need to be considered in defining phase angles; for example, although $+2/-2 = -2/+2 = -1$, $\tan^{-1}(+2/-2) = 135$ degrees and $\tan^{-1}(-2/+2) = 315$ degrees. The phase angles defined will be those of corrections having positive "parallel to reference axis" coordinates; for example, phase of unbalance about XX is that of a correction weight having a positive x coordinate. W, x, y , and z can be determined in terms of the parameters $D_x, D_y, D_z, \alpha, \beta$, and γ and an assigned value for R . Enough parameters to physically define a situation will suffice for a complete solution; for instance, the dynamic unbalance magnitude about one axis, plus the phase of unbalance about any two axes, is sufficient. Therefore any additional data would be redundant. This has significance for computer programming because input must be sufficient but must not be redundant, unless the program provides for using redundant data to compute alternative results. The added data would also be inconsistent insofar as it would not be 100 percent accurate.

From basic concepts or by definition:

$$\tan \alpha = \frac{y}{z} \quad (23)$$

$$\tan \beta = \frac{z}{x} \quad (24)$$

$$\tan \gamma = \frac{x}{y} \quad (25)$$

$$x^2 + y^2 + z^2 = R^2 \quad (26)$$

$$D_x = 2Wx \sqrt{y^2 + z^2} \quad (27)$$

$$D_y = 2Wy \sqrt{z^2 + x^2} \quad (28)$$

$$D_z = 2Wz \sqrt{x^2 + y^2} \quad (29)$$

Equations 23-25 and 27-29 can be combined into:

$$W = \frac{D_X \sin \alpha}{2xy} = \frac{D_Y \sin \beta}{2yz} = \frac{D_Z \sin \gamma}{2zx} \quad (30)$$

Equations 23-26 can be solved for x, y, and z. The results are:

$$\pm x = y \tan \gamma = \frac{z}{\tan \beta} = \frac{R \cos \beta \tan \gamma}{\sqrt{\cos^2 \beta + \tan^2 \gamma}} \quad (31)$$

$$\pm y = z \tan \alpha = \frac{x}{\tan \gamma} = \frac{R \cos \gamma \tan \alpha}{\sqrt{\cos^2 \gamma + \tan^2 \alpha}} \quad (32)$$

$$\pm z = x \tan \beta = \frac{y}{\tan \alpha} = \frac{R \cos \alpha \tan \beta}{\sqrt{\cos^2 \alpha + \tan^2 \beta}} \quad (33)$$

This solution, although correct, proved unsuitable for numerical calculation, because the trigonometric functions involved can have values

which are affected drastically by minor deviations of phase angles. A more useful solution, which involves the use of Equations 23-29, is:

$$|W| = \frac{D_X}{2x \sqrt{y^2 + z^2}} = \frac{D_Y}{2y \sqrt{z^2 + x^2}} = \frac{D_Z}{2z \sqrt{x^2 + y^2}} \quad (34)$$

$$|x| = \frac{1}{\sqrt{1 + \left(\frac{-D_X^2 + D_Y^2 + D_Z^2}{+D_X^2 - D_Y^2 + D_Z^2} \right) + \left(\frac{-D_X^2 + D_Y^2 + D_Z^2}{+D_X^2 + D_Y^2 - D_Z^2} \right)}} \quad (35)$$

$$|y| = \frac{1}{\sqrt{1 + \left(\frac{+D_X^2 - D_Y^2 + D_Z^2}{-D_X^2 + D_Y^2 + D_Z^2} \right) + \left(\frac{+D_X^2 - D_Y^2 + D_Z^2}{+D_X^2 + D_Y^2 - D_Z^2} \right)}} \quad (36)$$

$$|z| = \frac{1}{\sqrt{1 + \left(\frac{+D_X^2 + D_Y^2 - D_Z^2}{-D_X^2 + D_Y^2 + D_Z^2} \right) + \left(\frac{+D_X^2 + D_Y^2 - D_Z^2}{+D_X^2 - D_Y^2 + D_Z^2} \right)}} = 1 - x^2 - y^2 \quad (37)$$

The values for x, y, and z can be real only if the sum of any two of the quantities D_X^2 , D_Y^2 , and D_Z^2 exceeds the third quantity. Any real mass configuration will satisfy this condition in fact, but erroneous measurements may not. Methods for "normalizing" measured data to overcome this problem will be discussed later. The method of solving Equations 26-29 for Equations 35-37 is given in Appendix A.

To calculate W, x, y, and z from Equations 30-33 it is necessary to know the dynamic unbalance about one axis, plus the phase of unbalance about any two axes. As R may have any

convenient value, it is logical to let R equal one unit of length.

To use Equations 34-37 requires that dynamic unbalance about all three axes be known, but it is unnecessary to know any phase angles, except to define signs of W, x, y, and z. The signs of x, y, and z must define the unique location of a positive or negative W, and require a convention. Derivation of the convention depends on the established system geometry, and the established dynamic balance phase convention.

From Equation 30 it is evident that the expressions $D_X \sin \alpha / 2xy$, $D_Y \sin \beta / 2yz$ and

$D_z \sin \gamma / 2zx$ must all yield W as either positive or negative. A positive W means correction by weight addition (W at x, y, z and W at $-x, -y, -z$). A negative W means that correction could consist of weight removal, i.e., removing W at

x, y, z and at $-x, -y, -z$. However, this is not a practical correction means; therefore derivation of an equivalent correction by weight addition is necessary. With a three-dimensional model it can be demonstrated that

$-W, +x, +y, +z$	is equivalent to	$+W, -x, +y, +z$	or	$+W, +x, -y, +z$	or	$+W, +x, +y, -z$
$-W, -x, -y, -z$		$+W, +x, -y, -z$		$+W, -x, +y, -z$		$+W, -x, -y, +z$
		$+2W, 0, +y, -z$		$+2W, +x, 0, -z$		$+2W, +x, -y, 0$
		$+2W, 0, -y, +z$		$+2W, -x, 0, +z$		$+2W, -x, +y, 0$

It is now apparent why triaxial balancing requires either three or five weights. Static unbalance requires on weight N . Dynamic unbalance requires two weights if W from Equation 30 is positive; if W is negative the equivalent correction by weight addition requires four weights.

It will be shown that the sign of W , and therefore the number of correction weights, can be predicted from any two dynamic unbalance phase quadrants.

Figure 2 represents the octants of a sphere with the signs of $x, y,$ and z and the trigonometric quadrants of $\alpha, \beta,$ and γ . Data for each octant can be tabulated as in Table 1. By abstracting

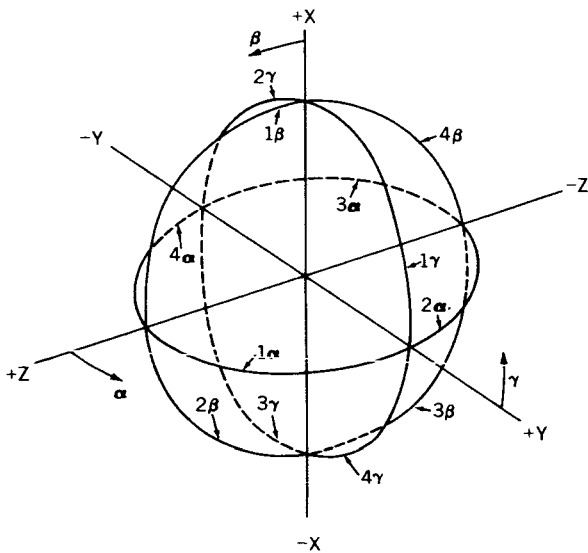


FIGURE 2.—Isometric representation of the coordinate system.

significant data from Table 1 and considering the implication of a negative W , Tables 2 and 3 may be derived.

In Table 1, A through H represent the octants of the sphere, and W must be defined in two diametrically opposed octants. Data for octants E-H apply to the $-x$ hemisphere. The second column gives the signs of $x, y,$ and z . Column 3 shows whether a positive W (correction by weight addition) or a negative W (correction by weight removal) is to be defined. Column 4 shows the quadrant of the phase of dynamic unbalance (about each axis); correction in the corresponding octant is implied. This is the quadrant of $\alpha, \beta,$ or γ in the octant *only* if the octant is on the *positive* side of the $x, y,$ or z axis. Because of the dynamic unbalance phase convention, if the octant is on the *negative* side of the unbalance phase quadrant is diametrically opposed to the octant phase quadrant; such cases are noted by the circle around the quadrant number. The concept involved requires three-dimensional visualizing to define the unbalance phase quadrant in each specific case, and column 4 lists the results of this exercise.

For convenience and emphasis the dynamic unbalance phase convention will now be repeated: α is the phase of a light spot with a positive x coordinate about axis XX ; β is the phase of a light spot with a positive y coordinate about axis YY ; γ is the phase of a light spot with a positive z coordinate about axis ZZ .

Since column 2 includes all octants, column 4 includes *all physically possible* $\alpha, \beta,$ and γ phase

Table 1
Spherical Octant Correction Signs.

Octant	Coordinate x y z	Positive or Negative Correction	Unbalance Phase Quadrant a β γ	Sine a β γ	Coordinate Product xy yz zx	W		
						$\frac{D_x \sin \alpha}{2xy}$	$\frac{D_y \sin \beta}{2yz}$	$\frac{D_z \sin \gamma}{2zx}$
A	+ + +	+	1 1 1	+ + +	+ + +	+	+	+
			3 3 3	- - -	+ + +	-	-	-
B	+ + -	+	2 4 ③	+ - -	+ - -	+	+	+
			4 2 ①	- + +	+ - -	-	-	-
C	+ - -	+	3 ② ④	- + -	- + -	+	+	+
			1 ④ ②	+ - +	- + -	-	-	-
D	+ - +	+	4 ③ 2	- - +	- - +	+	+	+
			2 ① 4	+ + -	- - +	-	-	-
E	- - -	+	① ① ①	+ + +	+ + +	+	+	+
			③ ③ ③	- - -	+ + +	-	-	-
F	- - +	+	② ④ 3	+ - -	+ - -	+	+	+
			④ ② 1	- + +	+ - -	-	-	-
G	- + +	+	③ 2 4	- + -	- + -	+	+	+
			① 4 2	+ - +	- + -	-	-	-
H	- + -	+	④ 3 ②	- - +	- - +	+	+	+
			② 1 ④	+ + -	- - +	-	-	-

quadrant combinations. Octants E-H are included for completeness but it is sufficient to define one of the two weights W, and the one in the +x hemisphere can be chosen. In other words x can be considered always positive, by convention.

Table 2 includes all physically possible dynamic unbalance phase quadrant combinations. There are 8 of these: 111, 243, 324, and 432, which imply correction by two weights, and 142, 214, 333, and 421, which imply correction by four weights, placed as indicated in Table 3. Note

that if the quadrant of a is known, there are only two possible quadrants for β and these are always adjacent quadrants. Also, a and β determine γ since $\tan \alpha \tan \beta \tan \gamma = 1$.

Column 5 of Table 1 shows the sign of $\sin \alpha$, $\sin \beta$, and $\sin \gamma$ corresponding to the quadrant tabulation of column 4. Column 6 gives the signs of the products xy , yz , and zx based on the column 2 signs for x , y , and z . Column 7 tabulates the signs of W, determined from Equation 30, by using xy , yz , and zx from column 6 and

Table 2

Corrections Based on Table 1.

Unbalance Phase Quadrant			Coordinates for Corrections W			W	Correction				
α	β	γ	x	y	z			$-x$	$-y$	$-z$	
1	1	1	+	+	+	-	-	-	+	Add as indicated (2 weights) Use Table 3 (4 weights)	
1	4	2	+	-	-	-	+	+			-
2	4	3	+	+	-	-	-	+	+	Add as indicated (2 weights) Use Table 3 (4 weights)	
2	1	4	+	-	+	-	+	-			-
3	2	4	+	-	-	-	+	+	+	Add as indicated (2 weights) Use Table 3 (4 weights)	
3	3	3	+	+	+	-	-	-			-
4	3	2	+	-	+	-	+	-	+	Add as indicated (2 weights) Use Table 3 (4 weights)	
4	2	1	+	+	-	-	-	+			-

Table 3

Equivalent Corrections by Weight Addition for Negative W Cases in Table 2.

Correction Choice	Quadrant			W			W'		
	α	β	γ	x	y	z	x'	y'	z'
A*	1	4	2	+	+	+	0	+	-
	2	1	4	+	+	-	0	+	+
	3	3	3	+	-	-	0	+	-
	4	2	1	+	-	+	0	+	+
B†	1	4	2	+	+	-	+	0	+
	2	1	4	+	+	+	+	0	-
	3	3	3	+	-	+	+	0	-
	4	2	1	+	-	-	+	0	+
C‡	1	4	2	+	-	+	+	+	0
	2	1	4	+	-	-	+	+	0
	3	3	3	+	+	-	+	-	0
	4	2	1	+	+	+	+	-	0

* $W' = 2W(y^2 + z^2)$, $|y'| = y/\sqrt{y^2 + z^2}$,

$|z'| = z/\sqrt{y^2 + z^2}$.

† $W' = 2W(z^2 + x^2)$, $|x'| = x/\sqrt{z^2 + x^2}$,

$|z'| = z/\sqrt{z^2 + x^2}$.

‡ $W' = 2W(x^2 + y^2)$, $|x'| = x/\sqrt{x^2 + y^2}$,

$|y'| = y/\sqrt{x^2 + y^2}$.

$\sin \alpha$, $\sin \beta$, and $\sin \gamma$ from column 5. The fact that the results uniformly conform to column 3 verifies the entire table. Note that D_x , D_y , and D_z are positive by definition.

APPLICATION OF GENERAL PRINCIPLES TO THE PROBLEM OF BALANCING A SPACE-CRAFT TRIAXIALLY

Expressing Triaxial Unbalance and Unbalance Tolerance

Just as single unbalance tolerance is often expressed as center of gravity displacement and principal axis tilt, triaxial unbalance tolerance basically limits center of gravity displacement (from a point), i.e., δ , and tilt of all three principal axes (θ_x , θ_y , θ_z). The values θ_x , θ_y , and θ_z are likely to be equal, or, in some cases, two of them may be equal and the third considerably smaller. Analytically, all three axes have identical significance, hence the triangular symmetry of the preceding derivations. In practice, unbalance tolerances may define a nonspherical volume limit for the center of gravity and principal axis limits within volumes other than circular cones.

In any event, in expressing static unbalance magnitude:

$$\delta = \frac{\sqrt{S_x^2 + S_y^2 + S_z^2}}{M \sqrt{2}} = \frac{NR}{M} \quad (38)$$

where R is the assigned value. Note that R, N, and M define unbalance magnitude. It is necessary to have x_N , y_N , and z_N to define the direction of unbalance.

It is pertinent to distinguish between dynamic unbalance and product of inertia. Although dynamic unbalance is a product of inertia, and

product of inertia is a dynamic unbalance, there is a conceptual difference.

An unbalanced object, with reference axes XX , YY , and ZZ , has principal axes $X'X'$, $Y'Y'$, and $Z'Z'$ skewed θ_x , θ_y , and θ_z from the reference axes. Two of the principal axes are axes of maximum and minimum inertia, and the three principal axes are mutually perpendicular.

The product of inertia concept is that the mass of the body is projected onto a plane defined by two perpendicular axes, such as XX and YY . The product of inertia I_{XY} is then $\int xy \, dm$; i.e., the integral summation of mass elements times the product of their XY coordinates. Products of inertia about pairs of principal axes are all zero. I_{XY} , I_{YZ} , and I_{ZX} will have positive or negative values.

The dynamic unbalance concept, for dynamic unbalance D_x/a about axis XX , is that all the mass of the body is projected onto a plane XP defined by axis XX and an axis PP in the YZ plane oriented so that the product of inertia I_{XP} is maximum.

$|I_{XP}|$ is the magnitude D_x of dynamic unbalance, and the orientation of PP in the YZ plane defines the phase angle a of dynamic unbalance. If an axis QQ is perpendicular to PP and in the YZ plane, then I_{XQ} will be zero. Dynamic unbalance is zero about all three principal axes, and some positive value for any other axis. It is always positive because its phase angle defines vector direction—a negative value at phase angle a is stated as an equal positive value at phase angle $(a+180^\circ)$. A phase convention is essential, and so is a convention on whether “unbalance” means existing net mass moment (heavy spot) or indicated correction (light spot). In this paper, unbalance means indicated correction.

It is noteworthy that, although only the principal axes of zero dynamic unbalance, there is for any axis AA one perpendicular axis BB such that the product of inertia I_{AB} is zero.

The distinction between product of inertia and dynamic unbalance has been clarified at some length because it is rarely explicitly stated. For the conventions established, it is apparent that:

$$D_x = \sqrt{I_{XY}^2 + I_{XZ}^2} / \tan^{-1} \frac{I_{XY}}{I_{XZ}}, \text{ and } I_{XY} = D_x \sin a = D_y \cos \beta, \quad (39)$$

$$D_y = \sqrt{I_{YZ}^2 + I_{XY}^2} / \tan^{-1} \frac{I_{YZ}}{I_{XY}}, \text{ and } I_{YZ} = D_y \sin \beta = D_z \cos \gamma, \quad (40)$$

$$D_z = \sqrt{I_{ZX}^2 + I_{ZY}^2} / \tan^{-1} \frac{I_{ZX}}{I_{ZY}}, \text{ and } I_{ZX} = D_z \sin \gamma = D_x \cos a. \quad (41)$$

The principal axis tilt in the XY plane is $(1/2)\tan^{-1}(2I_{XY}/|I_{XX}-I_{YY}|)$ and, for small tilt angles (<10 degrees), it may be considered $I_{XY}/|I_{XX}-I_{YY}|$ radians. Therefore under the assumption of a tilt angle <10 degrees and no cosine error (cosine error would actually be minimal)

$$\theta_x \approx \sqrt{\left(\frac{I_{XY}}{I_{XX}-I_{YY}}\right)^2 + \left(\frac{I_{XZ}}{I_{XX}-I_{ZZ}}\right)^2} \quad (42)$$

Equation 42 may be considered exact for all practical purposes since θ_x would be less than 10

degrees. By using Equations 39 and 41 in Equation 42 and making some rearrangement,

$$\frac{\theta_x}{D_x} = \sqrt{\left(\frac{\sin a}{I_{XX}-I_{YY}}\right)^2 + \left(\frac{\cos a}{I_{XX}-I_{ZZ}}\right)^2}. \quad (43)$$

Similarly

$$\frac{\theta_y}{D_y} = \sqrt{\left(\frac{\sin \beta}{I_{YY}-I_{ZZ}}\right)^2 + \left(\frac{\cos \beta}{I_{YY}-I_{XX}}\right)^2}. \quad (44)$$

$$\frac{\theta_z}{D_z} = \sqrt{\left(\frac{\sin \gamma}{I_{ZZ}-I_{XX}}\right)^2 + \left(\frac{\cos \gamma}{I_{ZZ}-I_{YY}}\right)^2}. \quad (45)$$

Equations 43-45 relate the tilt of each principal axis from an adjacent reference axis to the corresponding dynamic unbalance about the reference axis. That is with I_{XX} , I_{YY} , I_{ZZ} , and D_X/α known θ_X can be calculated and compared to tolerance, as can θ_Y , with D_Y/β known, or θ_Z , with D_Z/γ known. How to establish limits for D_X , D_Y , and D_Z from a knowledge of θ_X , θ_Y , and θ_Z is less apparent, but can be done as follows: Study of a general function $\sin^2 e/A + \cos^2 e/B$ shows that the maximum limit for D_X in Equation 43 would occur when either $\sin \alpha=0$ and $\cos \alpha=1$, or $\sin \alpha=1$ and $\cos \alpha=0$. It follows that the limits for D_X , D_Y , and D_Z are:

$$D_X \leq (I_{XX} - I_*) \theta_X \quad (46)$$

$$D_Y \leq (I_{YY} - I_*) \theta_Y \quad (47)$$

$$D_Z \leq (I_{ZZ} - I_*) \theta_Z \quad (48)$$

I_* must, in each case, be chosen from I_{XX} , I_{YY} , and I_{ZZ} to make D_X , D_Y , or D_Z a minimum but not zero. Note that these equations are predicated on the most unfavorable phase orientations of α , β , and γ . Larger D_X , D_Y , and D_Z may or may not be compatible with θ_X , θ_Y , and θ_Z limits, and more than 41.4 percent larger would never be compatible.

Triaxial dynamic unbalance could be expressed as D_X/α , D_Y/β , D_Z/γ . This fully defines all parameters, and principal axes tilts can be determined if I_{XX} , I_{YY} , and I_{ZZ} are known, but there is no indication of the necessary corrections. If separate corrections were applied for D_X , D_Y , and D_Z , six weights would be needed, weight addition would be excessive, and the result would be that *dynamic balance would not exist about XX, YY, or ZZ*, because of interaction between the corrections. Also D_X/α , D_Y/β , D_Z/γ contains redundant data - D_X/α and D_Y/β fully define the unbalance.

A better way to express triaxial unbalance, both static and dynamic, is by N , x_N , y_N , and z_N and W , x , y , and z . This defines the situation and also the necessary correction. Equations 23-29 enable derivation of D_X/α , D_Y/β , and D_Z/γ from W , x , y , and z . It is to be understood that N and

W are the surface of a sphere of unit radius (and therefore x_N , y_N , z_N , x , y , and z will all be less than one) and that correction W is needed at $+x$, $+y$, $+z$, and also at $-x$, $-y$, $-z$. N is directly indicative of the magnitude of triaxial static unbalance (x_N , y_N , z_N define its direction) and W is indicative of the magnitude of triaxial dynamic unbalance, with x , y , z defining its direction; x_N , y_N , z_N , and x , y , z are direction cosines of the unbalance vectors.

It is noteworthy that, although $N=0.707$ $\sqrt{S_X^2+S_Y^2+S_Z^2}$, W is neither equal to 0.707 $\sqrt{D_X^2+D_Y^2+D_Z^2}$ nor related to it in any simple fashion. W is actually

$$\sqrt{\frac{D_X^2 + D_Y^2 + D_Z^2}{8(x^2y^2 + y^2z^2 + z^2x^2)}}$$

where $x^2+y^2+z^2=1$.

Computing the Effect of Changes to a Spacecraft

Suppose the balance condition of a spacecraft is N , x_N , y_N , z_N , W , x , y , z , and a weight C is added at p , q , r . The new balance condition can be defined as follows below: (For convenience the symbols d , e , f will be used in place of x_N , y_N , and z_N and N' , d' , e' , f' , W' , x' , y' , z' will mean the new balance condition.)

Note that W , x , y , z includes all dynamic balance effects of N , d , e , f , and since W , x , y , z implies W , $+x$, $+y$, $+z$, plus W , $-x$, $-y$, $-z$, it has no static balance effect. For consistency, addition of C at $+p$, $+q$, $+r$ can be coded as $+C$, $-p$, $-q$, $-r$, $-C$, $+p$, $+q$, $+r$, which is actually the correction needed by the unbalance due to C (i.e., add C at $-p$, $-q$, $-r$ and then deduct C at $+p$, $+q$, $+r$ and at $-p$, $-q$, $-r$).

The problem can be stated as follows:

$$(N, d, e, f, W, x, y, z) \text{ (known input)} \oplus (+C, -p, -q, -r, -C, +p, +q, +r) = (N', d', e', f', W', x', y', z')$$

meaning two known unbalances are to be superimposed to give a third in the same form of expression as the first two. The problem is how to perform the superposition operation \oplus . The

information content of N, d, e, f, W, x, y, z is

$$\text{INPUT A} = \left\{ \begin{array}{l} (S_x/\underline{a})_A = N \sqrt{e^2 + f^2} / \tan^{-1} \frac{e}{f} , \\ *(S_y/\underline{b})_A = N \sqrt{f^2 + d^2} / \tan^{-1} \frac{f}{d} , \\ (S_z/\underline{c})_A = N \sqrt{d^2 + e^2} / \tan^{-1} \frac{d}{e} , \\ *(D_x/\underline{a})_A = 2Wx \sqrt{y^2 + z^2} / \tan^{-1} \frac{y}{z} , \\ *(D_y/\underline{b})_A = 2Wy \sqrt{z^2 + x^2} / \tan^{-1} \frac{z}{x} , \\ (D_z/\underline{c})_A = 2Wz \sqrt{x^2 + y^2} / \tan^{-1} \frac{x}{y} , \\ d^2 + e^2 + f^2 = x^2 + y^2 + z^2 = R_A^2 = 1 \text{ (by definition) .} \end{array} \right.$$

Note that, since $R=1$ by convention, the information content of N, d, e, f, W, x, y, z could be represented by N, d, e, W, x, y . The information content of $+C, -p, -q, -r, -C, p, q, r$ is

$$\text{INPUT B} = \left\{ \begin{array}{l} *(S_x/\underline{a})_B = +C \sqrt{q^2 + r^2} / \tan^{-1} \frac{-q}{-r} \text{ and } (D_x/\underline{a})_B = 2p(S_x/\underline{a})_B \\ *(S_y/\underline{b})_B = +C \sqrt{r^2 + p^2} / \tan^{-1} \frac{-r}{-p} \text{ and } (D_y/\underline{b})_B = 2q(S_y/\underline{b})_B \\ (S_z/\underline{c})_B = +C \sqrt{p^2 + q^2} / \tan^{-1} \frac{-p}{-q} \text{ and } (D_z/\underline{c})_B = 2r(S_z/\underline{c})_B \\ p^2 + q^2 + r^2 = R_B^2 \text{ (which may have any value).} \end{array} \right.$$

Input B could be coded as $-C, p, q, r$. Note that for input B, $a=\alpha, b=\beta, c=\gamma$, and the value of R_B does not affect the fact that $(S_x/\underline{a})_B, (S_y/\underline{b})_B, (D_x/\underline{a})_B$, etc. are the corrections for $+C$ at $+p, +q, +r$; i.e., $-C$ at $+p, +q, +r$. It has been shown earlier that $-C, p, q, r$ may actually imply five individual weight additions.

Some of the information content of inputs A and B is not independent and therefore redundant,

and only the items marked with an asterisk are needed to perform the operation input A \oplus input B. This involves vector addition of corresponding elements of the inputs, as follows:

$$R = 1 \text{ (by definition) ,}$$

$$(S_x/\underline{a})_A \leftrightarrow (S_x/\underline{a})_B = (S_x/\underline{a})_C ,$$

$$(S_Y/b)_A \leftrightarrow (S_Y/b)_B = (S_Y/b)_C .$$

$$(D_X/a)_A \leftrightarrow (D_X/a)_B = (D_X/a)_C .$$

$$(D_Y/b)_A \leftrightarrow (D_Y/b)_B = (D_Y/b)_C .$$

This is sufficient information to derive *output C* in the form $N', d', e', f', W', x', y', z'$ (redundant components are not included). It has already been established that $(S_X/a)_C$ and $(S_Y/b)_C$ will yield N', d', e', f' for output C (Equations 9-16). Also $(D_X/a)_C$ and $(D_Y/b)_C$ will yield W', x', y', z' for output C (Equations 30-37). Therefore output C is obtained in the form $N', d', e', f', W', x', y', z'$, or can be coded as N', d', e', W', x', y' (since $R=1$).

The \oplus operation is quite tedious by manual methods, but would be trivial for a suitably programmed computer, which could readily evaluate the balance condition of a complicated assembly from any number of inputs coded in the form $-C, p, q, r$. It could also evaluate total spacecraft weight (ΣC) and inertias about reference axes $XX, YY, \text{ and } ZZ$, which are $\Sigma C(p^2+q^2)$, $\Sigma C(q^2+r^2)$, and $\Sigma C(r^2+p^2)$ except that the inertias would not include the body-centered inertias of components. The dynamic unbalance results would also fail to include the body-centered dynamic unbalance of components.

These body-centered mass parameters could be included in a computer program with little additional complication, and so could inputs coded in polar rather than rectangular coordinates.

Prerequisite Revision of Measured Unbalance Data

To solve for either static or dynamic correction it is necessary to compensate for the fact that the input data is interdependent, partially redundant, and to some degree inaccurate, and therefore inconsistent. The inconsistency comes from measurement inaccuracies and also from the fact that the axes which are *theoretically* orthogonal and intersecting are neither *perfectly* orthogonal nor *exactly* intersecting.

To obtain an optimized solution by the methods to be described requires first that the input data be "normalized." Normalizing makes all the input data compatible and also averages out

random errors. All raw data is assumed equally reliable, and for vectors this implies that the angular error in degrees is approximately equal numerically to 0.6 the magnitude error expressed as a percentage (for example, a 2 percent error in magnitude is approximately equivalent to a 1 degree error in phase angle.)

It has been shown that only certain combinations of unbalance phase angle quadrants are physically possible. It can also be derived that for static unbalances

$$S_X \cos a = S_Y \sin b .$$

$$S_Y \cos b = S_Z \sin c .$$

$$S_Z \cos c = S_X \sin a .$$

as given in Equations 17-19, and for dynamic unbalances

$$D_X \cos a = D_Z \sin \gamma . \quad (49)$$

$$D_Y \cos \beta = D_X \sin a . \quad (50)$$

$$D_Z \cos \gamma = D_Y \sin \beta . \quad (51)$$

These relationships would hold exactly for perfect data, but are met only imperfectly by empirical measurements. Therefore, normalizing data means adjusting it to conform to Equations 17-19 and 49-51. This is best illustrated graphically, but an equivalent algebraic method can easily be derived.

Figure 3 shows *exact* unbalance vectors S_X/a (OX), S_Y/b (OY), and S_Z/c (OZ) on a common

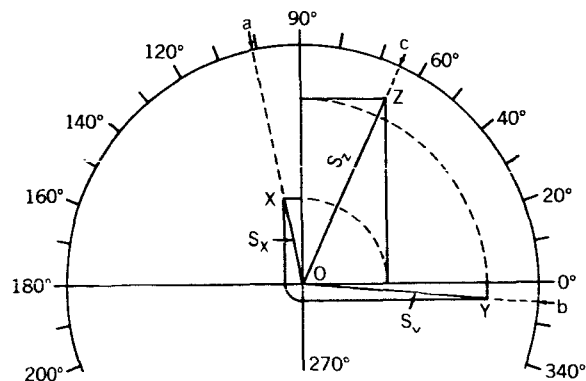


FIGURE 3.—Plot of exact unbalance vectors.

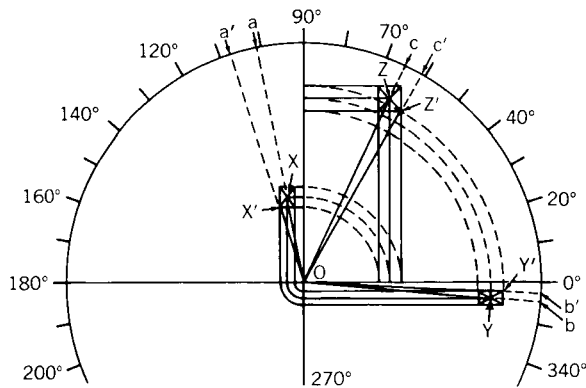


FIGURE 4.—Normalizing plot for inexact unbalance vectors.

polar plot. The rest of the diagram illustrates that these vectors conform to Equations 17–19.

Figure 4 shows, on a similar plot, measured values OX' , OY' , and OZ' of unbalance about the three axes. The rest of the diagram shows their nonconformity with the equations. The rectilinear projection of rotated components of the original vectors defines three rectangular areas, each having X' , Y' , or Z' as one corner, and X , Y , or Z as geometric center.

OX , OY , and OZ satisfy Equations 17–19 and are the normalized values of the original vectors OX' , OY' , and OZ' . All three original vectors are changed in both magnitude and phase. The normalized vectors are not necessarily exact, though they tend to define the overall condition better than the original measurements, but they are constrained to be consistent and therefore the calculation of unbalance correction is not complicated by multiple results based on the choice of input parameters from redundant data. Also, inaccurate measured data could appear to define a physically impossible situation (imaginary solutions of equations), whereas normalized data cannot.

The procedure for normalizing static unbalance data has been described. The procedure for dynamic unbalance data is the same, except that Equations 49–51 apply. In all cases measured data may be expected to conform approximately with the applicable equations. Extreme deviation would imply a major mistake rather than normal inaccuracy. With S_x/a , S_y/b , S_z/c , D_x/α , D_y/β , D_z/γ considered to be normalized

input data unbalance correction calculations can be made by using Equations 13–19 and 34–37.

Computing Dynamic Unbalance Corrections

An essential operation in determining dynamic unbalance correction is deriving a basic correction W , x , y , z from normalized data $D_x/\alpha D_y/\beta D_z/\gamma$. The symbol $\Delta_{\alpha\beta\gamma}^{xyz}$ will be used to denote this operation, which consists of solving Equations 34–37 for W , x , y , and z . In these equations x , y , and z can have real values only if the sum of any two of the quantities D_x^2 , D_y^2 , and D_z^2 exceeds the third quantity. This condition will be met by exact or normalized values, but may not be satisfied by measured values.

It is necessary to use previously derived tabulations to determine the signs of x , y , z , and W based on the quadrants of unbalance phase angles. Table 4, which is taken from Table 2, covers all cases for which W is positive.

Table 4

Coordinate Signs for Dynamic Correction with Two Weights.

Phase Quadrant			Sign			
α	β	γ	W	x	y	z
1	1	1	+	+	+	+
2	4	3	+	+	+	-
3	2	4	+	+	-	-
4	3	2	+	+	-	+

Although there is a possibility of these positive W cases occurring, a real spacecraft configuration would very often need the equivalent of a negative W correction, with $+W$ at x , y , z and $+W'$ at x' , y' , z' (Table 3).

Table 3 indicates the correction required at the surface of a sphere of unit radius, and the preferable choice between A , B , and C would be that for which W' is a minimum; i.e., A if y and z are both less than x , B if x and z are both less than y , and C if x and y are less than z .

The tables give the coordinates of correction weights in the positive X hemisphere; it is to be understood that equal but diametrically opposed weights are also implied in the negative X hemisphere (i.e., reverse signs of x , y , z , x' , y' , and z').

$\Delta_{\alpha\beta\gamma}^{xyz}$ is the computation of W, x, y, and z, and z, and W', x', y', z' in Table 3, or of W, x, y, z in Table 4, whichever is applicable to the phase quadrants of α , β , and γ . $\Delta_{\alpha\beta\gamma}^{xyz}$ will yield a tri-axial dynamic unbalance correction; however alternative corrections may be preferable because they involve less added weight, because weight addition may be more convenient, or because the accuracy of correction may be less dependent on precise placement of weights. Advantages of minimum weight and minimum dependence on weight placement accuracy will, in general, be

1. $x = +0.578, y = +0.578, z = +0.578$ ($\alpha = 45^\circ, \beta = 45^\circ, \gamma = 45^\circ$),
2. $x = +0.578, y = +0.578, z = -0.578$ ($\alpha = 135^\circ, \beta = 315^\circ, \gamma = 225^\circ$),
3. $x = +0.578, y = -0.578, z = -0.578$ ($\alpha = 225^\circ, \beta = 135^\circ, \gamma = 315^\circ$),
4. $x = +0.578, y = -0.578, z = +0.578$ ($\alpha = 315^\circ, \beta = 225^\circ, \gamma = 135^\circ$).

A weight W at one of these locations (plus an equal weight diametrically opposed on the "negative X" hemisphere) will cause $D_x = D_y = D_z = 0.942W = D_P$ with phase angles α, β, γ as indicated (0.578 is an approximation of $1/\sqrt{3}$ and 0.942 of $(2/3)\sqrt{2}$).

Figure 5 is a common plot of normalized initial unbalance vectors $D_x/\alpha(OX), D_y/\beta(OY),$ and $D_z/\gamma(OZ)$. $D_P(OP_{123})$ represents a triaxially symmetric vector, per case 1 above, $D_x'/\alpha'(P_{123}X), D_y'/\beta'(P_{123}Y),$ and $D_z'/\gamma'(P_{123}Z)$ represent the correction needed after applying D_P . Although

concurrent and associated with maximum angular displacement of correction weights from all the reference axes.

Alternative corrections are obtained by adding a triaxially symmetric unbalance vector and correcting for the modified unbalance situation.

A triaxially symmetric vector means an opposed pair of weights located so that they provide equal unbalance effect about all of the reference axes. There are only four possible locations with positive X coordinates on the "unit sphere" surface:

D_P must be at 45 degrees, its magnitude is arbitrary and therefore an infinite number of possible solutions exists. In all cases $D_x'/\alpha', D_y'/\beta', D_z'/\gamma'$ will be normalized, i.e., consistent, and the total ballast weight required will be a function of the magnitude of D_P : total weight = $(3/\sqrt{2})D_P$ plus the weight indicated by $\Delta_{\alpha'\beta'\gamma'}^{x'y'z'}$.

The total weight can be expressed in terms of D_P and known data, and solving for zero values of the derivative of this expression would yield the optimum value of D_P for minimum weight correction. However, this would be a prohibitively cumbersome operation since the total weight is

$$W_{TOTAL} = \frac{3}{\sqrt{2}} D_P + 2 \left(K_1 + K_2 + \frac{1}{2} \right) \sqrt{\frac{D_P^2 + D_z^2 - \sqrt{2} D_P D_z (\cos \gamma + \sin \gamma)}{K_z (K_x + K_y)}} \quad (52)$$

where

$$K_x = \frac{1}{1 + \left(\frac{-x^2 + y^2 + z^2}{+x^2 - y^2 + z^2} \right) + \left(\frac{-x^2 + y^2 + z^2}{+x^2 + y^2 - z^2} \right)}$$

$$K_y = \frac{1}{1 + \left(\frac{+x^2 - y^2 + z^2}{-x^2 + y^2 + z^2} \right) + \left(\frac{+x^2 - y^2 + z^2}{+x^2 + y^2 - z^2} \right)}$$

$$K_z = \frac{1}{1 + \left(\frac{+x^2 + y^2 - z^2}{-x^2 + y^2 + z^2} \right) + \left(\frac{+x^2 + y^2 - z^2}{+x^2 - y^2 + z^2} \right)}$$

(note that $K_x^2 + K_y^2 + K_z^2 = 1$). K_1 and K_2 are the smaller two of $K_x, K_y,$ and K_z if the phase quadrant combination of α', β', γ' is 142, 214, 333, or 421, but are both zero if the phase quadrant combination of α', β', γ' is 111, 243, 324, or 432;

$$x^2 = D_P^2 + D_x^2 - \sqrt{2} D_P D_x (\cos \alpha + \sin \alpha)$$

$$y^2 = D_P^2 + D_y^2 - \sqrt{2} D_P D_y (\cos \beta + \sin \beta)$$

$$z^2 = D_P^2 + D_z^2 - \sqrt{2} D_P D_z (\cos \gamma + \sin \gamma)$$

$$\alpha' = \tan^{-1} \frac{D_x \sin \alpha - 0.707D_p}{D_x \cos \alpha - 0.707D_p}$$

$$\beta' = \tan^{-1} \frac{D_y \sin \beta - 0.707D_p}{D_y \cos \beta - 0.707D_p}$$

$$\gamma' = \tan^{-1} \frac{D_z \sin \gamma - 0.707D_p}{D_z \cos \gamma - 0.707D_p}$$

Equation 52 holds only for case 1 of the triaxially symmetric vector D_p . It would have to be modified by geometrically appropriate changes of the signs of some of the trigonometric functions to be valid for cases 2, 3, or 4.

The total weight function was evaluated for specific values of D_x/α , D_y/β , and D_z/γ and various case 1 values of D_p . Setting $D_p=0$ gave the minimum number of correction weights, but not the minimum total weight. The total weight function had more than one minimum and at least one relative maximum. It became infinite when α' , β' , or γ' was 0° , 90° , 180° , or 270° . Further evaluation of the function could be made by expressing its derivative analytically, a formidable task, or by curve plotting based on iterative computation for incremental values of D_p . The

latter would be tedious but relatively straightforward. If D_x were larger than D_y and D_z , D_p could be limited to the value of D_x and increased in increments of $0.01 D_x$ from zero to D_x .

Figure 5 illustrates only case 1 of the four possible triaxially symmetric unbalance vectors. Each of the other three cases is represented on a common vector plot by equal vectors at 135° , 225° , and 315° , as shown in Figure 6, where OP_1 is the contribution to D_x , OP_2 the contribution to D_y , and OP_3 the contribution to D_z .

Figure 7 shows initial unbalances OX , OY , OZ with a case 2 triaxially symmetric vector represented by OP_1 , OP_2 , and OP_3 , and a case 3 symmetric vector represented by OP_1' , OP_2' , and OP_3' . Note that P_1 and P_1' are joined to X , P_2 and P_2' to Y , and P_3 and P_3' to Z .

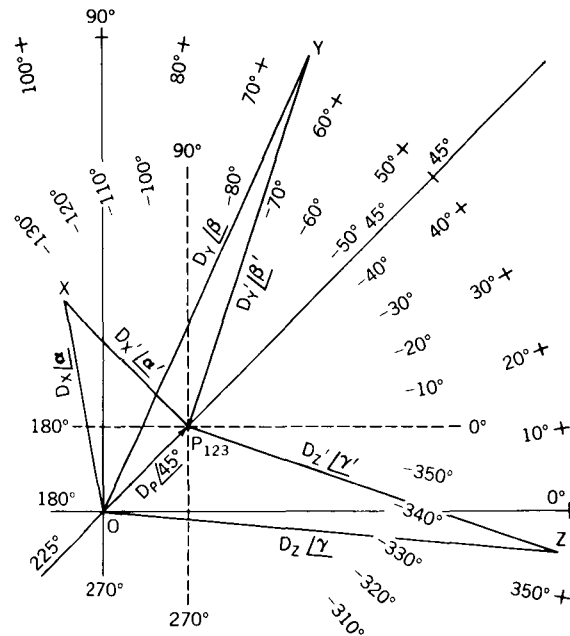


FIGURE 5.—Vector plot for case 1 of a triaxially symmetric vector D_p .

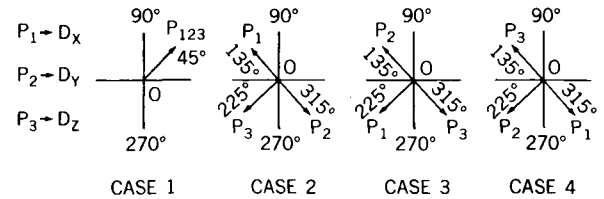


FIGURE 6.—Common vector plots of the four possible triaxially symmetric vector cases.

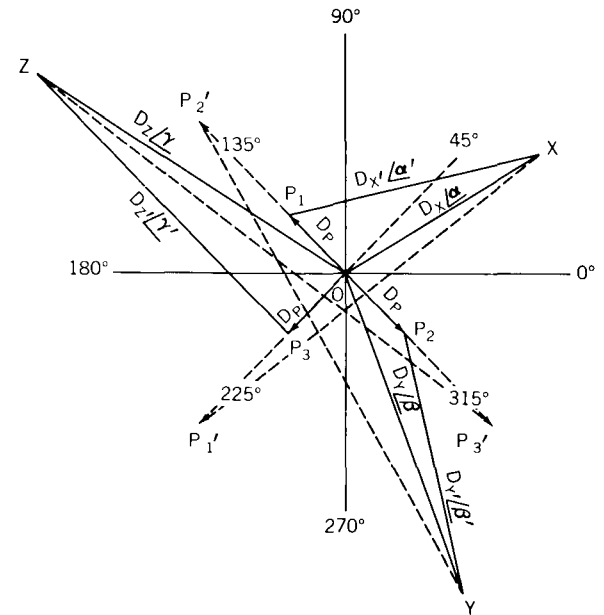


FIGURE 7.—Vector plots for case 2 (→) and case 3 (---) of triaxially symmetric vectors.

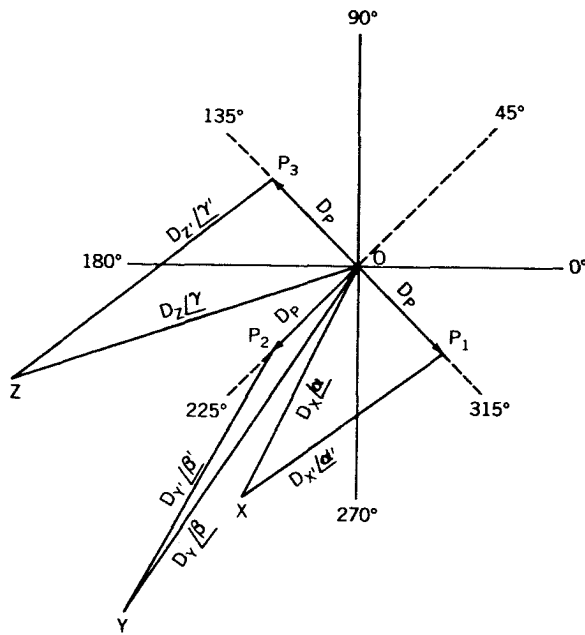


FIGURE 8.—Vector plot for case 4 of a triaxially symmetric vector D_P .

Figure 8 shows a similar plot for a case 4 symmetric vector. For all cases, the magnitude of the symmetric vector is arbitrary (OP_1 , OP_2 , and OP_3 must be equal but can have any length) and therefore each case represents a range of possible solutions. Also D_x/a' , D_y/b' , and D_z/g' will be normalized, and the total ballast weight needed will be $(3/\sqrt{2}) D_P$ plus the weight indicated by $\Delta_{a'\beta'\gamma'}^{xyz}$.

To consider all possibilities it would be necessary to evaluate all the cases for all values of D_P . However, some cases can be eliminated by inspection. If a , β , γ are all in the first quadrant, only case 1 need be considered, and if a , β , γ are all in the third quadrant (see Figure 8), only cases 2, 3, and 4 need be considered. Apart from these two general rules, a vector plot may allow some elimination. For example, case 3 could be eliminated in the situation shown by Figure 7, because a case 3 triaxially symmetric vector opposes all three of the original unbalance vectors. Of course, computer programming could be used to evaluate all possibilities and this might be better than including selective rejection rules in the program.

In summary, computation of the optimum dynamic unbalance correction consists of assuming

ranges of values for each of the four symmetric vectors, evaluating the weight needed for each symmetric vector plus the weight needed for correction of the remaining unbalance, and selecting the result which yields the minimum total weight to be added.

At this point it is pertinent to consider the advantages of finding an optimum (minimum weight) solution rather than, for instance, that given by $\Delta_{a\beta\gamma}^{xyz}$ with no symmetric vector.

The obvious advantage is saving weight, and the weight saved may be considerable. For a specific case (Appendix B) $\Delta_{a\beta\gamma}^{xyz}$ yielded 1886 grams, whereas an optimized solution necessitated only 1192 grams, and this was for an 83 kilogram spacecraft where the initial unbalance was comparatively small. Another advantage is that an optimized correction will, in practice, balance more precisely (assuming a "one shot" correction). The reason is that the D_P vector weight is located for minimum sensitivity to placement error and also induces the secondary corrections into locations of reduced sensitivity to placement error, i.e., it pulls them away from the reference axes. A given radial placement error has proportionally less effect at increased radius. Also, the probability of accurate placement is better for smaller weights. Advance provision for the eight possible D_P vector weights could well be advantageous.

Further justification for optimizing might be that it could easily be done by computer, and the incidental computation of many alternative corrections would, for an actual spacecraft, greatly assist in selecting an expedient correction.

For "single axis" balancing, weight optimizing, by judicious selection of correction planes, can be intuitive and alternative corrections can readily be devised if interference problems arise. But intuition is inadequate for the greater complexity of triaxial balancing, and manual computation of alternative corrections would be intolerably time consuming.

Correcting Unbalance

Unbalance can be corrected either by design or by ballasting. Balancing by design means placing or relocating components of an assembly so that the unbalance is within specified limits. It might also include such steps as changing the primary reference axes, or repositioning control

elements, to conform to the actual measured or computed situation. Most of this study is quite pertinent to the general problem of designing spacecraft to be adequately well balanced in the first place so that no correction is needed. Balance is always recognized as a design factor, and development of more comprehensive and sophisticated techniques for balancing by design might well meet the needs of large spacecraft with relatively liberal unbalance tolerances. Key factors would be comprehensive and systematic mass measurement of all components, good quality control of assembly, effective error analysis, and suitable data processing and computer programs.

Correction by ballasting means adding weight in locations defined by the expression of unbalance. These locations are much more restrictive for triaxial balancing than for single axis balancing. Each of the weights needed must lie on a specific radiant line from the reference axis

origin. "Splitting" weights to give the same vector effect is not impossible, but would need complicated analysis, especially for dynamic unbalance corrections.

A static correction equivalent to weight N at unit radius requires a weight of N/r at radius r .

The diametrically opposed radiants for dynamic unbalance correction are defined by $+x, +y, +z$ and $-z, -y, -x$. If at unit radius each correction weight would be W , and both weights are to be at the same radius P , each would have to be W/P^2 . If one weight L_1 is to be at radius r_1 , and the opposed weight L_2 at radius r_2 then:

$$L_1 r_1 = L_2 r_2 \quad (53)$$

Static balance must be preserved. Dynamic balance about axis XX must also be preserved. Let the included angle between the XX axis and the dynamic correction radiant be ϕ ; then

$$L_1 r_1^2 \cos \phi \sin \phi + L_2 r_2^2 \cos \phi \sin \phi = 2W(1)^2 \cos \phi \sin \phi \quad (54)$$

As $\cos \phi \sin \phi$ cancels out, other axes need not be considered. From Equations 53 and 54 it follows that

$$r_2 = \frac{2W}{L_1 r_1} - r_1 \quad (55)$$

$$L_1 = \frac{2W}{r_1(r_1 + r_2)} \quad (56)$$

If values are assumed for W , r_1 , and L_1 , Equations 53 and 55 will give r_2 and L_2 . Note from Equation 55 that $L_1 r_1^2$ must be less than $2W$; otherwise L_2 would have to be infinite for static balance to be maintained, since for $L_1 r_1^2 = 2W$, r_2 becomes zero. If values are given to W , r_1 and r_2 Equations 53 and 56 will give L_1 and L_2 . In practice all radii should be given maximum values consistent with the spacecraft configuration, to minimize the ballast, and any effect of static balance correction on dynamic balance must be considered.

CONCLUSIONS

This study reports the derivation of analytical relationships, coordinate system coding of unbalance corrections, and various related concepts.

Its intention was to promote understanding and suggest effective practical applications. Defining the $N, d, e, W, x, y,$ or $-C, p, q, r$ unbalance expression system and operation \oplus were major objectives. Its practical motivation was the needs of a particular spacecraft program (Appendix B).

Triaxial balancing is inherently more complicated than single axis balancing. However, balance tolerances are likely to be less stringent. Five problem areas are apparent:

1. How to make sufficient physical measurements to define initial unbalance. This study has defined what measurements are necessary, unless prior knowledge or acceptable estimates exist. These may be total weight, inertias about three axes, magnitude and phase of static and dynamic unbalance about one axis, and phase only of static and dynamic unbalance about a second axis. However, the effect of static unbalance correction on dynamic unbalances must be considered. This can be done by coding unbalances about $x=y=z=0$ as $(N, d, e, W, x, y)_A$ and $(N, d, e, -N, d, e)_B$ (due to static unbalance) and performing $(N, d, e, W, x, y)_A \oplus (N, d, e, -N, d, e)_B$, or by applying static unbalance correction *before* determining

dynamic unbalances. The relative merits and the best sequence of operations for either case depend on specific circumstances. There are other combinations of sufficient measurements, which suggests study of which measurements are easier to obtain with requisite accuracy. Redundant data will, if properly processed, improve the overall accuracy of balancing.

2. Relating measurements to unbalance tolerances. This study covers how it may be done.

3. Determining required corrections. This study proposes a system for expressing triaxial unbalance which implicitly defines corrections, and discusses latitude in actual weight location. The final choice depends on the specific spacecraft configuration.

4. Applying corrections. Because choice of location is very limited, this is likely to be more of a problem than in single axis balancing. Also, vector splitting techniques are difficult to evaluate, and should be considered with extreme caution, although the technique for computing the effect of changes to spacecraft can be used to check whether two proposed weights are in fact equivalent to an indicated single weight. Computer programming would be almost essential to apply this technique very extensively.

5. Evaluating changes to spacecraft, including addition of appendages, etc., which cannot be balanced on the spacecraft. This study covers how this may be done. Again, computer programming is extremely advantageous.

APPENDIX A

SOLUTION OF THE TRIAXIAL BALANCE EQUATIONS

If it is given that

$$D_x = a = 2Wx \sqrt{y^2 + z^2} \quad , \quad (A1)$$

$$D_y = b = 2Wy \sqrt{x^2 + z^2} \quad , \quad (A2)$$

$$D_z = c = 2Wz \sqrt{x^2 + y^2} \quad , \quad (A3)$$

$$x^2 + y^2 + z^2 = 1 \quad , \quad (A4)$$

it is obvious that

$$\frac{a}{2W} = x \sqrt{y^2 + z^2} \quad , \quad (A5)$$

$$\frac{b}{2W} = y \sqrt{x^2 + z^2} \quad , \quad (A6)$$

$$\frac{c}{2W} = z \sqrt{x^2 + y^2} \quad . \quad (A7)$$

Let

$$\left(\frac{a}{2W}\right)^2 = x^2(y^2 + z^2) = (xy)^2 + (xz)^2 = L \quad , \quad (A8)$$

$$\left(\frac{b}{2W}\right)^2 = y^2(x^2 + z^2) = (xy)^2 + (yz)^2 = M \quad , \quad (A9)$$

$$\left(\frac{c}{2W}\right)^2 = z^2(x^2 + y^2) = (xz)^2 + (yz)^2 = N \quad . \quad (A10)$$

Also set

$$(xy)^2 = P \quad ,$$

$$(xz)^2 = Q \quad ,$$

$$(yz)^2 = R \quad .$$

Therefore

$$P + Q = L \quad , \quad (A11)$$

$$P + R = M \quad , \quad (A12)$$

$$Q + R = N \quad , \quad (A13)$$

and so

$$L + M = 2P + Q + R \quad , \quad (A14)$$

From Equations A1–A3 it can be seen that

$$P = L - Q \quad , \quad (A15)$$

$$R = M - P \quad , \quad (A16)$$

$$Q = N - R \quad . \quad (A17)$$

Then from Equations A14, A16, and A17

$$\begin{aligned} 2P + Q + R &= 2P + N - R + M - P \\ &= 2P + N - M + P + M - P \\ &= 2P + N \quad . \end{aligned}$$

So $2P + N = L + M$; therefore

$$P = \frac{L + M - N}{2} \quad . \quad (A18)$$

By reverting to original symbols:

$$\begin{aligned} (xy)^2 &= \frac{\left(\frac{a}{2W}\right)^2 + \left(\frac{b}{2W}\right)^2 - \left(\frac{c}{2W}\right)^2}{2} \\ &= \frac{a^2 + b^2 - c^2}{2(2W)^2} \\ &= \frac{a^2 + b^2 - c^2}{8W^2} \quad . \quad (A19) \end{aligned}$$

From Equations A11 and A13

$$\begin{aligned} L + N &= 2Q + P + R \\ &= 2Q + L - Q + M - P \\ &= 2Q + L - Q + M - L + Q \\ &= 2Q + M \end{aligned}$$

so

$$Q = \frac{L + N - M}{2} \quad (\text{A20})$$

Again by reverting to original symbols:

$$\begin{aligned} (xz)^2 &= \frac{\left(\frac{a}{2W}\right)^2 + \left(\frac{c}{2W}\right)^2 - \left(\frac{b}{2W}\right)^2}{2} \\ &= \frac{a^2 + c^2 - b^2}{2(2W)^2} \\ &= \frac{a^2 - b^2 + c^2}{8W^2} \end{aligned} \quad (\text{A21})$$

Now, from Equations A12 and A13

$$\begin{aligned} M + N &= 2R + P + Q \\ &= 2R + L - Q + N - R \\ &= 2R + L - N + R + N - R \\ &= 2R + L \end{aligned}$$

so

$$R = \frac{M + N - L}{2} \quad (\text{A22})$$

Again, in reverting to original symbols:

$$\begin{aligned} (yz)^2 &= \frac{\left(\frac{b}{2W}\right)^2 + \left(\frac{c}{2W}\right)^2 - \left(\frac{a}{2W}\right)^2}{2} \\ &= \frac{-a^2 + b^2 + c^2}{8W^2} \end{aligned} \quad (\text{A23})$$

To obtain x^2 , y^2 , and z^2 , begin by dividing Equation A19 by Equation A23,

$$\begin{aligned} \frac{(xy)^2}{(yz)^2} &= \frac{\frac{a^2 + b^2 - c^2}{8W^2}}{\frac{-a^2 + b^2 + c^2}{8W^2}} \\ \frac{x^2}{z^2} &= \frac{a^2 + b^2 - c^2}{-a^2 + b^2 + c^2} \\ x^2 &= z^2 \left(\frac{a^2 + b^2 - c^2}{-a^2 + b^2 + c^2} \right) \end{aligned} \quad (\text{A24})$$

The value for y^2 can be obtained from Equations A21 and A23,

$$\begin{aligned} \frac{(xy)^2}{(xz)^2} &= \frac{a^2 + b^2 - c^2}{a^2 - b^2 + c^2} \\ y^2 &= z^2 \left(\frac{a^2 + b^2 - c^2}{a^2 - b^2 + c^2} \right) \end{aligned} \quad (\text{A25})$$

Likewise, from Equations A19 and A23

$$\begin{aligned} \frac{(yz)^2}{(xy)^2} &= \frac{-a^2 + b^2 + c^2}{a^2 + b^2 - c^2} \\ z^2 &= x^2 \left(\frac{-a^2 + b^2 + c^2}{a^2 + b^2 - c^2} \right) \end{aligned} \quad (\text{A26})$$

The procedure for deriving x^2 , y^2 , and z^2 can be used to derive the following:

$$\begin{aligned} \frac{x^2}{y^2} &= \frac{(xz)^2}{(yz)^2} = \frac{a^2 - b^2 + c^2}{-a^2 + b^2 + c^2} \\ x^2 &= y^2 \left(\frac{a^2 - b^2 + c^2}{-a^2 + b^2 + c^2} \right) = z^2 \left(\frac{a^2 + b^2 - c^2}{-a^2 + b^2 + c^2} \right) \end{aligned} \quad (\text{A27})$$

$$\frac{y^2}{x^2} = \frac{(yz)^2}{(xz)^2} = \frac{-a^2 + b^2 + c^2}{a^2 - b^2 + c^2} ,$$

$$y^2 = x^2 \left(\frac{-a^2 + b^2 + c^2}{a^2 - b^2 + c^2} \right) = z^2 \left(\frac{a^2 + b^2 - c^2}{a^2 - b^2 + c^2} \right) ; \quad (\text{A28})$$

$$\frac{z^2}{y^2} = \frac{(xz)^2}{(xy)^2} = \frac{a^2 - b^2 + c^2}{a^2 + b^2 - c^2} ,$$

$$z^2 = y^2 \left(\frac{a^2 - b^2 + c^2}{a^2 + b^2 - c^2} \right) = x^2 \left(\frac{-a^2 + b^2 + c^2}{a^2 + b^2 - c^2} \right) . \quad (\text{A29})$$

It can be seen that Equations A27 and A24 are the same, Equations A28 and A25 are the same, and Equations A29 and A26 are the same. By using Equations A4, A28, and A25,

$$x^2 + x^2 \left(\frac{-a^2 + b^2 + c^2}{a^2 - b^2 + c^2} \right) + x^2 \left(\frac{-a^2 + b^2 + c^2}{a^2 + b^2 - c^2} \right) = 1 ,$$

$$x = \frac{1}{\sqrt{1 + \left(\frac{-a^2 + b^2 + c^2}{a^2 - b^2 + c^2} \right) + \left(\frac{-a^2 + b^2 + c^2}{a^2 + b^2 - c^2} \right)}} . \quad (\text{A30})$$

Similarly

$$y = \frac{1}{\sqrt{1 + \left(\frac{a^2 - b^2 + c^2}{a^2 + b^2 - c^2} \right) + \left(\frac{a^2 - b^2 + c^2}{-a^2 + b^2 + c^2} \right)}} . \quad (\text{A31})$$

$$z = \frac{1}{\sqrt{1 + \left(\frac{a^2 + b^2 - c^2}{-a^2 + b^2 + c^2} \right) + \left(\frac{a^2 + b^2 - c^2}{a^2 - b^2 + c^2} \right)}} . \quad (\text{A32})$$

Inspection of Equations A30, A31, and A32 will show that their solution will result in real numbers only if the sum of any two of the quantities a^2 , b^2 , and c^2 exceeds the third quantity.

APPENDIX B

TRIAxIAL BALANCING OF SAN MARCO FLIGHT UNIT 2

The San Marco project is a joint program of Italy and the United States and the balancing of the Italian built spacecraft, performed by NASA at the Goddard Space Flight Center, was aided by the cooperation of personnel of the Italian Space Commission. Triaxial balancing was necessary for proper functioning of the drag measurement equipment on the spacecraft. San Marco flight unit 2 was balanced about its three reference axes by using stub shaft attachments to the spacecraft instead of the "wrap around" fixture used for earlier San Marco balance operations. The unit was balanced twice and both operations will be discussed here. The first operation reduced the dynamic unbalance to less than 25 oz.-in.² about all three axes; the second to less than 15 oz.-in.² Both operations reduced center of gravity displacement to less than 0.005 in.—the best value obtainable because of nonconcurrence of the reference axes.

First Balancing Operation

The stub shafts and spacecraft attachment adapters were individually statically balanced before assembly. Their static unbalance was minor and their design was assumed to preclude any appreciable dynamic unbalance. Alignment checks were made to ensure that the axes of rotation essentially coincided with the spacecraft reference axes. Deviations noted were considerably less than for the "wrap around" fixture used for flight unit 1 and did not exceed the practical limitations imposed by machining tolerances.

Measurements of initial static and dynamic unbalance about each reference axis were used to compute required corrections. Experience with unit 1 had shown that minor inconsistencies in measured data could affect the correction computation drastically, and even render the basic equations incapable of rational solution (as mentioned in the text). Therefore, a technique was developed to "normalize" (i.e., make consistent

by averaging out the incompatibilities between redundant data) the measured data prior to calculation. The alternative solution of the basic set of dynamic balance equations was also derived and used and proved more amenable to arithmetic operations.

Corrections were applied to the spacecraft and residual unbalance was measured about all three axes. Results were: Spacecraft center of gravity displacement from reference axes intersection was reduced from approximately 0.2 to 0.003 in.; dynamic unbalance, axis XX, was reduced from 520 oz.-in.² to 25 oz.-in.²; dynamic unbalance, axis YY, was reduced from 1210 oz.-in.² to 22 oz.-in.²; dynamic unbalance, axis ZZ, was reduced from 1100 oz.-in.² to 20 oz.-in.².

The weight of the spacecraft (including shell) after balancing was 177.7 ± 0.1 lb. Approximately 8.5 lb (4.8 percent of the total) was added for balancing.

Second Balancing Operation

Between the first and second balancing operations the spacecraft had minor component modifications, but the initial unbalance was assumed to be small. Since the earlier dynamic unbalance correction was not optimized for minimum weight, it was decided to remove it, apply the equivalent optimized correction determined as described in the text, and then measure and correct the remaining unbalance. Substitution of the optimized for the original dynamic correction gave a net weight reduction of 0.95 lb, and unbalance was then measured as follows:

1. Static balance was 30 oz. in. about axis XX, 21 oz. in. about YY, and 27.5 oz. in. about ZZ.
2. Dynamic unbalance was 160 oz. in.² about axis XX, 124 oz. in.² about YY, and 139 oz. in.² about ZZ.

The apparent static unbalance was largely due to nonconcurrence of the actual axis of spin. For the spin axis it was noted that the spacecraft in-

terface (at the separation plane) was a very loose fit on the adapter, having at least a 0.020 in. lateral slop, and was actually located (non-repeatably) by the external marmon clamp.

The dynamic unbalance was consistent with the expected results of modifications to the spacecraft; therefore the equivalence of the optimized to the original correction was essentially demonstrated.

The remaining dynamic unbalance did not warrant computing an optimized correction or consolidating the correction with existing weights; therefore a basic correction (no triaxially symmetric vector) was applied; this added 0.33 lb. Some final static balancing involved removal of 0.31 lb from existing balance weights, the overall operation reduced the total balancing weight on the spacecraft by 0.93 lb. Balancing required weight additions at 13 discrete locations, and because of spacecraft structural considerations 25 individual weights were actually attached. The final balance operation took 3 days and about 75 man-hours effort, much less than the earlier operation, because techniques had been improved and the spacecraft had less initial unbalance.

The residual unbalance obtained was as follows: Static unbalance was 8.6 oz. in. about spin axis XX, 9.7 oz. in. about YY, and 8.5 oz. in. about ZZ. (At least 70 percent of this static unbalance was the unavoidable consequence of nonconcurrency of the reference axes.) Dynamic unbalance was 10 oz. in.² about XX, 12 oz. in.² about YY, and 15 oz. in.² about ZZ. The final weight of the

spacecraft, including shell and antennas, was 177.5 lb. About 4.25 percent of this was balance weight (considerably less than the 4.8 percent needed by the first balancing operation).

Conclusions

Entirely new techniques were developed and applied to achieve the desired balance of the San Marco. This was the most complex precision balancing operation yet performed. The minor additional complications of concurrent control of inertias and total weight would have posed the ultimate complexity of complete control of all mass properties, so almost any balancing problem would be soluble by the general methods developed for San Marco.

The first balancing operation corrected unbalance with the minimum *number* of correction weights, and the second with a larger number of weights but with less total added weight.

The residual static unbalance was negligible because of the greater effect of nonconcurrency of reference axes. The residual dynamic unbalance, being equivalent to principal axis angular deviations between 0.15 and 0.02 degree, was low enough that non-orthogonality of reference axes may well have been significant.

Therefore, the San Marco flight unit 2 was balanced as well as the geometric errors of its own structure permitted. Effects due to unbalance could not have exceeded those due to geometric inaccuracies.

SPACE VACUUM INVESTIGATIONS

HENRY MAURER, JR.

Inaccessible-Non Repairable-Costly: The meaning of these words is absolute in the event of a complete failure of an orbiting spacecraft. Adequate design, or rather, minimizing chance of failure, challenges all scientists and engineers. Where experience has not been developed, the risk involved in a new application must be weighed against the consequence of a design failure. In space work, the spectrum of consequence is extremely wide. It ranges from loss of scientific data on one hand, to loss of life on the other. There is little question that knowledge of the space environment and its effects must be determined to produce reliable spacecraft. Tests to prove the design must be performed with equipment capable of simulating the effects of the environment in which the design is to operate. This does not necessarily mean that the natural environment must be duplicated. The degree of environmental simulation that should be achieved depends primarily upon the significance of the physical effects it causes. The purpose of space vacuum simulation in an environmental test program should be accurately conceived and the testing phased to obtain maximum information from the limited samples available for test.

This discussion is confined to some of the factors involved in simulating the effects of space vacuum environment. This survey is by no means complete, but will touch on some of the highlights in this field:

1. A description of the environment
2. The level of simulation necessary to evaluate its effects
3. Test equipment capability
4. Experience in test
5. Experience in orbit
6. Observations & future emphasis

THE ENVIRONMENT

The space vacuum environment covers the range from 10^{-3} mm Hg to 10^{-16} mm Hg or lower for interplanetary space. At lower altitudes, it's composition is a many varied thing and is affected by altitude, latitude, time of day, season of the year and disposition of the sun—whether at sun spot maximum or sun spot minimum. The earth's atmosphere extends to great altitudes. This fact is confirmed by various upper atmospheric and ionospheric observations, some of which have been made by the University of Michigan's High Altitude Engineering Laboratory. The atmosphere is essentially a compressible fluid in hydrostatic equilibrium, it's density varies roughly exponentially with height above the earth's surface, and no precise upper boundary has yet been defined. Our knowledge of the pressure, density, kinetic temperature, and composition is rather scant for altitudes above 90 km. The work to date by the Committee on Extending the Standard Atmosphere, COESA, has resulted in a proposed standard up to 32 km, a tentative standard to 90 km, and a speculative standard from 90 km to 700 km.

Suffice to say, then, that from 90 km to higher altitudes, models of the atmosphere are based primarily on some sounding rocket flights, drag data from satellites, and kinetic theory. The direct measurements made by Explorer XVII over the altitude range 256 km to 866 km were in general agreement with kinetic theory. However, mass spectrometer data has not as yet been published.

Figure 1 shows an approximation of the pressure distribution $\left(P = nKT \text{ or } \frac{nMV^2}{3}\right)$. This model is based on several sources and particularly data appearing in the 1962 issue of Astronautical Engineering Handbook. It represents the probable range of the environment. From the figure,

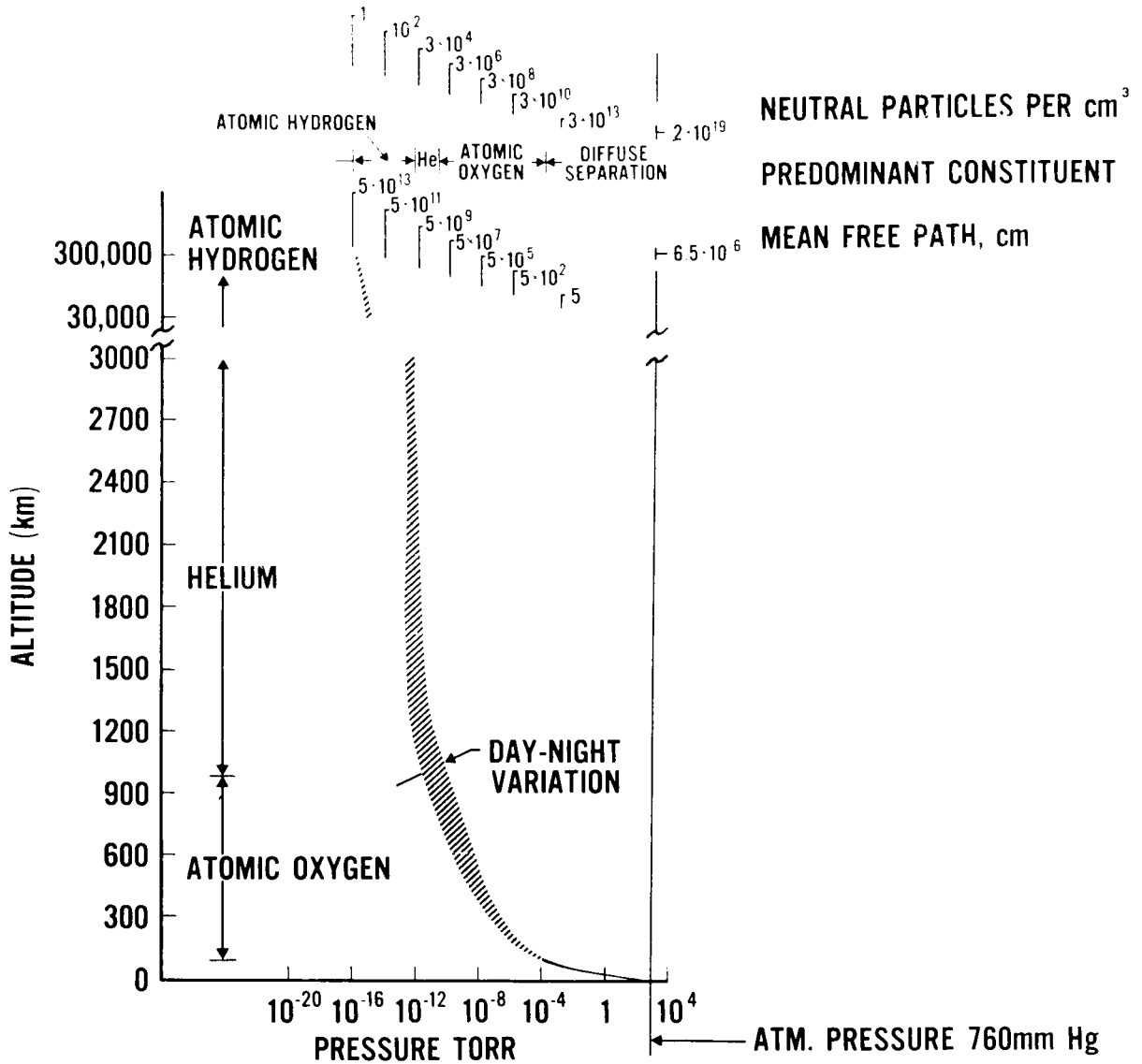


FIGURE 1.

diffuse separation is shown to start at ≈ 100 Km, predominant constituents are:

- Atomic oxygen from 100–1,000 km
- Helium from 1,000–3,000 km
- Atomic hydrogen 3,000–to higher altitudes

Estimated number densities of particles ($n = P/M$) and mean free path between collisions

$$\left(M.F.P. = \frac{5 \times 10^{-3}}{P \text{ mmHg}} \right)$$

is shown at the top of the figure.

An approximation of the vertical temperature and pressure distribution through the atmosphere for altitudes up to 700 km is shown in Figure 2.

It is significant to point out that the range of typical spacecraft missions covers the entire range from 10^{-4} to 10^{-16} torr. Of 260 satellites orbited by the U. S. and the U. S. S. R., approximately 75 percent had apogees exceeding 700 km. Therefore, our interest covers the entire range of 10^{-4} to 10^{-16} torr and it's possible effects on spacecraft.

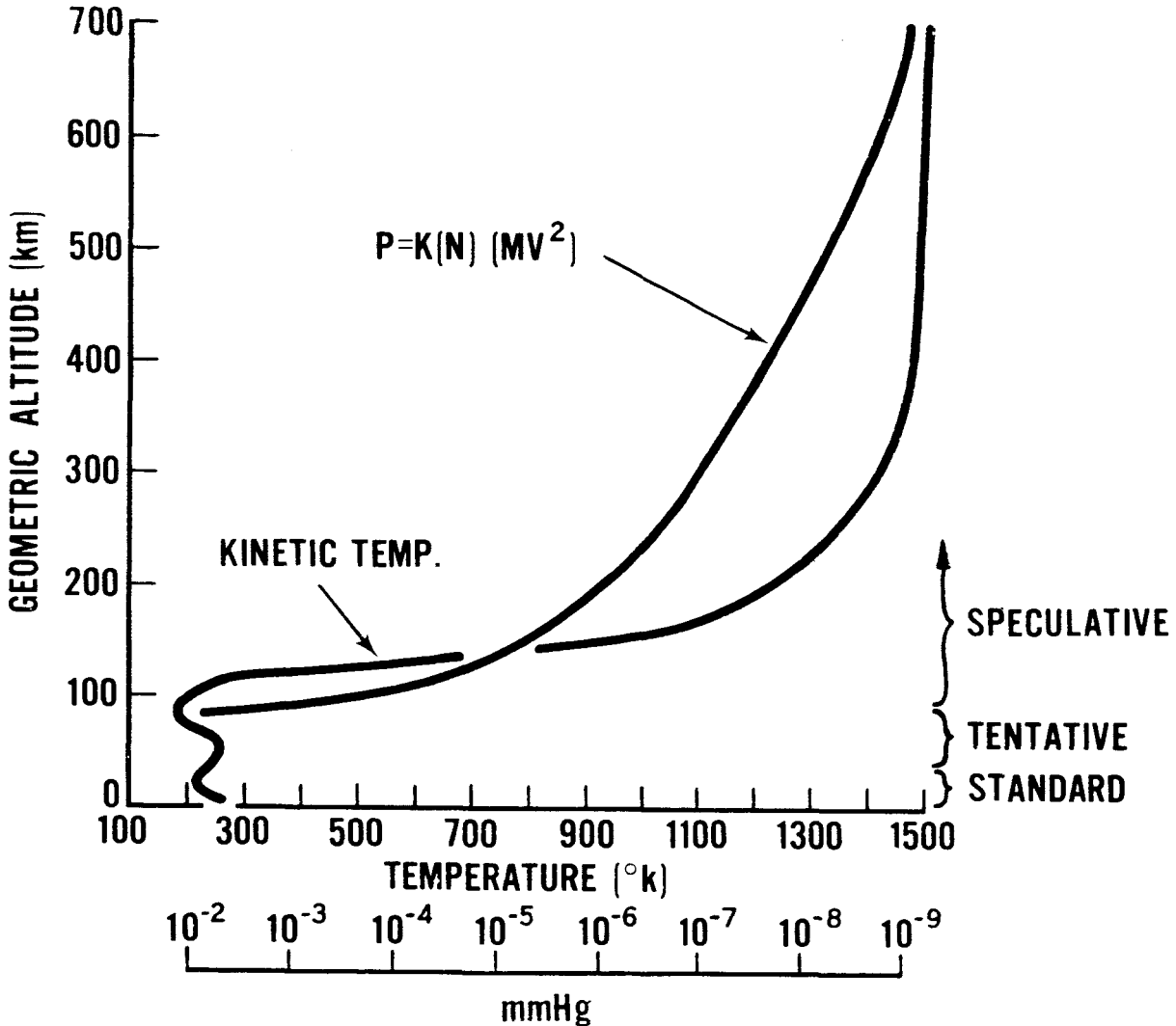


FIGURE 2.—Vertical temperature and pressure distribution through the atmosphere.

WHAT ARE THE EFFECTS OF THE VACUUM ENVIRONMENT?

One of the most complete surveys on this subject was conducted by Jaffee & Rittenhouse of the JPL (their original publication appeared in the 1962 issue of the AIAA Journal and is currently being updated). In general, the survey indicated that materials undergo weight loss due to evaporation, sublimation and decomposition. For some materials, decomposition can cause several orders of magnitude greater loss than that due to sublimation alone. Metals and alloys are quite stable. Sublimation of volatile constituents is of concern particularly where there is the possibility

of the material condensing on colder surfaces. Frictional behavior of contacting metal surfaces in a space environment is a problem with uncertainties resulting from the lack of reliable data on friction in ultra high vacuum. Among semiconductor materials, selenium, phosphides, and arsenides have high sublimation rates in high vacuum at moderate temperatures. Most organic engineering insulators are unaffected by the vacuum of space except at high temperatures. Because of the complex composition of most oil and grease lubricants, computed estimates of their evaporation rates are not reliable. Polymers show high decomposition rates in vacuum.

A number of investigators have generalized on the subject of the degree of vacuum simulation required to reveal a particular effect. Little has occurred that would cause one to change the criteria which had been stated earlier.

Table 1 summarizes the criteria. (Santler, Simons, et al).

CORONA

In order to evaluate corona or glow discharge effects, pressure simulation over the range of atmosphere to 10^{-4} is desired. Corona occurs when the ionization potential of a gas or air is exceeded. The glow process is the visual release of energy through deionization. In an air or oxygen environment, corona causes the generation of small quantities of ozone which attacks common insulation materials. In high voltage applications, the pressure, electrode separation distance, gas composition and dielectric strength of the material separating the electrodes determines the breakdown voltage. When deionization extends from one electrode to the other, the integrity of the dielectric is lost and sparkover can occur. Even where sparkover does not occur, energy loss due to corona can completely carbonize some dielectrics making them useless. The corona release is described by Patchens Law—the breakdown voltage is a function of Pressure (P) \times conductor Separation distance (D) and decreases with this product ($P\times D$) to a certain minimum, after which it rises again as $P\times D$ decreases further.

Heat Transfer

To minimize gas conduction and convection between mating parts and their surroundings, pressure in the range of 10^{-4} to 10^{-5} mm Hg. is

required. At this level of pressure, conduction only occurs at solid surface sites. Thus, radiative transfer predominates. This level of vacuum is necessary for heat balance investigations.

SUBLIMATION AND EVAPORATION

Sublimation and evaporation can be defined as the rate at which molecules leave a surface when exposed to vacuum. The equation of state for this process was given by Langmuir as:

$$W = \frac{P_{\gamma}}{17.14} \left[\frac{M}{T} \right]^{\frac{1}{2}}$$

W = Rate of Evap. or sublim. Gms/cm²/sec

P_{γ} = Vapor Pressure of Material mm Hg. at Temp. T .

M = Molecular Wt.

T = Absolute Temp. °K

or as

$$S = 1.85 \times 10^6 \frac{P_{\gamma}}{\rho} \left[\frac{M}{T} \right]^{\frac{1}{2}}$$

S = Rate of Sublim. Cm/Yr.

ρ = Density of Solid Matl.

These equations are valid when none of the molecules leaving the surface return to it. If pressure is held at levels of 10^{-5} mm Hg. or Lower, the mean free path of vapor molecules between collisions will be several meters or more. (This assumes a reasonable size chamber) It is important that where studies are done to establish sublimation rate, return of sublimed molecules should be minimized. Chamber walls colder than the specimen would help in this case.

TABLE 1

Effect	Pressure Simulation
1. Corona—Electrical Discharge.....	ATM. to 10^{-4} mm Hg.
2. Heat Transfer, Radiation Predominates.....	10^{-4} to 10^{-5} mm Hg.
3. Evaporation, Sublimation.....	$<10^{-5}$ mm Hg. (MFP $>$ mean distance to condensing wall)
4. Decomposition.....	$<10^{-5}$ mm Hg. (High pumping speeds)
5. Liberation of absorbed gas.....	$<10^{-5}$ mm Hg. (High pumping speeds)
6. Liberation of adsorbed (adhesion to surface) and chemisorbed gas.....	10^{-7} to 10^{-16} mm Hg.

DECOMPOSITION

Decomposition amounts to the breaking up of the base compound into elements or simpler compounds, some of which may be more volatile than the original base material. Weight loss by this mechanism may be several orders of magnitude greater than by sublimation or evaporation. Pressure in the range of 10^{-5} mm Hg. or lower is adequate as in the case of evaporation and sublimation.

ABSORBED GAS

The liberation of absorbed gas occurs by diffusion, the rate being dependent upon the temperature of the material and the partial pressure of the particular gas above the surface of the material. Pressures in the range of 10^{-5} mm Hg or lower is considered adequate to reveal this effect.

ADSORBED

Gases held by adsorption and chemisorption adhere to the surface and in combined form throughout the volume of the material. Release of these gases is of ultimate importance in evaluating the effectiveness of the lubrication or rather the seizure tendency of metals that are in mutual contact.

Clean surfaces do not occur until the pressure is reduced below 10^{-7} mm Hg and cannot be maintained for any appreciable time unless the pressure is in the range of 10^{-10} to 10^{-16} torr.

Note: 10^{-6} — 2 sec. (Jaffee, Rittenhouse)
 10^{-9} — 25 min. (Jaffee, Rittenhouse)
 10^{-10} — 10 hrs. (Sheffield)

Thus, the level of vacuum simulation required must be defined in terms of the intent of the investigation to be performed. Where corona, thermal & gross outgassing effects are of major importance simulation at 10^{-7} torr is reasonable. For surface effects studies, pressures below 10^{-12} torr are required to provide a reasonable test duration to observe the effect.

VACUUM SIMULATION

The "state of art" in vacuum technology today has made available large vessels of 50,000 cu. ft. or more for testing of full spacecraft systems. These chambers are capable of ultimate pressures in the range of 10^{-7} to 10^{-9} torr. Smaller chambers

utilizing bakeout and cryogenic pumping techniques with cold walls at 4.2°K are reported capable of pressures below 10^{-13} torr. These chambers are extremely few in number. One of the largest I am aware of has a working volume of ≈ 240 cu. ft. (LRC) (It is important to note at this point that our ability to create vacuum below 10^{-10} torr exceeds reliable means for measuring it.) So far as pumping techniques are concerned the next figure shows the variety most commonly used. (Figure 3)

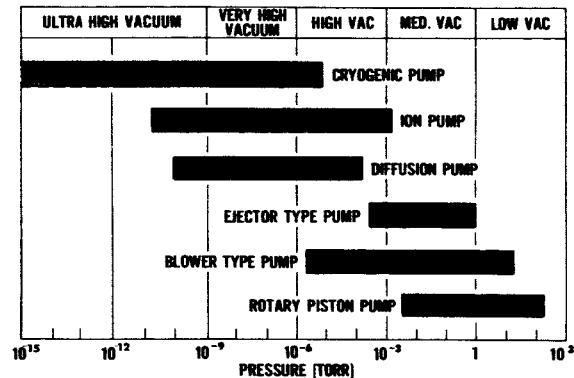


FIGURE 3.—Pumping methods for space chambers.

The first three methods are useful in the viscous and transitional flow regions and are often used for evacuation of the bulk gas load. Of the three, the steam ejector provides the most rapid evacuation. The upper three methods used are molecular pumps. For speed of evacuation, high volumetric flow is important and therefore, it has become common practice to use a combination of molecular pumps.

Good vacuum simulation implies that the vacuum system remove the molecules in such a way that probability of their returning to the test item is small. Ideally, the entire wall surrounding the test item should pump all gases that impinge on it on first impact. Since molecular flow is a random process, a distributed high pumping speed enhances the chamber's capability to deal with heavy outgassing loads. The selection of the pumping method is dependent upon the intended usage of the system and that method which is most economical for the case in question. To amplify on this particular point... in the testing of optical systems where it is established

that vacuum simulation be in the order of 10^{-7} torr or lower (to evaluate outgassing and thermal effects, as well as functional performance of the optical device) several methods of pumping are available. The most significant criteria for selecting the method to be used could well be based upon the pumping system which would reduce the chances for contamination of the optics. A pumping process that does not use a fluid pumping medium may well be the best choice.

For large space chambers capable of testing complete spacecraft systems, the trend has been toward using multiple pumping methods. The variability and magnitude of the gas load to be handled during test negates the possibility of using only one type of pumping system to evacuate the chamber. The major drawback to single method pumping is the time for evacuation and the ultimate pressure capability. Typically, large vacuum systems employ oil diffusion pumps backed by mechanical rotary piston and blower systems.

An important adjunct to diffusion pumping has been the use of cryogenic pumping, and more recently, cryosorption pumping. These methods of pumping have come into prominence concurrently with the development of reliable equipment for producing and controlling liquid helium. Cryogenic pumping is the method of removing gas from the vapor phase by means of condensation on cold walls. Cryosorption utilizes an adsorbent material in intimate contact with the surface of the cryogenically cooled condensing panels. The adsorbent has a high affinity for gases and is capable of more efficient pumping of some gases at temperatures above those required for cryogenic pumping. Cryogenic pumping speeds at ultra high vacuum are dependent upon the rate at which molecules of condensable gas strike the cryogenically cooled surfaces and also on the sticking coefficient of these molecules. The actual pumping speed is given by $S = f\bar{v}$ where \bar{v} is volume flow $\frac{\text{cm}^3}{\text{sec. } 2=2}$ cm^2 and f , the sticking coefficient, is dependent on the surface temperature, vapor pressure of the gas at this temperature, and kinetic energy of the molecules striking the surface. At LN_2 temperatures, 77°K , the main constituents of the cryopumped gas load are H_2O vapor CO_2 some O_2 and hydrocarbons. Pumping

of N_2 and oxygen can be accomplished on surfaces at 20°K . Gases that cannot be pumped at this temperature include hydrogen, helium and neon. Cryogenic pumping of hydrogen and neon requires temperatures at the boiling point of liquid helium, 4.2°K .

Recent experience with adsorbent materials used in cryosorption pumping has shown that at 10^{-8} Torr, hydrogen can be pumped at 20°K at a rate of approximately 7,000 liters per second per square foot of adsorbent area. There is reason to believe that this pumping speed will remain fairly constant until an appreciable area of the adsorbent surface has been utilized. This process appears very promising in its developmental stage. Adsorbent pumping of hydrogen at 20°K has been shown to be 16 times greater than cryogenic pumping at 4.0°K . The cost saving in cryogenic fluid is particularly appealing.

In spacecraft testing, a motivating force behind cryogenic pumping has been the increased demand for facilities that permit evaluation of the spacecraft's thermal systems. Thermal balance evaluation requires that black space be simulated at a level where the net rate of radiant heat exchange between the spacecraft and heat sink is near that which would occur in space. For a spacecraft operating in the range of 300°K , walls at liquid nitrogen temperature, (77°K), are adequate for this purpose. The additional benefit gained is the pumping of gases which condense at this temperature. The need for higher pumping speeds for gases that would not condense at 77°K (Nitrogen, Helium, Neon, Hydrogen) has led to the use of cryogenically cooled panels at 20°K . For economic reasons, removal of the radiant heat from the test object is usually done on a sink at 77°K .

A design approach for a cryogenic array which has gained wide application is shown in Figure 4.

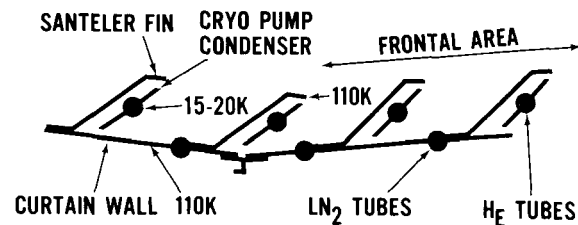


FIGURE 4.—Curtain wall showing "Santeler Array".

The chamber contains an intermediate shroud, cooled to liquid nitrogen temperatures. Fins projecting from the liquid nitrogen shroud are used to support the gaseous helium condensing panels. The fins absorb the thermal energy from the test item, while still providing admittance of the gas load to the 20°K surfaces. The pumping speeds of such an arrangement depend on the exact geometry; however, molecular capture probability of 45 percent has been demonstrated by experiment. (Santler)

VACUUM MEASUREMENT

As was mentioned earlier, methods for creating vacuum has progressed faster than instrumentation to measure it.

Figure 5 shows the range of applications of various gauges available today.

Classifying them according to effect or property measured . . .

- Hydrostatic pressure gauges include: Manometer McLeod
 - Thermal Conductivity Gauges: Thermocouple Pirani
 - Radiometer Knudsen
 - Ionization Gauge Alphanatron Bayert Alpert Cold Cathode Hot Cathode
- Measures ION current produced by any of several sources (figures 6 & 7)
- Mass spectrometers The Omegatron (Cyclotron) (figures 8 through 12)
 - RF Spectrometer Magnetic Sector
 - Time of Flight Quadrupole Mass Filter

Ionization gauges have wide application throughout the high vacuum field. Most recently the trend has been toward extensive development and use of mass analyzers since these instruments are capable of resolving the major constituents and partial pressure of the gas being pumped. Analysis of the pumped gases provides a powerful diagnostic tool for evaluating the performance of the vacuum system and the effect of vacuum on the test item.

The design requirements for mass spectrometers to be used on spacecraft test chambers differ from flight spectrometers for upper atmosphere research in that an extremely wide range of mass number is possible during the course of chamber evacuation and resolution up to mass 200 is

desired. Flight spectrometers are designed for good resolution to about mass 44—(CO₂) (Components such as H₂ He, CH₄, H₂O vapor, neon,

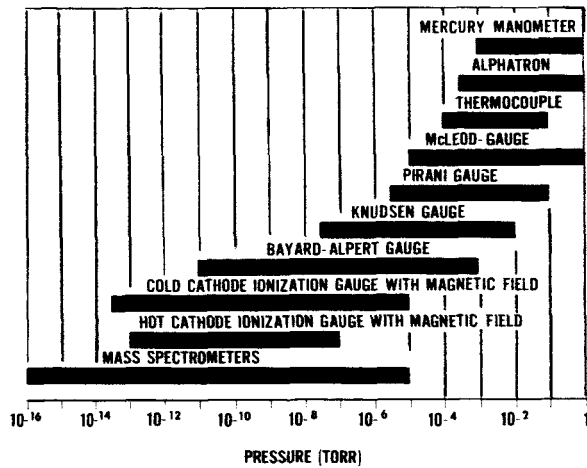


FIGURE 5.—Range of application for various vacuum gauges.

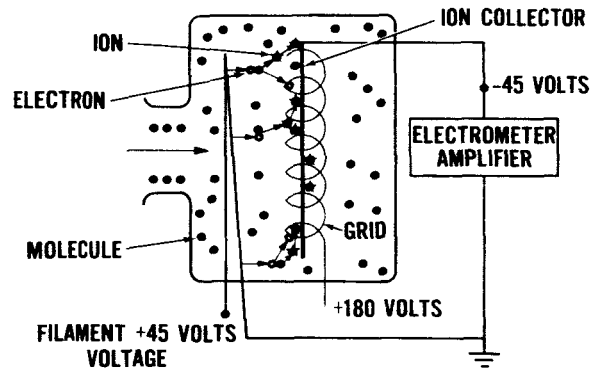


FIGURE 6.—Bayard-Alpert ionization gauge.

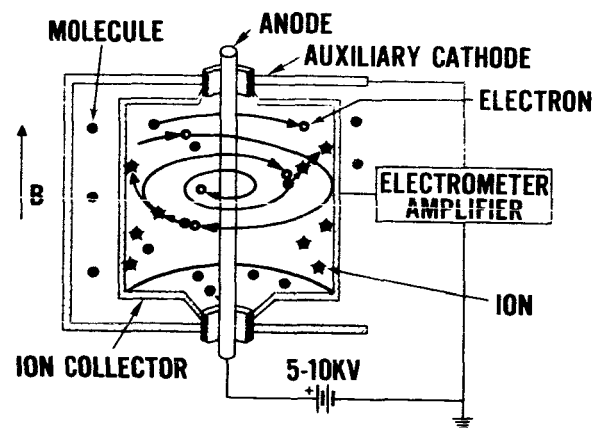


FIGURE 7.—Inverted cold cathode gauge (Redhead).

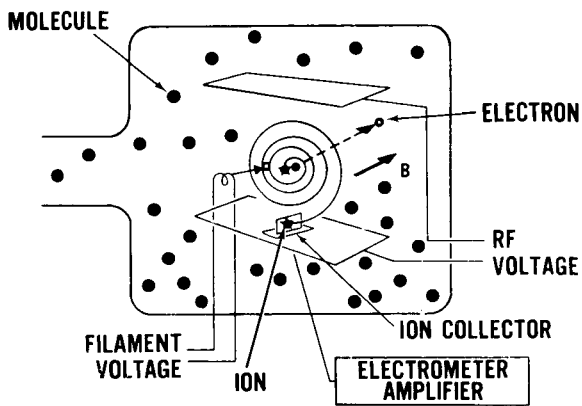


FIGURE 8.—Omegatron mass spectrometer (Cyclotron).

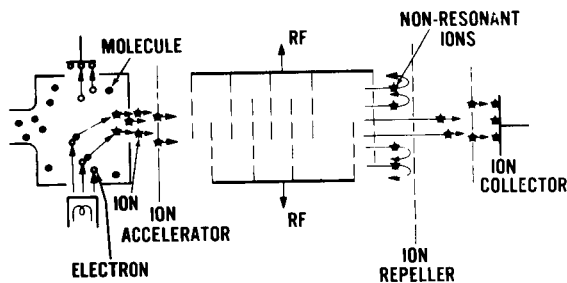


FIGURE 9.—R. F. mass spectrometer.

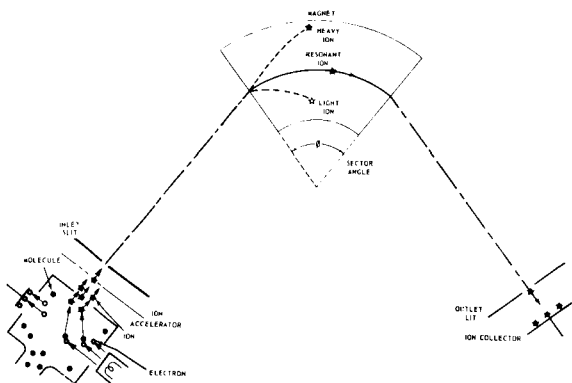


FIGURE 10.—Magnetic sector mass spectrometer.

O₂, N₂, CO, CO₂ lithium, NA, K, Fluorine and chlorine fall within this range).

The sensitivity required of flight spectrometers is of the order required for vacuum chambers for spacecraft testing. Since most testing is done in the range where total pressure is from 10⁻⁸ to 10⁻⁸ torr, a lower level of detectability of a minor constituent whose partial pressure may be as low as

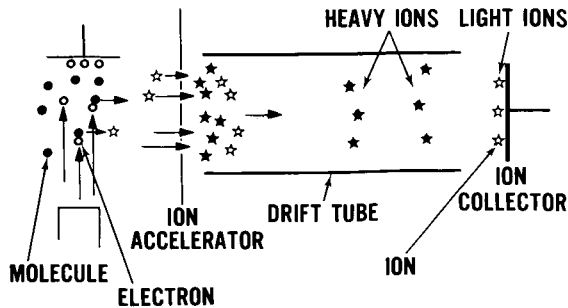


FIGURE 11.—Time of flight mass spectrometer.

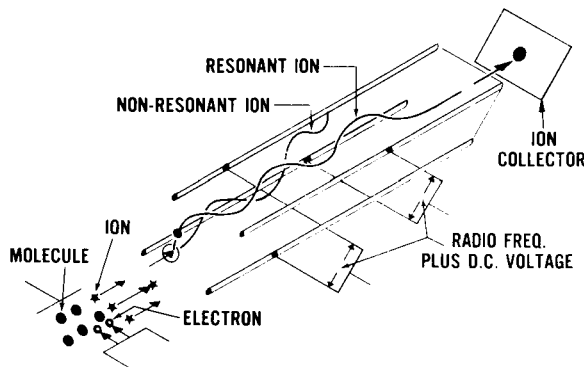


FIGURE 12.—Quadrupole mass filter.

10⁻¹⁶ torr is desired. This corresponds to detecting one molecule/CM³. Recent improvements in ultra high vacuum techniques have led to the increased use of secondary electron multipliers. Multiplier gains of 10⁷ are possible—which in turn makes it feasible to count single IONS arriving at the collector. The stability of the spectrometer electronics and isolation of surface effects in the sensor presents a most difficult problem.

Most space chambers are equipped with multiple gauges since the signal output of no single gauge covers the entire range of interest. Much can be said about vacuum gauging; however, an article written by Dr. Alpert, appearing in "Physics Today", is probably one of the better summaries. He states—"Two new classes of ionization instruments have recently been developed which have ultimate sensitivity below 10⁻¹¹ torr—one class for the measurement of total pressures with ultimate sensitivities in the 10⁻¹² to 10⁻¹⁵ torr range and a second class which involves mass analyzers capable of measuring partial constituents as low as 10⁻¹⁶ torr. However, in the range of pressures below 10⁻¹¹ torr

(and sometimes considerably higher) the surface effects, that is, the chemical and physical interactions which take place at the electrode surfaces of the instruments, begin to be comparable to, or to dominate, the volume effects which they are intended to measure."

In a more operational sense, the vacuum gauge situation at the present time is such that the physical standard for calibration of total pressure gauges is the McLeod gauge with a readable limit of $\approx 10^{-5}$ torr. All methods of calibration below this level are extrapolations from this standard. The use of instruments referred to by Dr. Alpert for measurement in the ultra high vacuum range are extremely delicate and at the present time, are definitely laboratory instruments.

Spacecraft Test Facilities at Goddard Space Flight Center

An example of a large space environment simulator employing multiple pumping is shown in Figure 13. This system is located at the Goddard Space Flight Center and has a usable working volume of approximately 50,000 cu. ft. The vacuum systems consist of 17-52,000 liters per second oil fractionating diffusion pumps which are backed by 8 rotary piston mechanical pumps and 8 blowers. Cryogenic pumping at 77°K is provided by circulating liquid nitrogen through shrouds located internal to the chamber on its top, bottom and sides. Pumping at 20°K is provided by circulating dense helium gas through cryogenic panels located in the lower half of the

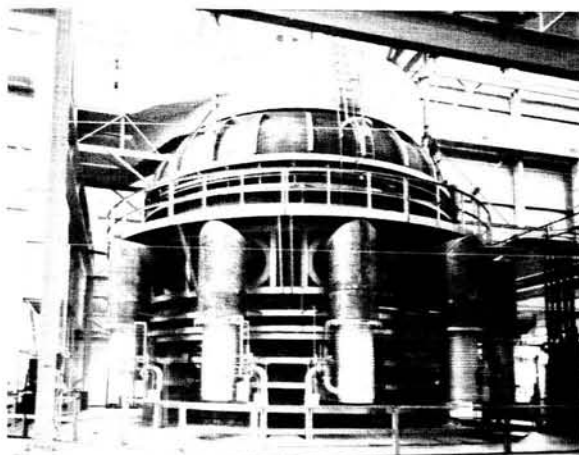


FIGURE 13.—GSFC 30' x 40' space environment simulator.

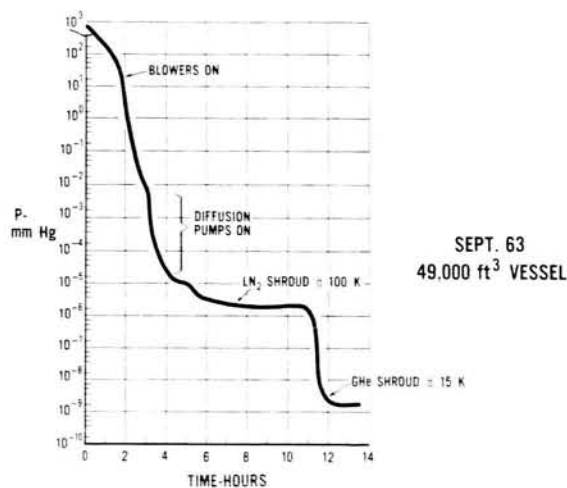


FIGURE 14.—SES Pumpdown Curve. Sept. 63 49,000 ft³ Vessel.

chamber. These panels are nested behind the liquid nitrogen panels as was previously shown. A typical pumpdown curve for this chamber is shown in Figure 14.

Generally speaking, the mechanical pumps reduce the chamber vacuum to approximately 10^{-3} torr, the liquid nitrogen panels and diffusion pumps reduce it to 10^{-7} torr and the dense gaseous helium panels reduce it further until the ultimate is reached. Ultimate pressure in the chamber has been monitored at 8×10^{-10} mm in 13 hours with no vehicle gas load. The pumping speed of this system below 10^{-3} torr is approximately 65 billion liters per second water vapor, and approximately 3 billion liters per second nitrogen. Ninety-two percent of the water vapor pumping capacity comes from the liquid nitrogen heat sink, while approximately 93 percent of the nitrogen pumping capacity is realized from the gaseous helium panels.

Other chambers available at Goddard to perform systems tests range in size from 7' x 8' to 12 to 15'. The spacecraft test experience which I am about to review is based upon tests run in these chambers.

Spacecraft Test Experience

Philosophically speaking, tests performed at the system level are intended to *demonstrate compatibility amongst the operating subsystems*. The level and duration of such tests are not directed toward

establishing the ultimate life or a statistical value for the reliability of the system. Under such a philosophy, the finding of weak links in the system and correcting them is accepted. This philosophy is predicated on the concept that critical materials and subsystems have in fact been tested and that failures are due to system interactions not previously tested for. Testing of prototype spacecraft is done at stress levels exceeding the expected environment while tests of the actual flight spacecraft is done at levels representative of maximum (95 percent confid) level expected during boost phase and orbital flight. Thermal vacuum testing is but one of the tests performed on a spacecraft.

After corona check, the vacuum conditioning is almost always combined with temperature conditioning to evaluate overall heat balance and function of the entire system.

Typically, the vacuum capability of the chamber under a full spacecraft gas load ranges from 10^{-5} through 10^{-7} torr. Failure experience data for 19 spacecraft systems has been accumulated for both the hot and cold conditioning. The temperature level for the hot condition was a maximum of 55°C and for the cold, -20°C . The total duration for hot and cold conditioning of prototype spacecraft was approximately 7 days, while for the flight units, was approximately 5 days. Approximately equal time was spent at hot and cold conditions. For the 19 systems tested, a total of 118 failures were noted. The average failure rate per day of exposure time for each temperature level is represented by Figure 15.

The failures observed were primarily due to component failures. The distribution of failures due to mechanical design, and thermal causes were approximately equal. It is significant to point out that the highest incidence of failure in the system occurs during the early conditioning period (first two days) after which the rate is reduced to a level of approximately one per day.

On the basis of these data, it is not possible to predict the reliability of the system for long term life. It is reasonable to predict that the system should not suffer from early orbital failures. This contention is supported by the orbital performance of satellites already tested and flown. (Remarks on the orbital performance representative of 14 of the systems tested is as follows:

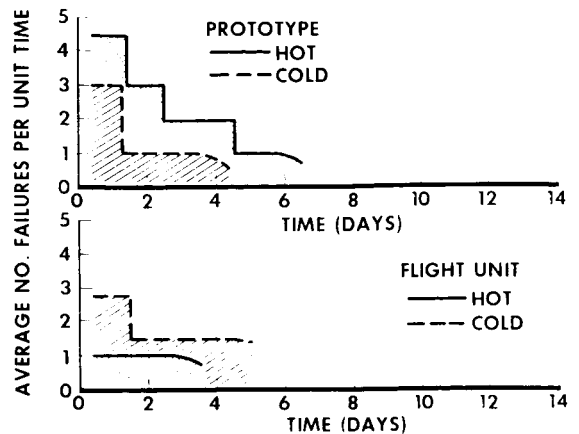


FIGURE 15.—Summary of spacecraft failures vs. time in thermal-vacuum tests.

Orbital performance not available for S-52-2 systems, S-48-3 systems).

This is but a small sample of the performance history of spacecraft flown by GSFC.

The data available on spacecraft performance in orbit is difficult to generalize on. Failures have occurred in both the experiments and some in-line equipment. However, there has been greater incidence of failure in the case of the experiments. Most post launch failure analysis studies have indicated the probable cause was failure of electronic components. However, the exact cause has most often not been identified. In addition, some equipment, particularly experiments have been flown where there had been an earlier history of poor performance in tests. Flight of these items was done on a high risk basis. In most cases, usable data was obtained. The exact role the vacuum environment played in the orbital failures is not known.

Turning to a more direct area of concern... 10 Spacecraft, Explorer XI, XIV, XV, OSO-1, Tiros I, II, III, Tiros IV, VI, Ariel I, selected at random were reviewed for possible orbital failure attributed to bearing or friction problems.¹ The performance data shows the life of bearing elements used in tape recorders and camera subsystems has been in excess of $7\frac{1}{2}$ months; motor failure has occurred in less than 4 months. These data are for sealed and oiled tape recorders. Only one spacecraft reviewed had bearings and friction-

¹ Private communication with Mr. C. Vest of Mech. System Branch.

Spacecraft Title	Design Life	Orbital Life	Remarks (as of Feb. 20, 1964)
P-14 Explorer X	60 hours	60 hours	Sphere temperature increased continuously for several hours after satellite injection. The RB vapor magnetometer became intermittent. Failure was attributed to change in absorbtivity of the sphere caused by outgassing of the nose cone.
S-3 Explorer XII	1 year	112 days	All experiments functioned properly during 112 day period. The exact cause of failure has not been determined.
S-51 Ariel I	1 year	130 days	One experiment—Lyman Alpha failed on launch. The experiments and spacecraft subsystems operated satisfactorily until early July (75 days). On July 12, the system began to go into 18 hour periods of undervoltage. Intermittent operation from that time on, has been experienced and is attributed to damage to the solar array by the high altitude detonation which occurred on July 9, 1962. Partial operation of the experiments and systems has been experienced and as of April 17, 1963 the spacecraft had a total equivalent 24 hour operating time of 174 days. As of February 20, 1964, the satellite completed its 137th undervoltage condition and is sending data essentially from one experiment approximately 50 percent of the time that the satellite is operating. It is interesting to point out that the tape recorder motor and associated bearings is still operating.
IMP Explorer XVIII	1 year	Operating satisfactorily as of 2/20/64 (90 days)	One experiment failed 20 hours after launch due to failure of mechanical programmer to step. Intermittent operation of the programmer has been experienced since this time. All other experiments and subsystems are functioning properly.
S-3a Explorer XIV Launch 10/2/62 OK	1 year	Operation Satisfactory for 300 days	First malfunction occurred after 75 days—could have been either an experiment (I & E) or the spacecraft encoder. Intermittent operation continued for two weeks then cleared up—same malfunction occurred again after 5½ months in orbit—lasted 16 hours. The spacecraft and experiments continued to operate satisfactorily from that time on.
S 52 S-48	Not yet flown		

al surfaces directly exposed to the space vacuum environment. In this case, MOS_2 was used as the lubricant. Operation was in excess of 17 months from data available for the survey. The orbital vacuum was approximately 10^{-8} torr (600 Km).

The lack of diagnostic instrumentation to monitor the change in frictional properties of components precludes any analysis except on a go-no-go basis. On the basis of these data, the performance of sealed and oil mechanical subsystems

appears to be somewhat better than the performance of electronics. Orbital data are not sufficient to draw any conclusion for frictional surfaces exposed directly to the space vacuum.

So far as surface effects are concerned, spacecraft have run some 10°C to 15°C hotter than the maximum temperature predicted by the thermal analysis. It has not been established whether this is caused by the environment, contamination of the coating, or by the accumulation of tolerances involving α/ϵ determination of the coatings, the fabrication process or the computational routines used for the thermal design.

Despite these uncertainties, spacecraft have been successful. On the basis of utilizing current design experience, the life expectancy of scientific spacecraft equipment is at least 6 months, and probably a year, where good quality control is maintained for both the parts and workmanship.

Observations & Future Emphasis

Based on past performance and State-of-Art in vacuum testing the following point of view is offered:

1. Tests being performed at the spacecraft systems level appear adequate to isolate possible early failures due to corona, thermal causes and incorrect parts application.

2. Long term effects of sublimation, and decomposition are not being tested for at the system level of assembly and at the present level of spacecraft instrumentation, neither is our means to diagnose this as a cause for failure in orbit.

3. The effects of vacuum on determining bearing life and wear characteristics of moving surfaces in mutual contact is not conclusive. Orbital data is primarily for sealed and oiled applications. A real problem still exists for applications where sealing cannot be done and particularly where

contamination cannot be tolerated. The contamination problem has greatest significance in the case of optical surfaces and electrical contacts.

4. So far as vacuum exposure of spacecraft equipment is concerned:

- a. Systems tests at 10^{-5} to 10^{-7} mm Hg appear reasonable at this time.
- b. Long term effects due to decomposition and outgassing should be attacked via tests on materials and electronics components and subsystems. Failure mechanism should be defined. Vacuum of 10^{-7} to 10^{-16} torr should be developed and utilized in these studies.
- c. The bearing and wear problem deserves closest attention and testing should be performed in the UHV range, 10^{-10} to 10^{-16} mm Hg.

5. So far as vacuum techniques are concerned:

- a. Reliable high speed pumping systems should be developed that will reduce the contamination problem. (Chamber contributed and self contamination of the test item due to rebounding molecules).
- b. Creation of UHV in small systems for parts and component testing should be pursued.
- c. Instrumentation to measure both type and quantity of gas constituents in a system from 10^{-7} through 10^{-16} torr should be developed to the point where they are reliable and easy to operate. Means for calibrating such instruments should be developed concurrently.

6. So far as test techniques are concerned, a good deal of effort should be directed toward understanding basic effects and establishing test levels and durations consistent with revealing these effects.

LAUNCH ENVIRONMENT PROFILES FOR SOUNDING ROCKETS AND SPACECRAFT*

W. J. NEFF AND R. A. MONTES DE OCA†

The Launch Phase Simulator, which is being developed by the Goddard Space Flight Center must provide several different functions—individually or in combination—for simulating the environmental conditions encountered by the payload of sounding rockets and unmanned spacecraft from liftoff to last stage burnout. A survey has been made of the sounding rockets and spacecraft launch vehicles that NASA is utilizing or will use in the near future, and representative profiles are plotted to show their time history and interrelation of parameters. The following parameters are included in the profiles: acceleration, (longitudinal and lateral), pressure, vibration, acoustics, heating, and spin. From these profiles, an overall envelope of launch environmental parameters has been constructed.

INTRODUCTION

An environmental test facility called the Launch Phase Simulator is being developed by Goddard Space Flight Center (GSFC). The purpose of this facility is to provide simulation of the environmental conditions encountered by the payloads of sounding rockets and unmanned spacecraft during the launch phase of flight. The most important feature of the launch environment is that mechanical input (vibration, shock, acoustics, spin) and thermal-vacuum inputs are imposed on a payload in a varying acceleration field. The fundamental design criterion of the Launch Phase Simulator is this combination of environmental inputs.

To provide simulation of the environmental parameters acting during the launch phase, the magnitudes of these parameters and the extent to which they are combined must be known. Therefore, a survey has been made of the sounding rockets and unmanned spacecraft launch vehicles that NASA is using or will use in the near future. The results of this survey are presented herein as launch environment profiles for representative missions. The parameters given in these profiles include acceleration, pressure, vibration, acoustics,

heating, and spin. From these profiles, an overall envelope of the launch environment parameters can be constructed.

The launch profiles have been prepared to assist in developing the Launch Phase Simulator facility. The material has been organized to fulfill the need for an overall envelope of launch environment parameters. However, this material is not considered suitable for specific applications, such as programming individual spacecraft environmental tests.

LAUNCH ENVIRONMENT PROFILES FOR SOUNDING ROCKETS

Launch environment profiles for the following sounding rockets have been included in the study:

Name	Designation
Nike-Asp	ASPAN 150
Aerobee 150 A	---
Javelin	Argo D-4
Aerobee 300 A	Spaerobee
Iris	52 KS 3850, Marc 13A1
Nike-Cajun	---
Journeyman	Argo D-8
Journeyman B	TS609 (Blue Scout, Jr.)

*Published as *NASA Technical Note D-1916*, January 1964. This work was accomplished for the Goddard Space Flight Center, Test and Evaluation Division, by Booz-Allen Applied Research, Inc., Bethesda, Maryland, as a task under contract NAS 5-2415.

†Booz-Allen Applied Research, Inc.

The data gathered for this work are incomplete with regard to some parameters, and a summary of availability is presented in Table 1. The

following paragraphs contain general comments for each of the parameters.

Table 1
Availability of Data on Launch Profile Parameters for Sounding Rockets.*

Sounding Rocket	Acceleration, Longitudinal	Acceleration, Lateral	Altitude (Pressure)	Vibration	Acoustics	Heating	Roll Rate (Spin)	Dynamic Pressure	Mach Number
Nike-Asp	x		x			x	x		
Aerobee 150 A	x		x			x	x		
Javelin	x		x	x	x	x	x	x	x
Aerobee 300 A	x		x			x	x		
Iris	x		x	x		x	x		
Nike-Cajun	x		x			x	x		
Journeyman	x		x			x	x	x	x
Journeyman B	x		x	x		x	x		

*Availability of data is denoted by the letter "x."

1. Acceleration

(a) Longitudinal: Complete data are available, either calibrated or tested. Test results are in good agreement with the calibrated data. Table 2, which shows the maximum acceleration achieved for different stages of

various rockets with nominal payloads, illustrates the wide range of maximum accelerations and time of occurrence.

(b) Lateral: No record of lateral acceleration for these rockets was found during the survey.

Table 2
Maximum Stage Acceleration Versus Time for Sounding Rockets.*

Sounding Rocket	1st Stage		2nd Stage		3rd Stage		4th Stage	
	Accel. (g)	Time (sec)	Accel. (g)	Time (sec)	Accel. (g)	Time (sec)	Accel. (g)	Time (sec)
Nike-Asp	39.6	2.8	<u>49.3</u>	23.1	---	---	---	---
Aerobee 150 A	<u>11</u>	2.5	9.1	55	---	---	---	---
Javelin	17.5	4	16.3	13	<u>34.8</u>	28	12.6	87
Aerobee 300 A	10.9	2	8.5	51	<u>57.5</u>	54	---	---
Iris	<u>13.9</u>	1	9.6	52	---	---	---	---
Nike-Cajun	41.7	2.8	<u>64.1</u>	24	---	---	---	---
Journeyman	9.6	2.5	14.8	43	<u>33.8</u>	51	12.7	97
Journeyman B	7.6	2.6	7.1	72	<u>27.3</u>	105	12.9	146

*The maximum acceleration for each rocket is underlined.

2. Altitude (pressure)

Altitude versus time and the corresponding atmospheric pressure profile are depicted for each rocket. These profiles vary with payload.

3. Vibration

Vibration profiles are not plotted because of the lack of appropriate data. However, vibration inputs may be expected during the whole launch phase and will be particularly severe during the burning time of certain stage motors. The list of data sources (Bibliography, page 799) includes several references for vibration data.

4. Acoustics

Data on acoustic excitation have been presented for the Javelin (Argo D-4) only.

5. Heating

(a) External: The external temperatures plotted for each sounding rocket have been presented for an assumed low launch angle (approximately 70 degrees) and a nominal payload. At higher launch angles the temperature will be lower.

(b) Internal: Heating of the payload compartment after launch is a function of the compartment temperature prior to launch, vehicle flight path, duration of flight, heat output of the payload, and compartment configuration. Since the powered flight time for sounding rockets is very short (in general, less than 100 sec), the payload compartment temperature rise will be small.

6. Roll rate (spin)

Roll rates are presented as a function of time. There is a wide range, from 0 spin to 12 rps, occurring at different values of acceleration and pressure.

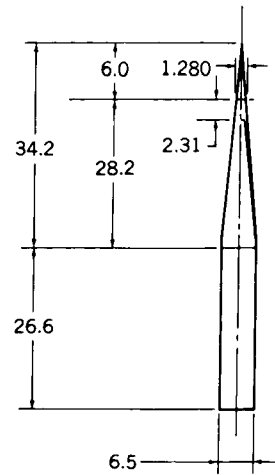
7. Dynamic pressure and Mach number

Dynamic pressure and Mach number are also plotted for those rockets on which data are available.

Figures 1 through 8 give the launch profiles for the sounding rockets. Descriptive data for each rocket accompany the figures. Profiles are given for nominal payload unless otherwise specified.

Nike-Asp

Designation: ASPAN 150
 Launch vehicle
 1st: Nike M5 Booster
 2nd: Asp-I
 Payload size: See sketch
 Payload weight (lb)
 Minimum: 25
 Nominal: 50
 Maximum: 100



(Dimensions in inches)
 (Nominal available payload volume = 0.64 cu ft)

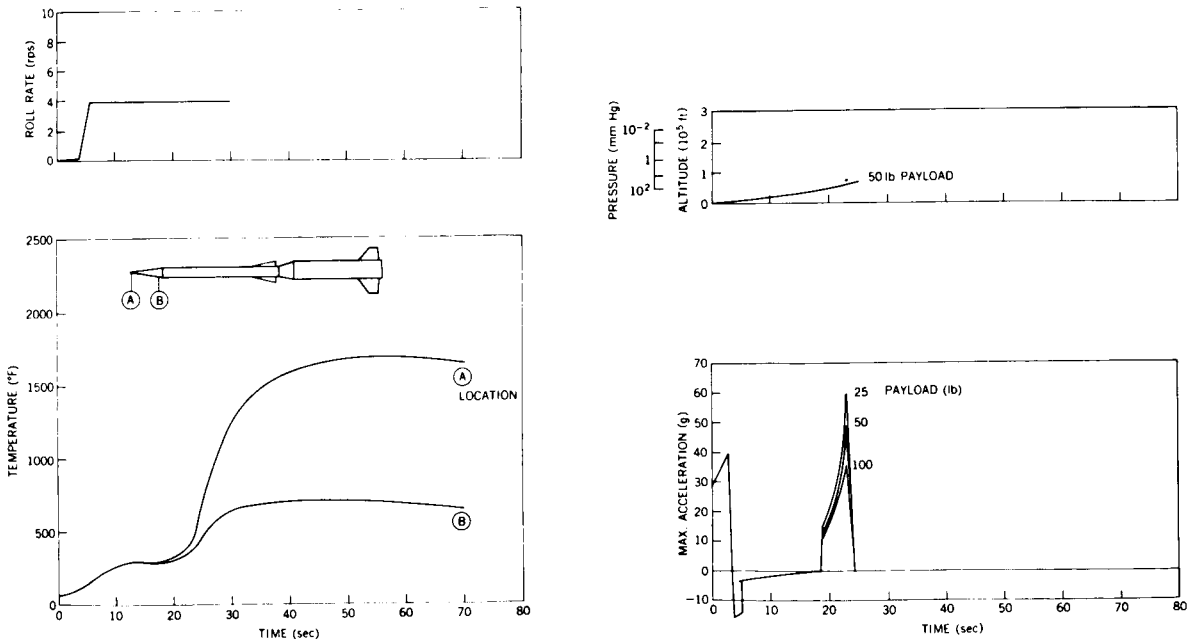
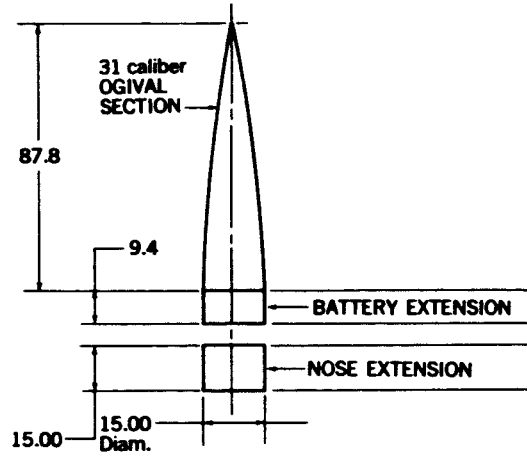


FIGURE 1.—Launch profile for Nike-Asp sounding rocket.

Aerobee 150 A

Launch vehicle
 1st: 2.5 KS 18000
 2nd: Aerobee 150 A
 Payload size: See sketch
 Payload weight (lb)
 Minimum: 100
 Nominal: 200
 Maximum: 300



(Dimensions in inches)
 (Nominal available payload volume, including extensions = 6.90 cu ft)

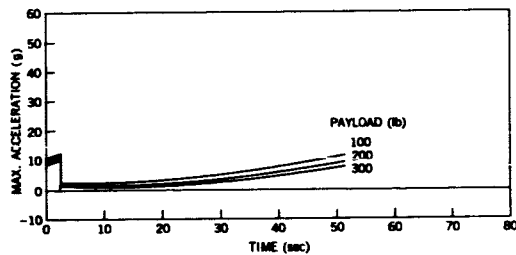
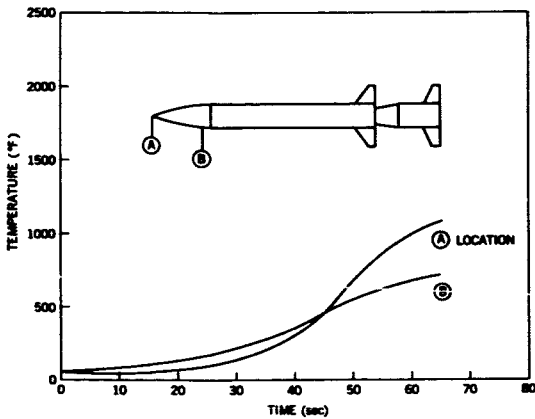
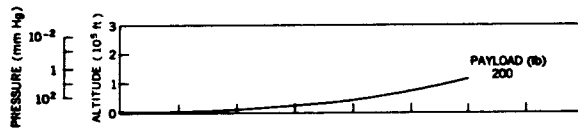
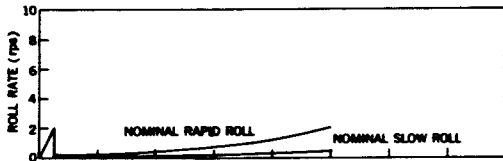
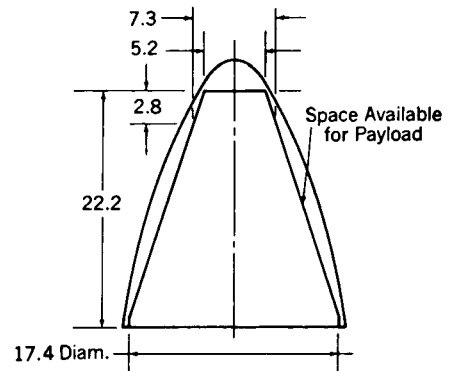


FIGURE 2.—Launch profile for Aerobee 150 A sounding rocket.

Javelin (Argo D-4)

Designation: Argo D-4
 Launch vehicle
 1st: Honest John M6 Booster
 2nd: Nike M5 Booster
 3rd: Nike M5 Booster
 4th: Altair (X248-A6)
 Payload size: See sketch
 Payload weight (lb)
 Minimum: 20
 Nominal: 125
 Maximum: 175



(Dimensions in inches)
 (Nominal available payload volume = 1.4 cu ft)

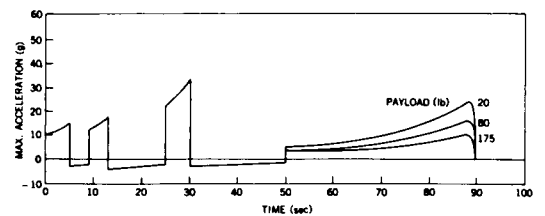
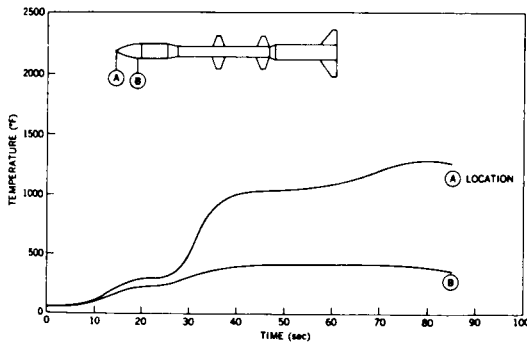
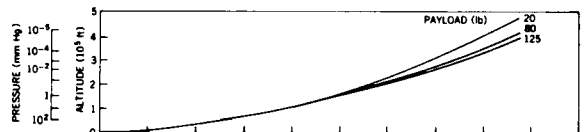
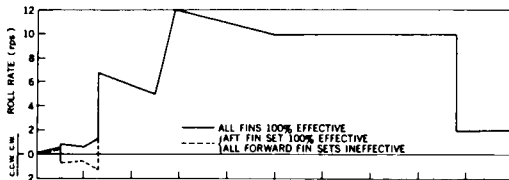
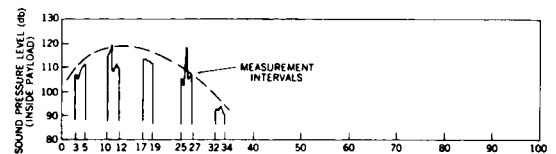
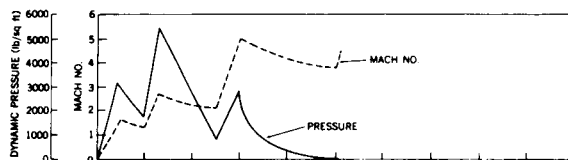
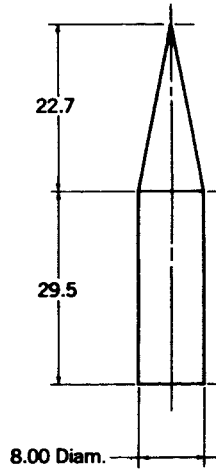


FIGURE 3.—Launch profile for Javelin sounding rocket (Argo D-4).

Aerobee 300 A

Designation: Spaerbee
 Launch vehicle
 1st: Aerojet (2.5 KS 18000)
 2nd: Aerobee 150 A
 3rd: Sparrow III (1.8 KS 7800)
 Payload size: See sketch
 Payload weight (lb)
 Minimum: 20
 Nominal: 60
 Maximum: 100



(Dimensions in inches)
 (Nominal available payload volume = 0.9 cu ft)

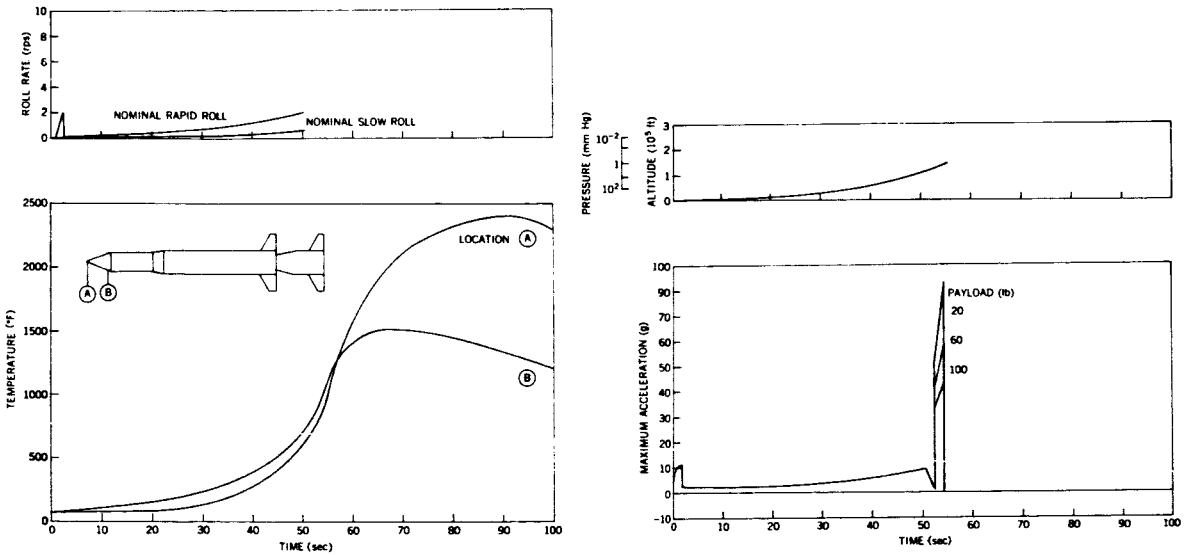
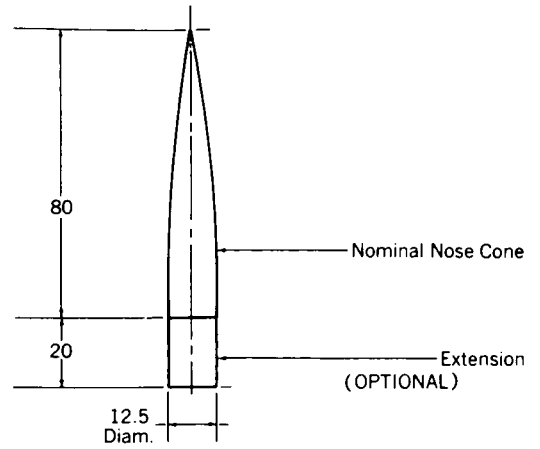


FIGURE 4.—Launch profile for Aerobee 300A sounding rocket.

Iris

Designation: 52 KS 3850, Marc 13A1
 Launch vehicle
 1st: 0.8 KS 18800, Marc 14B1
 2nd: 52 KS 3850
 Payload size: See sketch
 Payload weight (lb)
 Minimum: 75
 Nominal: 100
 Maximum: 200



(Dimensions in inches)

(Nominal available payload volume = 4.5 cu ft;
 with 20-in. cylindrical extension = 5.8 cu ft)

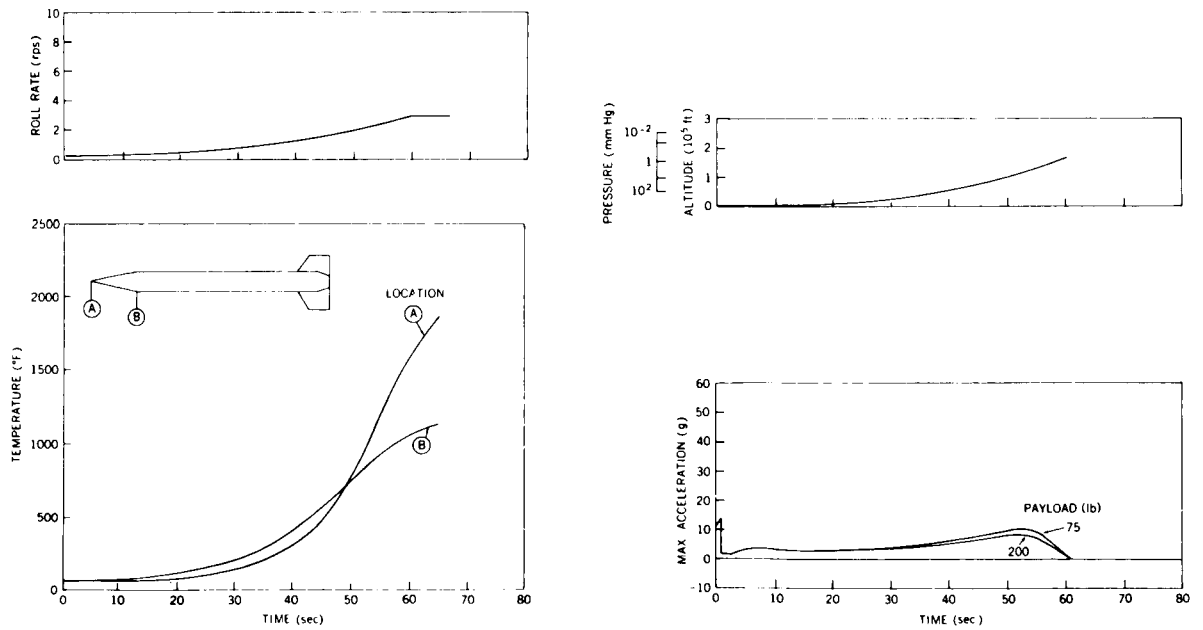
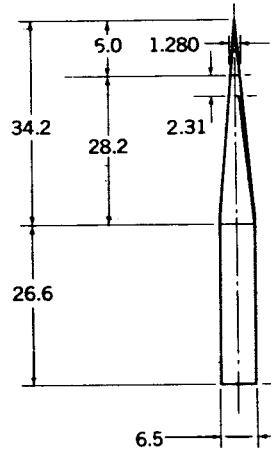


FIGURE 5.—Launch profile for Iris sounding rocket.

Nike-Cajun

Launch vehicle
 1st: Nike M5 Booster
 2nd: 2.8 KS 8000 Cajun
 Payload size: See sketch
 Payload weight (lb)
 Minimum: 25
 Nominal: 50
 Maximum: 100



(Dimensions in inches)
 (Nominal available payload volume = 0.64 cu ft)

ROLL RATE (rpm)
 The nominal roll rate is zero
 Note, however, that random perturbations such as launcher tip off, fin misalignment, and winds may produce scatter in roll orientation.

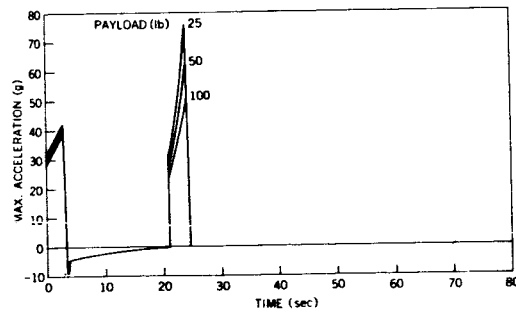
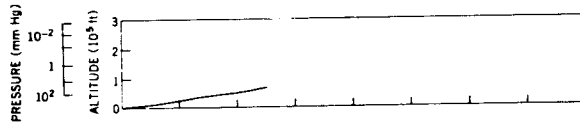
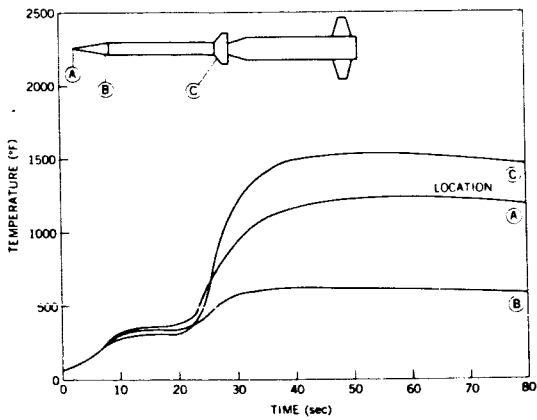
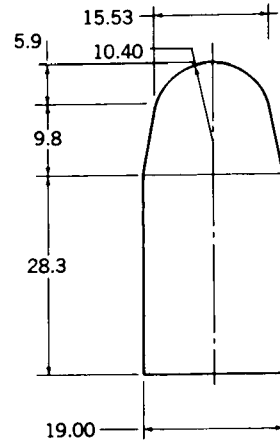


FIGURE 6.—Launch profile for Nike-Cajun sounding rocket.

Journeyman (Argo D-8)

Designation: Argo D-8
 Launch vehicle
 1st: Pollux XM-33E6 plus two auxiliary
 Recruit (XM-19E1)
 2nd: Lance XM-25
 3rd: Lance XM-25
 4th: Altair X248-A6
 Payload size: See sketch
 Payload weight (lb)
 Minimum: 75
 Nominal: 125
 Maximum: 175



(Dimensions in inches)
 (Nominal available payload volume = 7 cu ft, estimated)

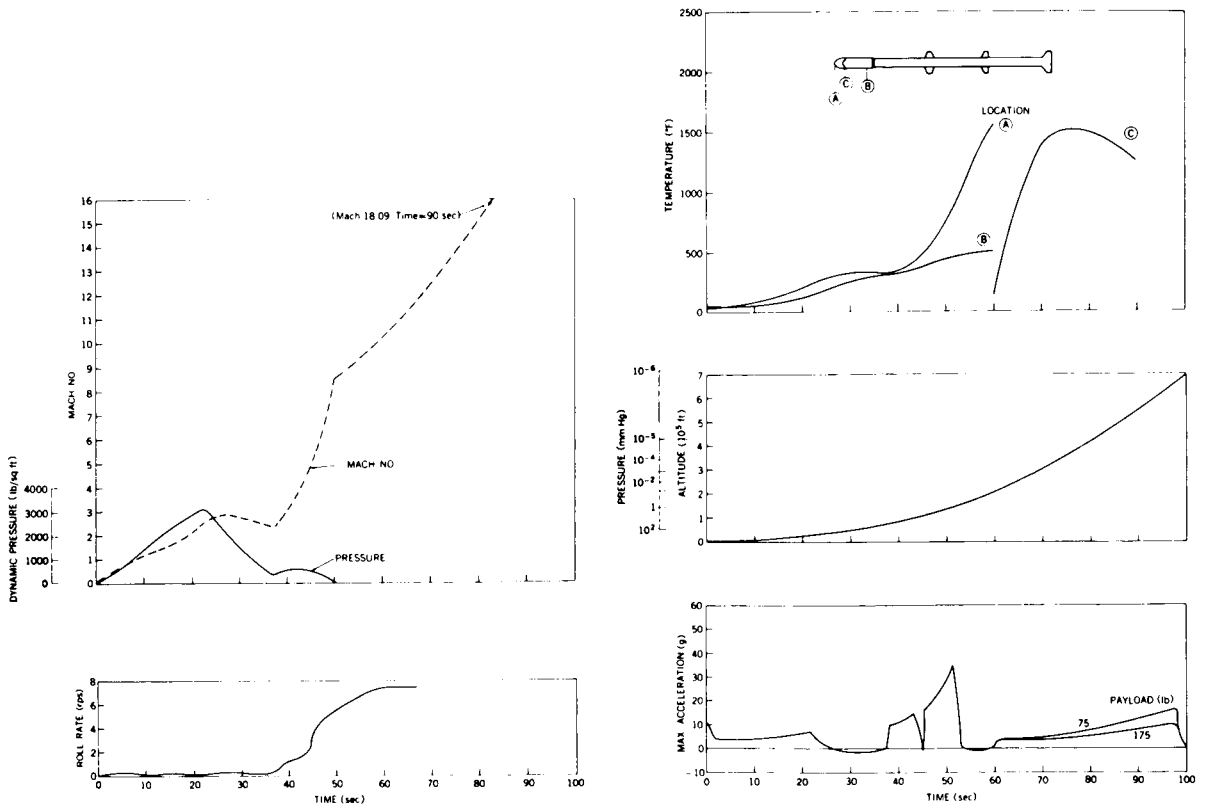


FIGURE 7.—Launch profile for Journeyman sounding rocket (Argo D-8).

Journeyman B

Designation: TS 609A (Blue Scout, Jr.)

Launch vehicle

1st: Castor (XM33-E8)

2nd: Antares (X254-A1)

3rd: Alcor (AJ10-41)

4th: Cetus (NOTS 100A)

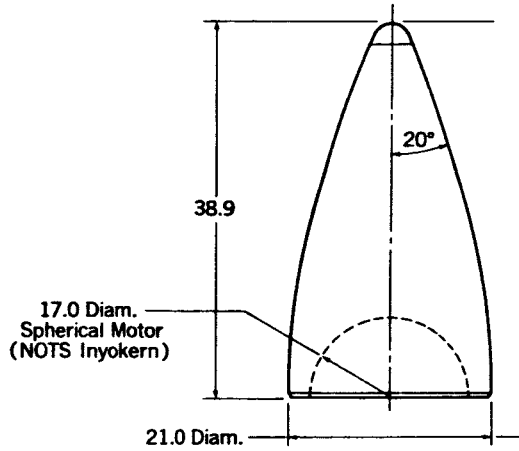
Payload size: See sketch

Payload weight (lb)

Nominal: 30

Intermediate: 50

Maximum: 100



(Dimensions in inches)
(Nominal available payload volume = 2.8 cu ft, estimated)

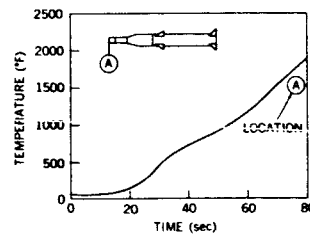
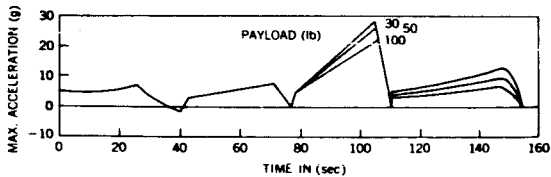
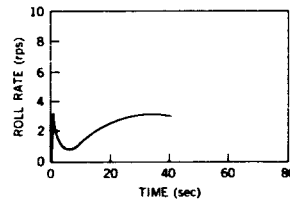
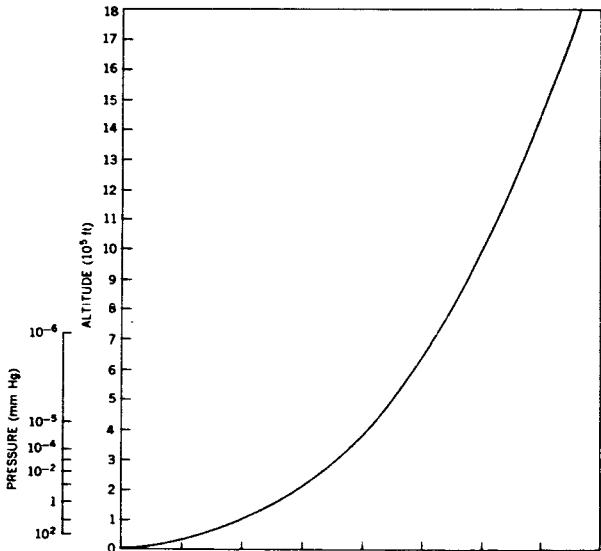


FIGURE 8.—Launch profile for Journeyman B sounding rocket.

LAUNCH ENVIRONMENT PROFILES FOR SPACECRAFT

Launch environment profiles for the spacecraft presently programmed for development by GSFC are dependent to a large degree on the characteristics of the launch vehicles, with smaller variations over the range of weights and trajectories of the particular spacecraft. This overall launch profile review is based, therefore, on available data relating to the following launch vehicles:

Scout	Atlas-Agena B
Delta	Centaur
Thor-Agena B	

The general configurations of these launch vehicles are well-known and not repeated here. Among the specific characteristics affecting launch profiles, the following should be noted:

1. All launch vehicles are liquid propellant except Scout (four-stage, all solid propellant) and the third stage of Delta.
2. Delta and Thor-Agena B have essentially the same first stage.
3. Thor-Agena B and Atlas-Agena B have essentially the same second stage.
4. Atlas-Agena B and Centaur have essentially the same second stage.
5. Scout and Delta have essentially the same final stage.

All the launch vehicles above may utilize variable burning schedules for the upper stages, depending on the particular orbit or trajectory desired. This will affect the time history of acceleration and vibration, and the relation of pressure to acceleration and vibration (and, to a lesser extent, the relation of temperature to vibration and acceleration). It will not affect the relation of vibration to acceleration. Almost all the acoustic and aerodynamic vibration and heating inputs occur during the first stage burning and are nearly the same for all launch payloads, varying only with shroud characteristics.

Representative launch profiles for Scout, Delta, Thor-Agena B, Atlas-Agena B, and Centaur vehicles are given in Figures 9 through 13.

Figures 9 through 13 present maximum longitudinal acceleration in g's, altitude in feet, and dynamic pressure in pounds per square foot. The

notations indicate staging, Mach 1, ejection of aerodynamic shroud, and spin-up, as well as levels and sources of vibration, acoustic, and heating inputs. The data sheets accompanying each plot give further information on the profiles, including magnitudes of vibration and acoustic inputs if available. The anticipated launch pad environment prior to liftoff is also given if known.

The data sheets also include a listing of typical NASA payloads launched or programmed for each launch vehicle. Manned spacecraft are not included, since these are not under the cognizance of GSFC.

The following notes apply to Figures 9 through 13:

1. Representative profiles were chosen primarily on the basis of data availability. In general, the continuous function data (acceleration, altitude, dynamic pressure) are from measured or computer trajectory print-outs for the selected spacecraft—launch-vehicle combinations. The other environmental data (vibration, acoustics, heating, and spin) are from GSFC internal reports and correspondence, and from interviews with project personnel.

2. Continuous function data are plotted against time from liftoff, in seconds. Diagonal rulings indicate a time break. Event timing is given in the horizontal bar chart of operational phases just above the plot. In addition to the plotted functions of acceleration (longitudinal), altitude, and dynamic pressure, the ambient pressure can be obtained from the altitude scale by using the following conversion table:

Pressure (mm Hg)	Altitude	
	(ft)	(10 ⁵ ft)
10 ²	49,000	0.49
10 ¹	98,000	0.98
10 ⁰	147,500	1.47
10 ⁻¹	213,000	2.13
10 ⁻²	255,000	2.55
10 ⁻³	295,000	2.95
10 ⁻⁴	337,000	3.37
10 ⁻⁵	425,000	4.25
10 ⁻⁶	720,000	7.20
10 ⁻⁷	1,145,000	11.45
10 ⁻⁸	1,735,000	17.35

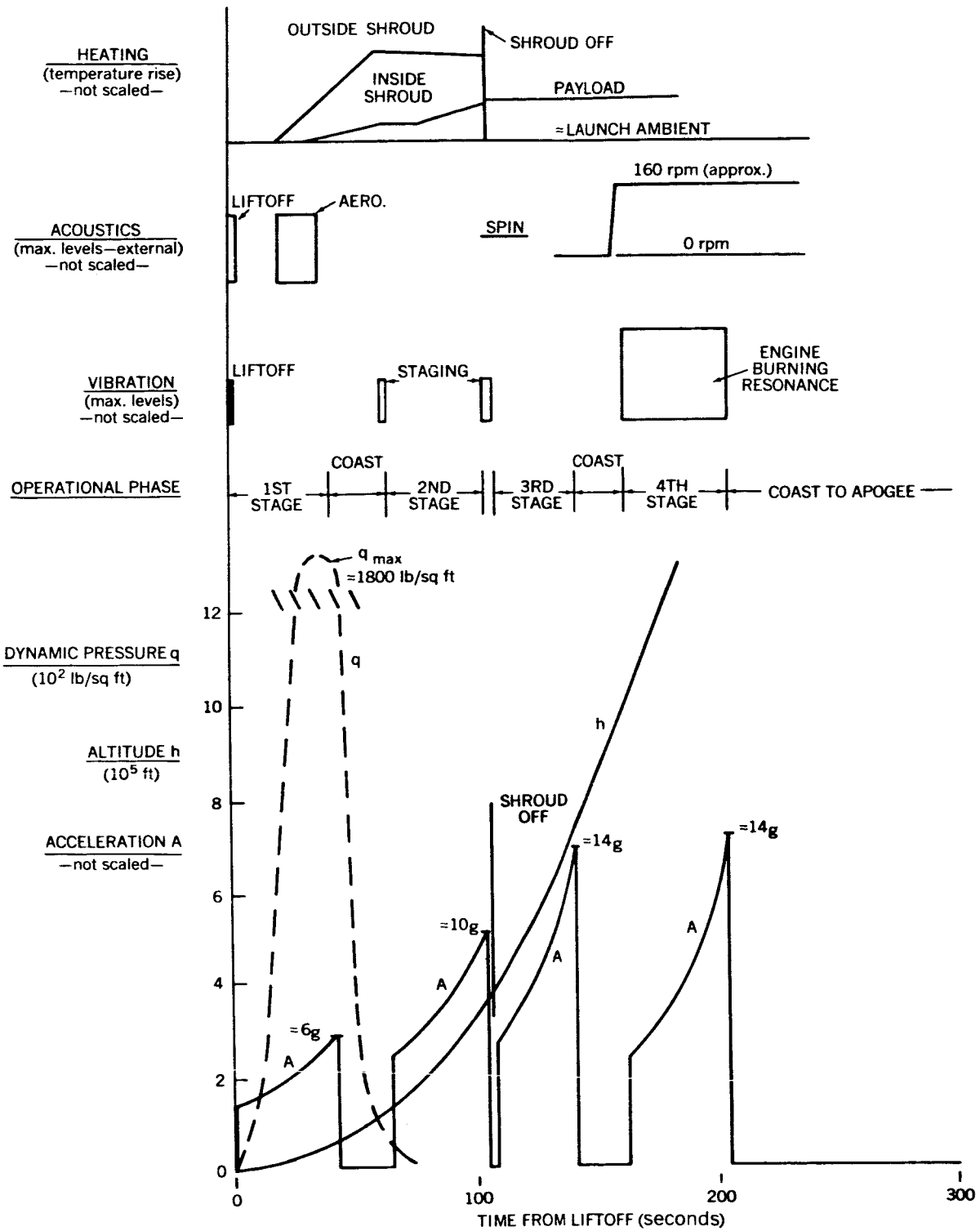


FIGURE 9.—Representative launch profile for Scout vehicle (St-9/p 21a profile; payload, 150 lb.).

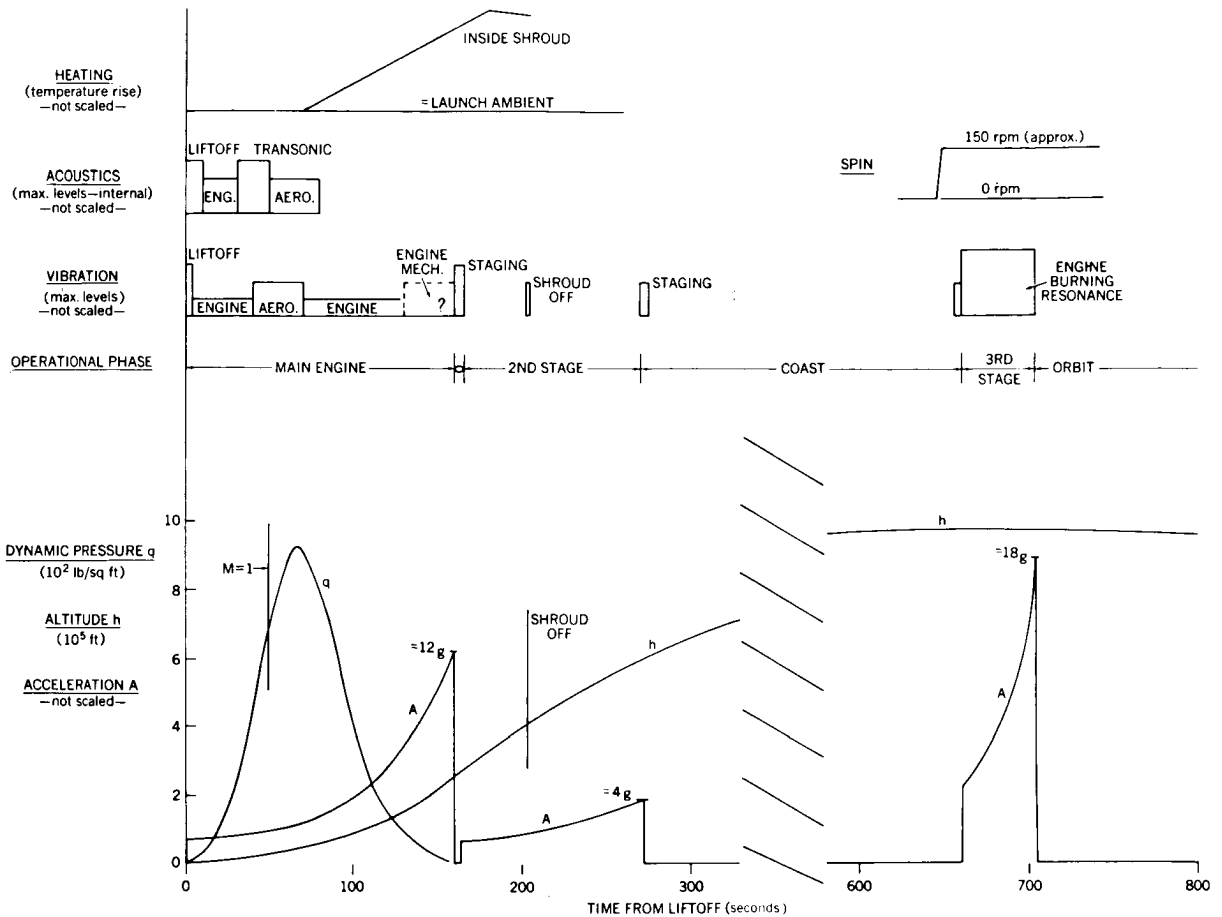


FIGURE 10—Representative launch profile for Delta vehicle (S-3 profile; payload, 85 lb.).

3. Other environmental data are given only in qualitative form, to show how they fit into the time profile of the launch. Height of the bars for vibration and acoustic levels is indicative of order of magnitude and is consistent throughout the five profiles. Further information on these parameters (including quantitative values where available) is given on the data sheets accompanying each figure.

4. Variations of payloads and trajectory requirements from those plotted in the figures will affect the early portion of the profile very little. Variations in payload weight will affect *final* stage burnout acceleration and altitude. Variations in trajectory will affect altitude and magnitude of heating. Variation in vehicle configuration (primarily shroud changes) will affect heating and acoustics. In particular, payload changes within

a given vehicle configuration will have little effect on all profile data of the vehicle's first stage (1st and 2nd stages for Scout). Consequently most vibration, acoustic, and heating inputs—which occur early in the launch phase—are similar for all payloads. For this reason it was considered acceptable to extrapolate vibration, acoustic, and heating data for payloads other than those used for the continuous function data.

5. Vibration data are generally referenced to the payload support structure. In particular, the vibration data for Delta, Thor-Agena, and Atlas-Agena are based on this reference. Vibration data for Scout are from measurements within the lower payload section, but this is at the same location relative to the scientific payload.

6. Acoustic data are referenced to the payload location of the launch vehicle. Data labeled

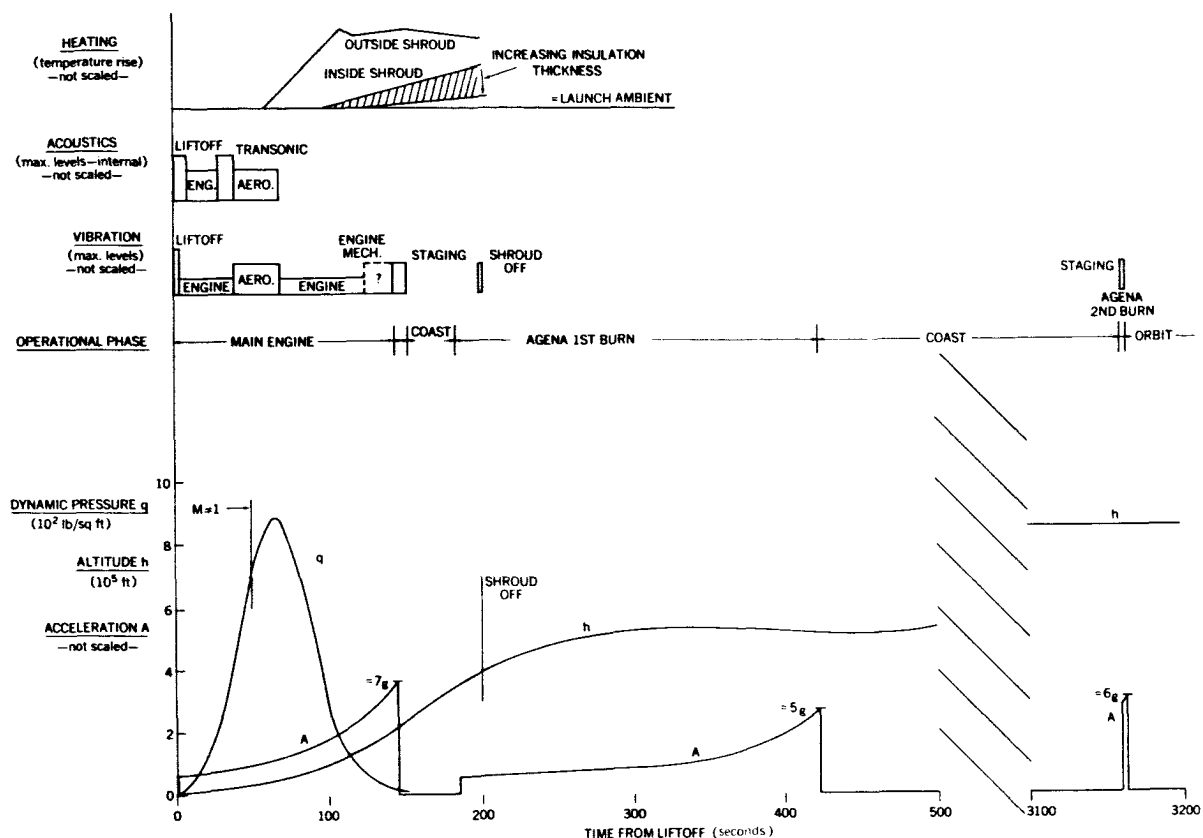


FIGURE 11.—Representative launch profile for Thor-Agena B vehicle (POGO profile; payload, 1000 lb.).

“external” are referenced to the outer surface of the shroud. Data labeled “internal” are referenced to the inner surface of the shroud.

7. Heating data are given for the external and internal surfaces of the nose cone (shroud) at the junction of the conical and cylindrical surfaces (reference point 2 feet aft of nose, for Scout). The notes accompanying the Atlas-Agena profile indicate the wide variation in external shroud temperature at different locations.

Scout (Representative Profile)

Scout.....Four-stage, solid propellant
 Acceleration.....Longitudinal:
 max. = 14.5 g, 4th stage
 13.8 g, 3rd stage (estimated)
 Lateral: none recorded, but see
 “Vibration”
 Altitude.....At end of boosted flight: 1,670,000 ft
 Dynamic Pressure...Maximum ≈ 1800 lb/sq ft at T = 35 sec
 Vibration.....Dominated by characteristic burning
 resonances of 4th stage motor (X-248).

Vibration (Cont.)...Tangential mode of resonance has frequency range of 2350 to 3700 cps and levels to 50 g-rms. Longitudinal mode has fundamental frequency at 580 cps and second harmonic at 1160 cps, with level up to 17.8 g-rms at 580 cps and 8.9 g-rms at 1160 cps—on the longitudinal axis.

Approx. vibration limits from St-9 data (total duration, up to 40 sec):

- Longitudinal:
- 580 cps, to 17.8 g-rms
- 1160 cps, to 8.9 g-rms
- 2350-3700 cps, to 29.4 g-rms

- Lateral:
- 580 cps, to 5.9 g-rms
- 1160 cps, to 2.5 g-rms

(Further vibration data on X-248 motor and Scout rocket are available in NASA-GSFC Test and Evaluation memos 621-36 (St-9) and 621-4 (St-7); see Bibliography.)

Acoustics.....Powerplant noise at launch T=0 approx. flat at 140 db/octave band from 40 to 2000 cps, reduced by 15 to 20 db after 2 sec.

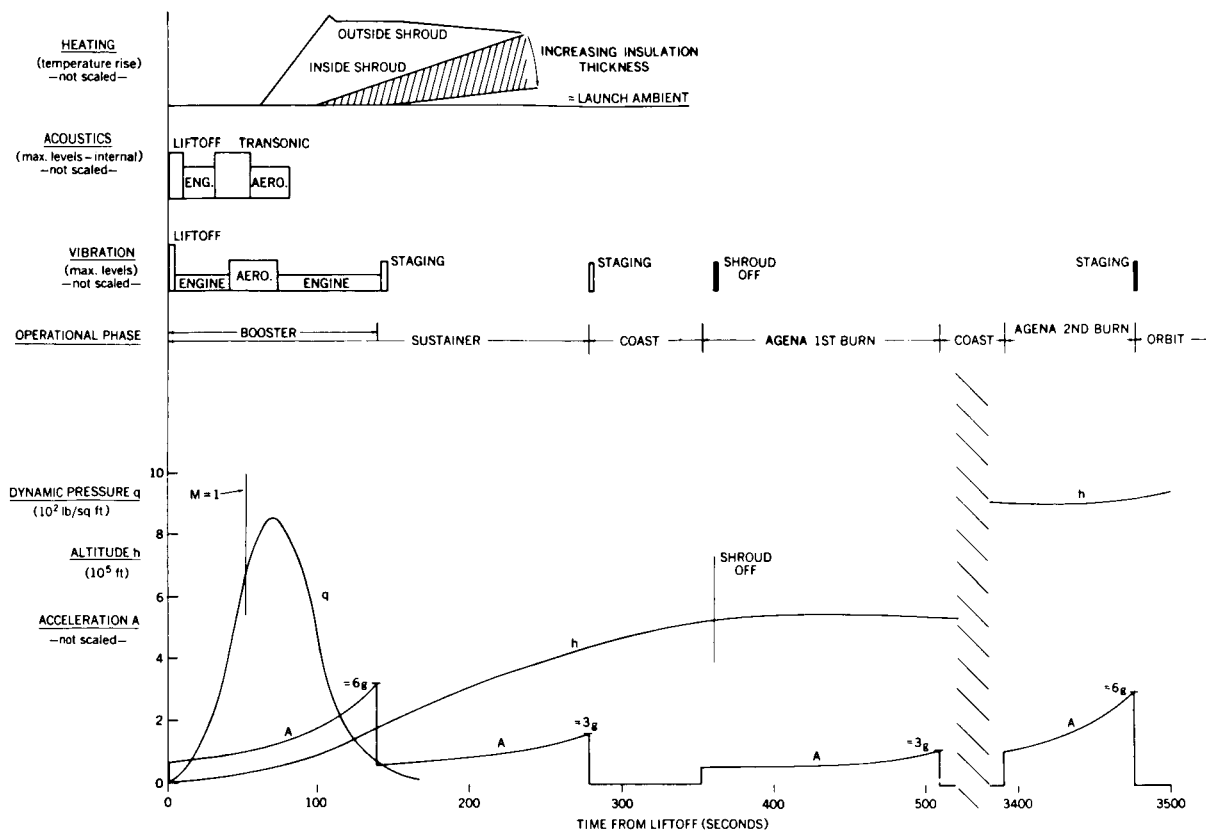


FIGURE 12.—Representative launch profile for Atlas-Agena B vehicle (EGO profile; payload, 1000 lb.).

Acoustics (Cont.)— Aerodynamic noise peaks at maximum dynamic pressure q_{max} ($T=35$ sec) at approx. 140 db/octave band from 800 to 8000 cps (these are external levels).

Heating— Outside heat shield at 2 ft behind nose: Temperature rises from ambient to about 600°F at $T=50$ sec (probe trajectory) and drops off slowly to 500°F at $T=100$ sec.

Inside heat shield at same point: Temperature rises from ambient to about 100°F at 100 sec.

Heat shield ejected before 3rd stage ignition ($T=100$ sec): Payload temperature on St-7 gave erratic readings, but steady-state maximum for $T=200$ sec to $T=600$ sec was 120°F.

Spin— Approx. 160 rpm, initiated just before 4th stage ignition and continuing after final burnout.

Pad environment prior to $T=0$ — Air conditioning required for 4th stage motor.

Typical payloads (in lb) for Scout (NASA only); capability—150 lb to 300-mile orbit and 250 lb for ad-

vanced versions:

Ariel II (S-52; UK-2)	165
Polar Ionosphere Beacon (S-66)	70
Fixed Frequency Topside Sounder (S-48)	110
Micrometeoroid Satellite (S-55)	135

Delta (Representative Profile)

Delta— Three-stage; 1st and 2nd, liquid propellant; and 3rd, solid propellant

1st stage is Thor engine, essentially same as in Thor-Agena B vehicle

3rd stage is X-248, same as 4th stage of Scout

Acceleration— Longitudinal: max. ≈ 18.0 g, 3rd stage ≈ 12.5 g, 1st stage
Lateral: none recorded, but see "Vibration"

Altitude— At end of powered flight, $T \approx 700$ sec: 970,000 ft

Dynamic Pressure— Maximum ≈ 930 lb/sq ft at $T=66$ sec

Vibration— As with Scout, dominated by X-248 burning resonances at 580 cps, and 3700 cps in first half of burning. (See vibration data for Scout; also

Vibration (Cont.)...NASA-GSFC Test and Evaluation memo 621-37 on Delta-9 (S51/UK-1) test—see Bibliography)

Other vibrations: aerodynamic inputs near Mach 1 and q_{max} during first stage, and a possible longitudinal mode near end of 1st stage burning, T = 130 to 160 sec, up to ± 2.8 g-peak. Other tests (AVT-1) on NASA vehicles using same 1st stage give up to 7 g-rms longitudinal near Mach 1 and 3 g-rms lateral at the same time.

Approx. maximum limits, based on these limited data:

Stage	Frequency (cps)	Maximum Acceleration (g-rms)
Longitudinal		
1st Stage (Mach 1, q_{max})	400-2100	7
3rd Stage (X-248 resonance)	580 1160 2350-3700	18 9 30
Lateral		
1st Stage (Mach 1, q_{max})	400-2100	3
3rd Stage	580 1160 2350-3700	6 2.5 52

(See data sheets for Scout and Thor-Agena B)

Acoustics.....Maximum noise levels at liftoff T=0 to T=10 sec, also at possible shroud coincidence-resonance at Mach 0.5. Other high noise levels near Mach 1 and possibly at q_{max} . Estimated overall noise levels at liftoff up to 145 db external, possibly up to 8 db higher at Mach 0.5; noise decays rapidly after q_{max} (T=66 sec).

N. B.
Changes with
varying shroud
configuration

Noise spectra peak at 150 to max 300 cps external and about one octave higher, internal.

Heating.....Inside shroud (only available data): Temperature at junction of nose cone and cylindrical section rises from ambient to 450°-500°F at about T=190 sec, and slowly decreases (data for Ariel I shroud and trajectory). This is for fiber glass shroud (0.10 in.) without insulation in cylindrical section.

Spin.....Approx. 150 rpm, initiated before 3rd stage ignition and continuing after final burnout.

Pad environment prior to T=0.....Air conditioning required for 3rd stage motor.

Typical payloads (lb) for Delta (NASA only); capability—500 lb to 300-mile orbit:

Tiros	285
Orbiting Solar Observatory (S-17)	440
Relay (A-15)	170
Syncom	70
Telstar	170
Interplanetary Monitoring Probe (S-74)	125
Atmospheric Structure Satellite (S-6)	375
Ariel I (S-51; UK-1)	150
Energetic Particles Satellite (S-3a; S-3b)	100
Echo	200

Thor-Agena B (Representative Profile)

Thor-Agena B.....Two-stage, liquid propellant
1st stage, essentially same as 1st stage of Delta vehicle

2nd stage, essentially same as 2nd stage of Atlas-Agena B vehicle

Acceleration.....Longitudinal: max. ≈ 7.5 g, 1st stage ≈ 6.0 g, 2nd stage

Lateral: none feroced

Altitude.....At end of boosted flight, T=3160 sec: $\approx 851,000$ ft

Dynamic Pressure...Maximum ≈ 900 lb/sq ft at T=66 sec

Vibration.....Highest vibration input (low and high frequency) at liftoff, other low frequency vibrations at staging; high frequency vibration peak at Mach 1 to q_{max} . Possible low frequency longitudinal mode, resulting from engine mechanical system, existing for about 20 sec prior to 1st stage

Frequency, Sinusoidal (cps)	Peak (g)
Longitudinal	
8-16	± 1.5
16-22	± 2.8
22-100	± 1.5
100-250	± 2.2
250-400	± 3.3
400-2000	± 5.0
Lateral	
5-100	± 1.0
100-400	± 1.5
400-2000	± 5.0

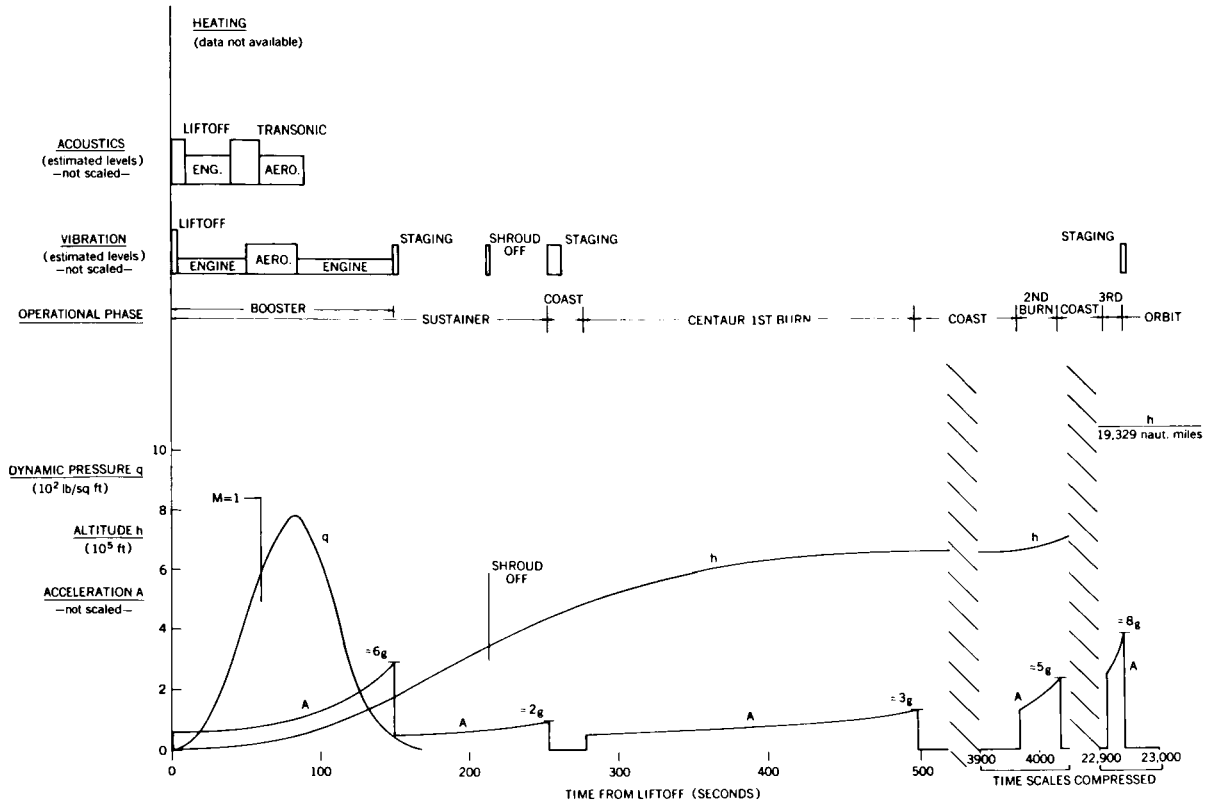


FIGURE 13.—Representative launch profile for Centaur vehicle (24-hr orbit profile; payload, 850 lb.).

Vibration (Cont.)—(main engine) cutoff. This mode may give up to ± 2.8 g-peak at 16 to 22 cps. Approx. limits from recent test data given on page 793.

Acoustics—(Approx. same as given for Delta, but more reliable here.)

Maximum noise levels at liftoff, $T=0$ to $T=10$ sec, also at possible shroud coincidence-resonance at Mach 0.5. Other high noise levels near Mach 1 and possibly at q_{\max} .

Estimated overall noise levels at liftoff up to 145 db external and 130 db internal, possibly up to 8 db higher at Mach 0.5; noise decays rapidly after q_{\max} ($T=66$ sec).

Noise spectra peak at 150 to 300 cps external and about one octave higher, internal.

Heating—External temperature at junction of nose cone and cylindrical section rises ambient up to $T=60$ sec, to about 300°F at $T=110$ sec, then decreases slowly. Temperature inside shroud at same point rises slowly from ambient, to 100° to 230°F , depending on the thickness of the insulation.

Heating (Cont.)—The shroud is ejected at $T=200$ sec, carrying most of the heat input with it.

Spin—Not used during powered flight.

Pad environment

prior to $T=0$ —Probably air-conditioned on pad, but may be unprotected at times.

Typical Thor-Agena B payloads (NASA only); capability—1600 lb to 300-mile orbit:

POGO (S-50)	1000 lb, 4 ft diam. by 10 ft long
Echo-2	650 lb
Nimbus	650 lb, 5 ft diam. by 10 ft long
Alouette (S-27; Canadian)	300 lb
POGO in shroud	1400 lb, 5.5 ft diam. by 18 ft long
Nimbus in shroud	1100 lb, 5.5 ft diam. by 16 ft long

Atlas-Agena B (Representative Profile)

Atlas-Agena B Two-stage, liquid propellant
1st stage is essentially same as in Centaur vehicle.

2nd stage is essentially same as in Thor-Agena B vehicle.

Acceleration—Longitudinal: max. ≈ 6.5 g, 1st stage
 ≈ 6.0 g, 2nd stage

Lateral: none recorded

Altitude—At end of boosted flight, $T=3475$ sec:
 $\approx 930,000$ ft

Dynamic pressure	Maximum ≈ 870 lb/sq ft at $T = 70$ sec.	2nd stage—high energy engine, same as upper stage engines of Saturn C-1 vehicle under development
Vibration	Highest vibration input at liftoff (high and low frequency excitation), low frequency vibrations at staging; high frequency vibrations in transonic- q_{max} regime. Approx. spectrum: Longitudinal: 10-250 cps, ± 1.5 g (sinusoidal) 250-400 cps, ± 2.5 g (sinusoidal) 400-3000 cps, ± 5.0 g (sinusoidal) Lateral: 2.5-250 cps, ± 1.0 g (sinusoidal) 250-400 cps, ± 2.0 g (sinusoidal) 400-3000 cps, ± 5.0 g (sinusoidal)	Acceleration-----Longitudinal: max. ≈ 8.0 g, 2nd stage ≈ 6.0 g, 1st stage Lateral: none recorded
Acoustics	Maximum noise levels at liftoff; and, in transonic range, other high levels may exist at q_{max} . External noise level near nose cone—cylindrical-section junction, about 128 db overall at liftoff and again at about Mach 0.8 ($T = 45$ sec), peaking between 600 and 1200 cps.	Altitude-----Final boost into 24-hr orbit, 19,329 nautical miles at $T \approx 22,900$ sec (6 hr, 22 min)
Heating	External temperature near junction of nose cone and cylindrical section rises from ambient (up to $T = 60$ sec) up to about 350°F at $T = 105$ sec, then decreases very slowly. Temperature inside shroud at same point rises slowly from ambient to $100^\circ\text{--}250^\circ\text{F}$, depending on the thickness of insulation. The shroud is ejected near first Agena ignition, carrying most of the heat input away. External temperatures are very much a function of the location of the measurement point. For example, the external temperature at the tip of the nose cone rises to near 1600°F at about $T = 160$ sec, then decreases to 800°F by the time of shroud ejection.	Dynamic pressure--Maximum ≈ 800 lb/sq ft at $T = 84$ sec Vibration-----Data not available (see Atlas-Agena B) 1st stage vibration extrapolated from Atlas-Agena B data sheet and probably quite similar. Vibration characteristics of 2nd stage engine are unknown.
Spin	Not used during powered flight.	Acoustics-----Data not available; external noise (see Atlas-Agena B) levels probably follow those for Atlas-Agena B. Noise attenuation of Centaur shroud is unknown at time of writing.
Pad environment	prior to $T = 0$ -----Probably air-conditioned on pad, but may be unprotected at times. Typical Atlas-Agena B payloads (NASA only); capability—5000 lb to 300-mile orbit: EGO (S-49) 1000 lb, 4 ft diam. by 10 ft long OAO (S-18) 3600 lb, 8 ft diam. by 12 ft long Ranger 700 lb (GSFC experiment) Mariner A 500 lb (GSFC experiment) Syncom (Advanced) 500 lb OAO in shroud 4700 lb, 10 ft diam. by 26 ft long EGO in shroud 1400 lb, 5.5 ft diam. by 18 ft long	Heating-----Data not available. It is interesting to note that both heating and cooling may be a problem. For example, the S-64 spacecraft for 24-hr orbit launch on Centaur (piggyback load) is limited to a maximum temperature exposure of 250°F and a minimum temperature of -140°F (solar paddles). Because of the long coast time in the launch vehicle shadow (spacecraft end away from sun), it must be protected against cooling as well as heating. Thus, the S-64 has two shrouds; one is jettisoned during sustainer burning, the other at spacecraft separation from 2nd stage.
Pad environment	prior to $T = 0$ -----Unknown	Spin-----Not used during powered flight.
		Pad environment
		prior to $T = 0$ -----Unknown
		Typical Centaur payloads (NASA only); capability—8500 lb to 300-mile orbit:
		Radiation Measurements 112 lb; piggyback load total
		Payload (S-64) 450 lb, 5 ft diam. by 10 ft long
		Surveyor 2000 lb (GSFC experiment)
		Mariner B 1300 lb (GSFC experiment)
		S-64 in shroud ≈ 1000 lb total, 10 ft by 18 ft long

Centaur (Representative Profile)

Centaur	Two-stage, liquid propellant
	1st stage, essentially same as 1st stage of Atlas-Agena B

DISCUSSION OF LAUNCH PROFILE PARAMETERS

The launch environment profiles given in Figures 1 through 13, with accompanying data, form the basis for an envelope of launch phase environmental parameters. This envelope is, in effect, the overall combination of flight level environmental inputs imposed on the payloads of

sounding rockets and unmanned spacecraft launch vehicles.

Changes in the levels of environmental parameters due to new developments in launch vehicle technology also should be considered. A review of this problem by the authors shows that the primary effect of anticipated developments in the period to 1970 will be the extension of acceleration levels for large spacecraft (those over 1000 lb.) This is included in the discussion of acceleration levels below. Notations on possible changes in levels of other inputs are also included.

The summarized levels of environmental parameters given here are classified primarily by spacecraft weight. Combinations of parameters that exist simultaneously during the launch phase are indicated. These are the most severe combinations for likely NASA sounding rocket and spacecraft vehicle configurations and missions,

although other combinations are theoretically possible.

The following parameters have been considered in the discussion:

1. Acceleration—longitudinal
2. Acceleration—lateral
3. Pressure (altitude)
4. Vibration
5. Acoustics
6. Heating
7. Spin

Acceleration—Longitudinal

1. Accelerations range from 60 g for payloads of 50 lb to 15 g for payloads of 300 lb for sounding rockets, and from 36 g for payloads of 50 lb to 6.5 g for payloads of 2000 to 4000 lb for spacecraft.

The envelope of maximum acceleration imposed on unmanned spacecraft during the launch phase

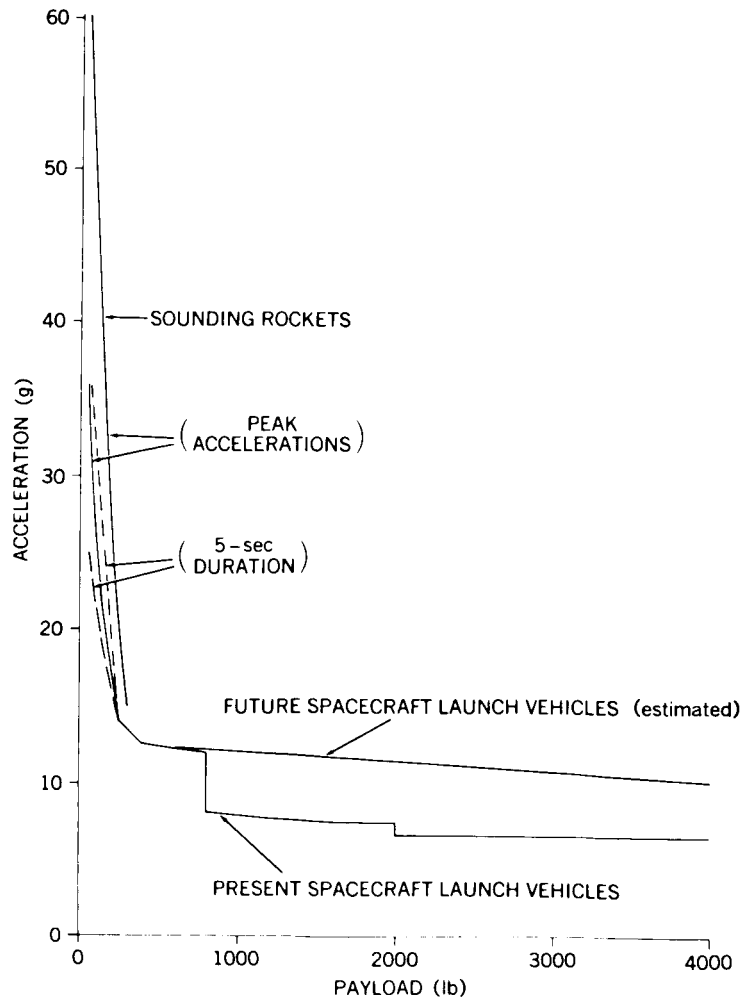


FIGURE 14.—Envelope of spacecraft and sounding rocket maximum acceleration during launch phase.

is given in Figure 14. This figure gives maximum longitudinal acceleration as a function of payload weight. The lower solid curve gives accelerations imposed by present spacecraft launch vehicles including Scout, Delta, Thor-Agena, Atlas-Agena, and Centaur. Programmed improvements in individual stage propulsion are included. The near-vertical solid curve gives accelerations imposed on sounding rocket payloads. The estimated accelerations for future NASA launch vehicles are given by the upper solid curve. These are based on a brief review of development programs for large boost vehicles including Saturn, Nova, Titan II, and Titan III and of anticipated developments in solid and liquid rocket engines.

The high acceleration peaks for the payloads less than 300 lb occur only momentarily at burnout of the boost rocket stages. The dotted curves in Figure 14, showing accelerations of 5-sec duration, give an indication of the sharpness of these peaks and of the problem of simulating acceleration and combined environmental inputs for this range of payloads.

Accelerations for spacecraft launch vehicles range from 36 g for payloads of 50 lb to 12 g for payloads of 400 to 800 lb, 8 g for payloads of 800 to 2000 lb, and 6.5 g for payloads of 2000 to 4000 lb.

Estimated accelerations for future launch vehicles range from about 12 g for payloads of 1000 lb to 10 g for payloads of 4000 lb.

2. The anticipated levels of acceleration imposed by future launch vehicles will not be significantly higher than those at present except that the payload range of 1000 to 4000 lb may be subjected to accelerations on the order of 10 to 12 g.

Estimation of maximum launch acceleration levels that might be imposed on unmanned spacecraft by future launch vehicles was accomplished by a brief review of published material on vehicle and motor development program.

The trend in large solid rockets is toward application of built-up boosters in the liftoff stage, where size and weight are important. Liquid rockets such as the Agena series and the newer hydrogen/oxygen engines (e.g., Centaur) will continue to provide the upper stage propulsion for vehicles such as Titan III and Saturn C-1.

In general, the emphasis is toward large pay-

loads and low accelerations suitable for the manned space flight missions. It is true that high accelerations can be imposed on small payloads attached to large boosters having high thrust/weight ratios. However, the normal vehicle stage weights and thrust/weight ratios limit these accelerations for practical vehicle/payload combinations.

Acceleration—Lateral

Data on lateral accelerations during the launch phase have been conspicuously absent during this study.

Flight level lateral accelerations are usually much lower than those anticipated during transportation and handling of the payload prior to launch. For flight programmed pitch and yaw motions, the lateral acceleration loads are normally less than 0.2 g although, for one of the vehicles considered, these loads reached 0.4 g. These are all small compared with the maximum lateral accelerations due to handling, which are 1.5 to 2.0 g.

There appears to be no requirement for testing under combined longitudinal and lateral acceleration loadings. Payload test specifications may include provisions against conduct of such tests, which are considered to be unrealistic environments.

Pressure (Altitude)

The achievement of high altitude (low pressure) during the launch phase is associated with all acceleration levels and payload weights.

The following rates of pressure change apply:

Large launch vehicles: Delta and up, to 10^{-2} mm Hg in 150 to 180 sec

Small launch vehicles: Scout and sounding rockets, to 10^{-2} mm Hg in 60 to 100 sec

Maximum altitudes (pressure levels) reached during the launch phase are:

Large satellite vehicle: 10^{-7} mm Hg (except 24-hr orbit, 10^{-9} mm)

Probes: 10^{-8} mm Hg

Sounding rockets: 1 mm to 10^{-8} mm Hg

Most orbiting spacecraft exceed altitudes corresponding to 10^{-5} ; mm Hg.

Future levels are anticipated to be the same as present levels.

Vibration

Vibration inputs up to the maximum levels measured can be combined with the highest acceleration loads.

Vibration inputs from rocket combustion, mechanical, and aerodynamic sources can occur throughout all of the flight launch phase. Lateral and longitudinal vibrations are very often of the same order of magnitude, occurring simultaneously.

It is much more difficult to define an envelope of vibration levels. The following is a rough summary of vibration flight levels.

1. Payloads of approx. 1000 lb

Longitudinal:

- 10-250 cps, ± 2.5 g (peak sinusoidal)
- 250-400 cps, ± 3.3 g (peak sinusoidal)
- 400-2000 cps, ± 5.0 g (peak sinusoidal)

Lateral:

- 10-250 cps, ± 1.5 g (peak sinusoidal)
- 250-400 cps, ± 2.0 g (peak sinusoidal)
- 400-2000 cps, ± 5.0 g (peak sinusoidal)

Insufficient data exist for determining random acceleration levels.

2. Payloads up to 500 to 600 lb

In addition to the sinusoidal accelerations given under no. 1 above, it is necessary to consider the large accelerations that may arise from rocket motor resonance burning—specifically the X-248 motor, which is the final stage for Delta, Scout, Javelin, and Journeyman vehicles. This motor has burning resonances at approximately 580 and 1160 cps, and also a range between 2350 and 3700 cps. While these burning resonances are of only a few seconds duration, magnitudes of vibration are very high; for example, the following were measured on a Scout test (150-lb payload):

Longitudinal:

- 580 cps, to 18 g-rms
- 1160 cps, to 9 g-rms
- 2350-3700 cps, to 29 g-rms

Lateral:

- 580 cps, to 6 g-rms
- 1160 cps, to 3 g-rms
- 2350-3700 cps, to 52 g-rms

The Bibliography given in this report contains a number of references related to these engine burning resonances, on Scout and Delta launch vehicles and Javelin sounding rockets.

Vibration levels imposed by future launch vehicles are expected to be of the same order of magnitude as the present ones. It is possible that rocket motors having burning resonances similar to that of the X-248 (perhaps at different frequency ranges) may be used.

Acoustics

Acoustic excitations are at their highest level during the early phases of launch, with primary peaks at liftoff and in the Mach 1 to q_{max} regime.

For all vehicles except sounding rockets, the maximum acoustic inputs occur at acceleration levels less than about 6 g and at altitudes corresponding to pressure levels from atmospheric to about 100 mm Hg. For sounding rockets (payloads up to 300 lb) large acoustic inputs may be associated with accelerations up to 30 g.

Maximum sound pressure levels of overall acoustic inputs range up to about 155 db (external). Internal sound pressure levels are not very well defined, but mostly are 10 db or more below the maximum external sound pressure levels.

Acoustic inputs for future vehicles may be slightly higher than those for present vehicles. An accurate assessment is not possible at this time.

Heating

Aerodynamic heating during the launch phase is of short duration but can be quite intense on the exterior of the shroud. Most of the heat absorbed by the shroud during launch is carried away when the shroud is jettisoned.

The rates of temperature rise for the exterior of the shroud are generally within the following envelopes:

1. Sounding rockets: nose cone stagnation temperature from ambient to 1600°F in 60 sec; shroud temperature at forward cylindrical section to 800°F in 60 sec.
2. Spacecraft launch vehicles: nose cone stagnation temperature from ambient to 1600°F in 100 sec; shroud temperature at forward cylindrical section to 600°F in 100 sec.

Peak internal surface temperatures of the shroud are much lower than peak external

temperatures and are reached more slowly. For spacecraft launch vehicles, the maximum internal temperature in the forward cylindrical section is about 300°F, reached in 200 sec. After shroud ejection the payload is subjected to direct solar radiation and free molecular heating. Temperature change on the payload after shroud ejection is dependent on payload material, surface finish and color, and vehicle trajectory.

Heating inputs for future launch vehicles are not expected to exceed present levels.

Spin

Many payloads are subjected to spin during missions, but only a few are spun during the launch phase. Spin during the launch phase is often associated with high acceleration and vibration levels.

Most sounding rockets are spin-stabilized, with spin rates up to 300 to 600 rpm for payloads up to 300 lb. The spin rates may be associated with acceleration levels up to 30 g as well as high vibration, acoustic, and heating inputs.

Payloads on Scout and Delta launch vehicles are also subjected to spin, during last stage burning (X-248 motor). Spin rates are about 150 to 180 rpm for payloads up to 500 to 600 lb. These spin rates may be associated with maximum acceleration levels for the payload weight and with high vibration levels. Acoustic and heating inputs are not present.

There is no information on possible spin requirements that future launch vehicles may impose on larger payloads.

APPLICATION TO LAUNCH PHASE SIMULATOR

The survey of launch environment profiles summarized in this report can be used to assist in the development of design and operational requirements for the Launch Phase Simulator. The magnitude of each environmental parameter and the extent to which the parameters are combined can be determined from the profiles. Simulation of all parameters in "real time" corresponding to the launch phase of flight missions may not be possible in the Launch Phase Simulator. However, these data provide an overall envelope of the design and operation criteria for this new environmental test facility.

BIBLIOGRAPHY

The launch profile data in this report have been assembled from a wide variety of sources. While it would be very difficult to list references for each item of the profiles, the following have been primary data sources.

Sounding Rockets

Most of the profile data for sounding rockets is from the Sounding Rocket Handbook series prepared by Vought Astronautics, a Division of Chance Vought Corporation, for NASA under contract NAS 1-1013. Altogether, eighteen vehicles are covered in the Handbook series; eight of these were chosen for inclusion in this study. The data are fairly complete, except for vibration inputs.

Additional data on vibration inputs for sounding rockets are given in the following reports:

- OLESON, M. W., "Report on Acceleration and Vibration Data from Javelin (8.02) Vehicle," Naval Research Laboratory Memo Report 1074, July 1960.
- NAGY, J. A., "Vibration Experiments of Aerobee, NASA 4.20 and NASA 4.68," NASA-GSFC Test and Evaluation (T&E) Division Memo Report 621-35, May 18, 1962.
- ELSEN, W. G., "Vibration Experiments of IRIS, NASA 5.04," NASA-GSFC T&E Memo Report 621-41, May 28, 1962.

Spacecraft

Launch profile data for spacecraft other than sounding rockets are from two primary sources: Trajectory data for launch vehicles (acceleration, altitude, dynamic pressure) are adapted from computer trajectory print-outs contained in reports by the launch contractors, except that Scout data are computed from NASA test results. Data on vibration, acoustics, heating, and spin are mostly from NASA published and internal reports and from technical reports of vehicle contractors.

Among the sources consulted for vibrations and acoustic data are:

- NASA-Langley Scout Vehicle Manual, July 1960.
- POSNER, J., "Considerations Affecting Satellite and Space Probe Research with Emphasis on the Scout as a Launch Vehicle," NASA Technical Report R-97.
- MAYES, W. H., HILTON, D. A., and HARDESTY, C. A., "In-Flight Noise Measurements for Three Project Mercury Vehicles," NASA Technical Note D-997, January 1962.

BANGS, W. F., "Vibration Experiment of Scout ST-7/P-21," NASA-GSFC T&E Memo Report 621-4, November 27, 1961.

TERENIAK, W. B., "Preliminary Analysis of Echo A-12 (AVT-1) In-Flight Vibration Data," NASA-GSFC T&E Memo Report 621-10, February 2, 1962.

TERENIAK, W. B., "Scout ST-9/P21a Vibration Experiment," NASA-GSFC T&E Memo Report 621-36, May 19, 1962.

WILLIAMS, L. A., "Vibration Experiment of Delta 9 (S-51/UK-1)," NASA-GSFC T&E Memo Report 621-37, May 21, 1962.

Sources for heating inputs include:

NASA-Langley Scout Vehicle Manual, July 1960.

BOECKEL, J. H., "In-Flight Temperatures of the Scout ST-7/P-21 Ionospheric Probe," NASA-GSFC T&E Memo Report 320-34-61, January 4, 1962.

Spin data are taken from vehicle specifications and vibration test reports for Scout and Delta.

Other Data Sources

Other data sources include a literature survey of scientific and technical publications, Lockheed Missiles and Space Company internal reports and reports to NASA on Agena shroud heating, Douglas Aircraft Company reports to NASA on Delta shroud heating, review of NASA program planning reports, and inputs from other Booz-Allen Applied Research, Inc. programs in astronautics. These sources were particularly useful in extrapolating launch profile parameters to future launch vehicles. Discussions were held with members of the professional staff at GSFC, including T&E Division Test Coordinators and the Spacecraft System and Project Division Vehicle Managers.

The launch profiles contained in this report represent the best application of all these sources.

THREE AND FOUR COIL SYSTEMS FOR HOMOGENEOUS MAGNETIC FIELDS*

M. E. PITTMAN AND D. L. WAIDELICH

Very homogeneous fields are needed in the magnetic testing of spacecraft. One method of obtaining such fields is to use a number of circular coils on a common axis. The parameters of the best three and four coil systems were obtained by setting to zero as many terms as possible in the equation for the field along the axis of the system. The parameters are presented in the form of tables and curves.

INTRODUCTION

In testing the magnetic properties and instruments of spacecraft, there is a need to cancel the earth's magnetic field and then produce a very homogeneous controlled magnetic field. The working volume should be conveniently accessible and the electrical design and construction should be as simple as possible. Since the working volume required is large, it appears that the use of air-core coils is the only practical solution. Most of the previous work has been on circular coils and these will be considered here, although square coils might have some constructional advantages. The volume of homogeneity for two coils (Helmholtz pair) is so small that prohibitively large coils would be needed for the required uniformity. For four coils, a much larger volume of homogeneity may be obtained for a given size of the coils. A few special solutions have been given for the four coil system but no general solution over the whole range of parameters seems to have been made. It is the purpose of this paper to present such a solution and to indicate the various optimum values. The availability of a general solution will allow the design of a system when various factors such as the size or shape of the object being tested do not allow an optimum value to be used. In preparing for the solution of the four coil system, it was found that the solution of the three coil system was useful. The general solution of the three coil system is also presented here but access

to the volume of uniform field is quite limited because of the position of the center coil.

THEORY

The magnetic field of a single circular coil may be obtained by various methods, such as by the use of a scalar or vector magnetic potential (References 1 and 2). The magnetic field along the axis of the coil is

$$H = \sum_{n=0}^{\infty} a_n z^n, \quad (1)$$

where z is the distance along the axis measured from the origin 0 as shown in Figure 1, $z < b_1$. Also

$$a_n = \frac{NI(1-x^2)}{2b_1^{n+1}} P'_{n+1}(x) \quad (2)$$

where N is the number of turns on the coil, I is the current through the coil, and

$$P'_{n+1}(x) = \frac{dP_{n+1}(x)}{dx}$$

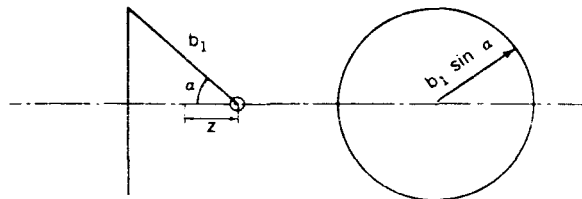


FIGURE 1.—The single circular coil.

*Published as NASA Technical Note D-2095, January 1964.

$P_n(x)$ is the n th order Legendre polynomial and

$$x = \cos \alpha .$$

Two or more of the coils on a common axis may be used to produce a more homogeneous magnetic field than is possible with one coil. This is done by making as many terms zero beyond a_0 in Equation 1 as is possible (References 3 and 4). The

same results may be obtained by use of a potential (References 5 and 6).

For a symmetrical four coil system as shown in Figure 2, the magnetic field along the axis is

$$H = \sum_{n=0, 2, 4, \dots}^{\infty} a_n z^n , \quad (3)$$

where

$$a_n = \frac{N_1 I_1 (1 - x_1^2)}{b_1^{n+1}} P'_{n+1}(x_1) + \frac{N_1 I_2 (1 - x_2^2)}{b_2^{n+1}} P'_{n+1}(x_2) . \quad (4)$$

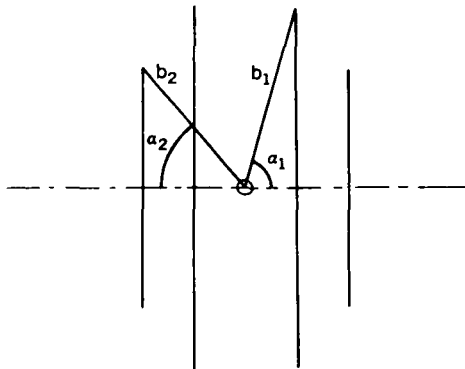


FIGURE 2.—The four coil system.

The terms for the odd values of n in Equation 3 are not considered since they become zero because of symmetry.

THREE COIL RESULTS

The three coil solution is obtained by setting $x_1 = 0$ or $\alpha_1 = 90^\circ$. This has the effect of making the two inner coils of Figure 2 become the one center coil of Figure 3. With three coils it is possible to make a_2 and a_4 in Equation 3 zero and then a_6 may be made a minimum. The solution is given in more detail in appendix A and results are shown in Figures 4 and 5 (A6 and B6 will be defined later). Additional results are presented in Table 1. The range of x_2 in Figures 4 and 5 is from 0.4472 to 0.7651 and the ratio of the radii $b = b_2/b_1$ ranges from infinity to zero as shown in Figure 4. This indicates that when x_2 is close to 0.4472 the diameter of the center coil should be

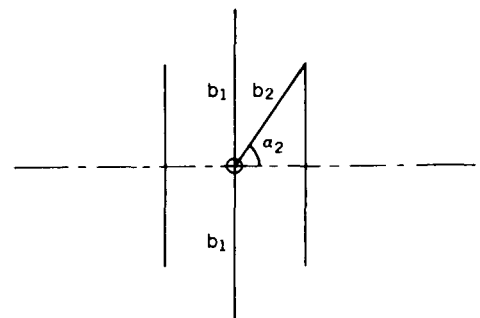


FIGURE 3.—The three coil system.

smaller than that of the two outer coils, whereas when x_2 is close to 0.7651 the diameter of the center coil should be larger than that of the two outer coils. From Table 1 all three coils should have equal diameters at $x_2 = 0.6051$ and at $x_2 = 0.6547$ the three coils should lie on the surface of a sphere. The range of the ratio of ampere-turns, $I = N_2 I_2 / N_1 I_1$ is from infinity to zero as shown in Figure 4. Where $x_1 = 0$ the two inner coils of a four coil system become the center coil of a three coil system. The ampere-turns on the center coil then would be $2N_1 I_1$, and the actual ratio of ampere-turns of an outer coil to the ampere-turns of the center coil is $N_2 I_2 / 2N_1 I_1 = I/2$. When x_2 is close to 0.4472, the number of ampere-turns of the center coil should be smaller than that of the outer coils, and when x_2 is close to 0.7651, the number of ampere-turns of the center coil should be larger than that of the outer coils. From Table 1 the three coils would have an equal number of ampere-turns at $x_2 = 0.6402$.

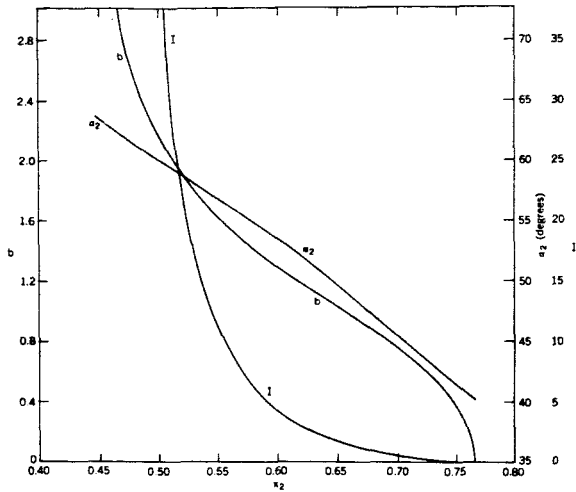


FIGURE 4.—Calculated values of α_2 , I , and b vs. X_2 for the three coil system.

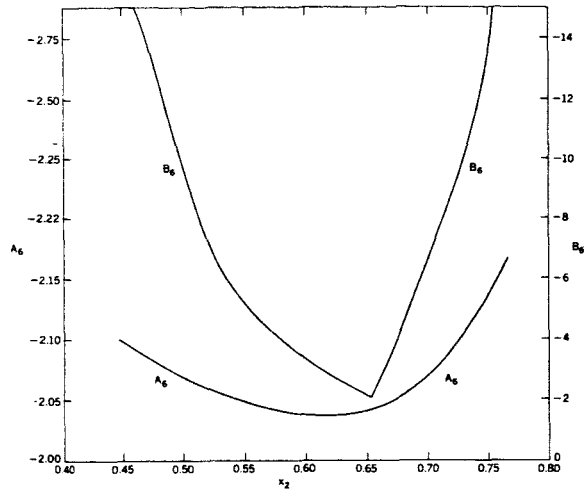


FIGURE 5.—Calculated values of A_6 and B_6 vs. X_2 for the three coil system.

The most homogeneous field would be the one which made the remainder of the series of Equation 3 a minimum; i.e., in

$$H = a_0 \left[1 + \frac{a_6}{a_0} z^6 + \frac{a_8}{a_0} z^8 + \dots \right] \quad (5)$$

the sum of the terms in a_6 , a_8 , and so on should be a minimum. Since the term containing a_6 usually is much larger than the sum of the remaining terms, making the a_6 term a minimum should give a close approximation to the optimum field. This

may be done in a number of ways depending upon which parameter or combinations of parameters are assumed to remain constant. As an example let

$$A_6 = b_1^2 b_2^4 \frac{a_6}{a_0} = \bar{b}^6 \frac{a_6}{a_0} \quad (6)$$

In Equation 6 if the mean radius from the center of the system $\bar{b} = (b_1 b_2^2)^{1/3}$ for the three coils were constant, the a_6 term would have its minimum at $x_2 = 0.6163$ as shown in Figure 5 and Table 1. As a second example let

$$B_6 = b_m^6 \frac{a_6}{a_0} = \begin{cases} b_2^6 \frac{a_6}{a_0} = b^2 A_6 & (b_2 \geq b_1) \\ b_1^6 \frac{a_6}{a_0} = b^{-4} A_6 & (b_1 \geq b_2) \end{cases} \quad (7)$$

If the larger radius from the center of the system to the coils b_m were constant, the a_6 term would have its minimum at $x_2 = 0.6547$ as shown in Figure 5 and Table 1. If the system must be limited to a certain largest volume, probably the second optimum, that of B_6 , would be the better of the two to use. It is possible to define other optima as well but it is believed that these two are the most useful and practical.

FOUR COIL RESULTS

With a four coil system such as that shown in Figure 2, it is possible to make a_2 , a_4 , and a_6 of Equation 3 equal zero and to make a_8 a minimum. This solution is carried out in more detail in Appendix A and the results are shown in Figures 6 and 7 and Table 2. The range of x_2 in Figures 6 and 7 is from 0.44721 to 0.87174 and the corresponding x_1 as shown in Figure 6 decreases from

Table 1
Particular Three Coil Systems.

x_2	a_2 (degrees)	b	l	A_6	B_6	Remarks
0.4472	63.42	∞	∞	-2.100	$-\infty$	End point
0.6051	52.76	1.256	3.763	-2.0369	-3.214	Coils of equal diameter, Barker's solution (Reference 3)
0.6163	51.96	1.197	3.076	-2.0364	-2.925	Minimum $ A_6 $
0.6402	50.19	1.074	2.000	-2.0388	-2.351	Coils of equal ampere-turns
0.6547	49.10	1.000	1.531	-2.043	-2.043	Minimum $ B_6 $, Maxwell's solution (Reference 6), coils on surface of sphere
0.7651	40.08	0.000	0.000	-2.167	$-\infty$	End point

0.20929 to a minimum of 0.20360 and then increases to a maximum of 0.44721. The ratio of the radii $b=b_2/b_1$ ranges from infinity to zero. Thus when x_2 is close to 0.44721, the diameter of the two inner or No. 1 coils should be smaller than that of the two outer or No. 2 coils, but when x_2 is close to 0.87174 the diameter of the No. 1 coils should be larger than that of the No. 2 coils. As

indicated in Table 2 at $x_2=0.68519$ all four coils have the same diameter, whereas at $x_2=0.76505$ the four coils lie on the surface of a sphere. At 0.85363 they lie in the same plane perpendicular to the axis of the system, the No. 1 coils having a diameter 3.76797 times that of the No. 2 coils. From $x_2=0.85363$ to 0.87174, the No. 2 coils are closer to the center of the system than the No. 1

Table 2
Particular Four Coil Systems.

x_1	x_2	b	l	A_8	B_8	Remarks
0.20929	0.44721	∞	∞	-2.35384	$-\infty$	End point
0.20360	0.51961	2.49155	46.00025	-2.34170	-90.24299	Minimum x_1
0.23629	0.68519	1.33407	2.26058	-2.47133	-7.82798	All coils have equal diameters, Barker's solution (Reference 3)
0.26786	0.74207	1.09795	1.000	-2.23448	-3.24721	All coils have equal ampere-turns Braunebek's solution (References 6 and 7)
0.27235	0.74842	1.07127	0.90406	-2.22196	-2.92648	Optimum using both A_8 and A_{10}
0.27505	0.75208	1.05576	0.85165	-2.21988	-2.75803	Optimum using A_8 only
0.28523	0.76505	1.000	0.68211	-2.25510	-2.25510	All coils on surface of a sphere. Optimum using B_8 . McKeehan's solution (References 5 and 6)
0.39864	0.85363	0.46699	0.024569	-2.34225	-49.24825	Both No. 1 and No. 2 coils lie in the same plane perpendicular to the axis of the system
0.44721	0.87174	0.000	0.000	-2.35384	$-\infty$	End point

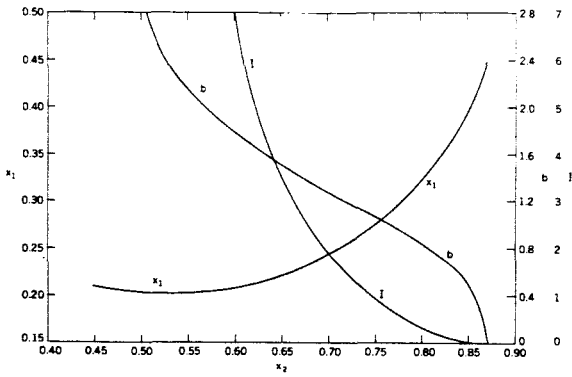


FIGURE 6.—Calculated values of x_1 , I , and b vs. x_2 for the four coil system.

The most uniform field is the one which makes the remainder of the series of Equation 3 a minimum; i.e., in

$$H = a_0 \left[1 + \frac{a_8}{a_0} z^8 + \frac{a_{10}}{a_0} z^{10} + \dots \right] \quad (8)$$

the sum of the terms involving a_8 , a_{10} , etc. should be made a minimum. A good approximation to this minimum should be that which makes the a_8 term a minimum. This may be done in a number of ways, one of which is to make the mean radius from the center of the system, $\bar{b} = (b_1 b_2)^{1/2}$, constant and define

$$A_8 = b_1^4 b_2^4 \frac{a_8}{a_0} = \bar{b}^8 \frac{a_8}{a_0} \quad (9)$$

Figure 7 shows a curve of A_8 plotted against x_2 and Table 2 gives some values of A_8 including those at the end points. The minimum value of A_8 as given in Table 2 and shown in Figure 7 is -2.21988 , which occurs at $x_2 = 0.75208$ and $x_1 = 0.27505$.

Another way of making the a_8 term a minimum is to put

$$B_8 = b_m^8 \frac{a_8}{a_0} = \begin{cases} b_2^8 \frac{a_8}{a_0} & (b_2 \geq b_1) \\ b_1^8 \frac{a_8}{a_0} & (b_1 \geq b_2) \end{cases} \quad (10)$$

where b_m is the larger radius from the center of the system to the coils. A curve of B_8 against x_2 is shown in Figure 7 and some values of B_8 are given in Table 2. One of the values in Table 2 is the minimum value of $B_8 = -2.25510$, which occurs at $x_2 = 0.76505$ and $x_1 = 0.28523$. If more than one term is considered in the series of Equation 8, the minimum depends upon the magnitude of z . For example, consider the a_8 and a_{10} terms. They may be written as

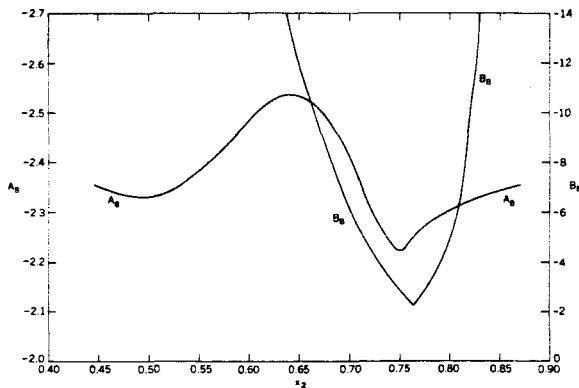


FIGURE 7.—Calculated values of A_8 and B_8 vs. x_2 for the four coil system.

coils. It is interesting to notice that if the No. 1 and No. 2 coils are interchanged the two end points of Table 2 become identical. The range of the ratio of ampere-turns, $I = N_2 I_2 / N_1 I_1$, goes from infinity to zero as shown in Figure 6. When x_2 is close to 0.44721 the number of ampere-turns of the No. 1 (inner) coils is smaller than that of the No. 2 (outer) coils, and for x_2 close to 0.87174 the number of ampere-turns of the No. 1 coils is larger. All coils have equal ampere-turns at $x_2 = 0.74207$.

$$\begin{aligned} \bar{b}^8 \frac{a_8}{a_0} \left(\frac{z}{\bar{b}}\right)^8 + \bar{b}^{10} \frac{a_{10}}{a_0} \left(\frac{z}{\bar{b}}\right)^{10} &= A_8 \left(\frac{z}{\bar{b}}\right)^8 + A_{10} \left(\frac{z}{\bar{b}}\right)^{10} \\ &= c \left(\frac{z}{\bar{b}}\right)^8 \end{aligned} \quad (11)$$

where

$$A_{10} = \bar{b}^{10} \frac{a_{10}}{a_0} = b_1^5 b_2^5 \frac{a_{10}}{a_0},$$

$$C = A_8 + A_{10} \left(\frac{z}{\bar{b}} \right)^2.$$

The minimum value of C is the same as the minimum value of A_8 when $z=0$ but when $z/\bar{b} = 0.2167$, the minimum value of C occurs at $x_2 = 0.74842$ and $x_1 = 0.27235$ as given in Table 2.

There are several other well-known four coil solutions which do not have as great a uniformity as the solutions given in Figures 6 and 7 and Table 2 because only two coefficients of Equation 3, a_2 and a_4 , are made zero. Neumann (Reference 6) and Fanselau (References 4 and 8) put x_1 and x_2 at the roots of $P'_5(x) = 0$ to make a_4 zero, chose the ampere-turns of both No. 1 and No. 2 coils to be equal, and found b by putting $a_2 = 0$. Further details are given in Table 3. Fanselau (Reference 10) in another solution made a_4 zero by using the roots of $P'_5(x) = 0$. Both coils on one side of the system were put in the same plane perpendicular to the axis of the system and a_2 was made zero by choosing $I = 28.2897$. The ratio of the diameter of the smaller coil to that of the larger was 0.250495 and $b = 0.372830$. Fanselau also indicated that a similar solution could be found where both sets of coils would have the same radius. Several additional four coil solutions with both a_2 and a_4 zero are given by McKeehan (Reference 6). Scott used four coils, the inner pair having a smaller diameter than the outer pair (Reference 11). His solution had both a_2 and a_4 zero. Franzen (Reference 12) used Garrett's theory (Reference 5) to develop a theory on a finite coil cross-section for a four coil system.

In Table 3 the specifications of several coil systems are presented along with an indication of how large a sphere about the center of the system will have a given homogeneity. Several three and four coil systems are given together with the two coil system of Helmholtz and the presently known six and eight coil systems. These are compared on a basis of an inhomogeneity of 10^{-5} . An example will now be given by using the minimum A_8

solution for the four coil system:

$$\left| A_8 \left(\frac{z}{\bar{b}} \right)^8 \right| = 10^{-5} \quad (12)$$

and

$$\left(\frac{z}{\bar{b}} \right) = 0.21677 \text{ (or 21.677 percent) .}$$

This means that if $\bar{b} = 10$ feet and only the A_8 term is considered, the inhomogeneity of the magnetic field inside a sphere 2.1677 feet in radius is less than or equal to 10^{-5} or 0.001 percent of the magnetic field at the center of the system. If B_8 is used

$$\left| B_8 \left(\frac{z}{b_m} \right)^8 \right| = 10^{-5} \quad (13)$$

and

$$\left(\frac{z}{b_m} \right) = 0.20944 \text{ (or 20.944 percent) .}$$

The number of ampere-turns in the middle coil of the three coil system should be $2N_1I_1$. For the six coil systems, the x 's for the Neumann-Fanselau and the Williams-Cain solutions are the roots of $P'_7(x) = 0$. Also, for the eight coil systems the x 's for the Neumann-McKeehan and the Williams-Cain solutions are the roots $P'_9(x) = 0$. Similar solutions with the roots of $P'_{n+1}(x) = 0$, where n is even, could be found for systems containing ten, twelve, or more coils (Reference 9). In fact Garrett mentions the fact that the solution for a sixteen coil system could be found readily (Reference 5). The Williams-Cain solution always gives the optimum solution for minimum B_{2n} .

CONCLUSIONS

Complete solutions along with tables and curves that should be useful in design work have been presented for the three and four circular-coil systems with zero for the cross-sectional areas of the coils. Comparisons of these results have been made with the two coil (Helmholtz) and the known six and eight coil systems. At the present time the following are needed:

1. A complete study of the six coil problem over the ranges of all of the parameters.
2. An investigation of the square or rectangular coil systems.

Table 3
Specifications of Various Coil Systems.

Source	Ampere (Reference 6)	Helmholtz (Reference 6)	Barker (Reference 3)	Maxwell (Reference 6)	Minimum A_0	
Number of Coils	1	2	3	3	3	
Assumptions	None	None	Equal diameter coils	Coils on surface of sphere	None	
The a 's that are zero	None	a_2	a_2, a_4	a_2, a_4	a_2, a_4	
Coefficient of next term, A_n	-1.5	-1.8	-2.03689	-2.0428571	-2.036426	
x_1	0	0.4472136	0.0	0.0	0.0	
x_2	-	-	0.605108	0.6546537	0.61627	
b	-	-	1.25606	1.00	1.19697	
l	-	-	3.76323	1.53125	3.075846	
For an inhomogeneity of 10^{-3} (percent)	z/b	0.26	4.86	13.0368	13.031	13.0373
	z/b_n	0.26	4.86	12.083	13.031	12.279

Source	Neumann-Fanslau (References 6 and 8)	Braunbek (Reference 7)	Barker (Reference 3)	McKeehan (References 5 and 6)	Minimum A_0	
Number of coils	4	4	4	4	4	
Assumptions	Coils have equal ampere-turns	Coils have equal ampere-turns	Coils have equal diameters	Coils on surface of sphere	None	
The a 's that are zero	a_2, a_4	a_2, a_4, a_6	a_2, a_4, a_6	a_2, a_4, a_6	a_2, a_4, a_6	
Coefficient of next term, A_n	-1.289	-2.23448	-2.47133	-2.255102	-2.2219679	
x_1	0.285232	0.26786	0.23629	0.2852315	0.2723547	
x_2	0.765055	0.74207	0.68519	0.7650553	0.7484183	
b	1.136009	1.09795	1.33407	1.00	1.0712777	
l	1.00	1.00	2.26058	0.6821109	0.9040608	
For an inhomogeneity of 10^{-3} (percent)	z/b	14.07	21.672	21.480	21.607	21.677
	z/b_n	13.20	20.682	18.583	21.607	20.944

Source	Neumann-Fanslau (References 6 and 8)	Braunbek-McKeehan (References 6 and 7)	Williams-Cain (Reference 9)	Neumann-McKeehan (Reference 6)	Williams-Cain (Reference 9)	
Number of coils	6	6	6	8	8	
Assumptions	Coils have equal ampere-turns	Coils have equal ampere-turns	Coils lie on surface of sphere	Coils have equal ampere-turns	Coils lie on surface of sphere	
The a 's that are zero	a_2, a_4, a_6	a_2 to a_{10}	a_2 to a_{10}	a_2 to a_8	a_2 to a_{10}	
x_1	0.20929922	0.190655	0.20929922	0.1652754	0.1652754	
x_2	0.59170018	0.550274	0.59170018	0.4779250	0.4779250	
x_3	0.87174003	0.843307	0.87174003	0.7387739	0.7387739	
x_4	-	-	-	0.9195342	0.9195342	
b_2/b_1	1.071723	1.046147	1.00	1.0222398	1.00	
b_3/b_1	1.242359	1.157907	1.00	1.1827288	1.00	
b_4/b_1	-	-	-	1.2382935	1.00	
$N_2 I_2 / N_1 I_1$	1.00	1.00	0.8270469	1.00	0.891626	
$N_3 I_3 / N_1 I_1$	1.00	1.00	0.5108492	1.00	0.686604	
$N_4 I_4 / N_1 I_1$	-	-	-	1.00	0.406992	
For an inhomogeneity of 10^{-3} (percent)	z/b_n	31	34	36	16	47

3. An analysis of the effect of finite cross-sectional area of the coils.

REFERENCES

1. SMYTHE, W. R., "Static and Dynamic Electricity," New York: McGraw-Hill, 1950, pp. 270-275.
2. BOOKER, H. G., "An Approach to Electrical Science," New York: McGraw-Hill, 1959, pp. 449-472.
3. BARKER, J. R., "New Coil Systems for the Production of Uniform Magnetic Fields," *J. Sci. Instrum.*, 26(8):273-275, August 1949.
4. SAUTER, A., and SAUTER, F., "Die Erzeugung von Möglichst Homogenen Magnetfeldern durch Stromsysteme," *Zeits. Physik*, 122:120-136, 1944.
5. GARRETT, M. W., "Axially Symmetric Systems for Generating and Measuring Magnetic Fields. Part 1," *J. Appl. Phys.*, 22(9):1091-1107, September 1951.
6. McKEEHAN, L. W., "Combinations of Circular Currents for Producing Uniform Magnetic Fields," *Rev. Sci. Instrum.*, 7(3):150-153, March 1936.
7. BRAUNBEK, W., "Die Erzeugung Weitgehend Homogener Magnetfelder durch Kreisströme," *Zeit Physik*, 88(5/6):399-402, April 3, 1934.
8. FANSELAU, G., "Die Erzeugung Weitgehend Homogener Magnetfelder durch Kreisströme," *Zeits. Physik*, 54:260-269, 1929.
9. WILLIAMS, V. L., and CAIN, J. C., "Homogeneous Magnetic Fields Produced by Circular Current Loops," unpublished memorandum, Goddard Space Flight Center, Greenbelt, Maryland.
10. FRANSELAU, G., "Über die Homogenität des Magnetfeldes bei Symmetrischer Spulenordnung," *Zeits. Geophysik*, 9(4/5):236-237, 1933.
11. SCOTT, G. G., "Compensation of the Earth's Magnetic Field," *Rev. Sci. Instrum.*, 28(4):270-273, April 1957.
12. FRANZEN, W., "Generation of Uniform Magnetic Fields by Means of Air-Core Coils," *Rev. Sci. Instrum.*, 33(9):933-938, September 1961.

APPENDIX A

SOLUTION OF THE COIL SYSTEMS

Three Coil System

Put $x_1=0$ in Equation 4 and let $b=b_2/b_1$ and $I=N_2I_2/N_1I_1$. Then for $n=2$ and $n=4$ in Equation 4

$$-\frac{3}{2} b^3 + I(1 - x_2^2)P'_3(x_2) = 0, \quad (A1)$$

$$\frac{15}{8} b^5 + I(1 - x_2^2)P'_5(x_2) = 0, \quad (A2)$$

where

$$P'_3(x_2) = \frac{3}{2}(5x_2^2 - 1)$$

and

$$P'_5(x_2) = \frac{15}{8}(21x_2^4 - 14x_2^2 + 1).$$

From Equations A1 and A2

$$\frac{I(1 - x_2^2)}{b^3} = \frac{+\frac{3}{2}}{P'_3(x_2)} = \frac{-\frac{15}{8}b^2}{P'_5(x_2)}. \quad (A3)$$

Therefore,

$$b^2 = \frac{1 - 14x_2^2 + 21x_2^4}{1 - 5x_2^2} \quad (A4)$$

$$I = \frac{b^3}{(1 - x_2^2)(5x_2^2 - 1)} \quad (A5)$$

In Equation A4, b^2 must be positive, so that only solutions must have $0.0 \leq x_2 \leq 0.2852$ or $0.4472 \leq x_2 \leq 0.7651$ where 0.4472 is the root of $P'_3(x_2)=0$ and 0.2852 and 0.7651 are the roots of $P'_5(x_2)=0$. The range of values of x_2 from 0.4472 to 0.7651 produces positive values of I and the range from 0.0 to 0.2852 produces negative values of I . A negative value of I means that the current direction in the center coil is reversed from those in the two outer coils.

For the a_6 term of Equation 3 let

$$A_6 = b_1^2 b_2^4 \frac{a_6}{a_0} = b^4 b_1^6 \frac{a_6}{a_0}, \quad (A6)$$

where a_6 and a_0 are given by Equation 2. Then, by the use of Equations 4, A4, and A5

$$A_6 = -\frac{7}{16} \left(\frac{15x_2^6 - 9x_2^4 + 5x_2^2 + 5}{x_2^2 + 1} \right). \quad (A7)$$

The value of x_2 at which Equation A7 has its minimum value is 0.6163. This value of x_2 is in the range for which I is positive. Hence, for the three coil solution at least, making all currents flow in the same direction will produce a more homogeneous field. Another minimum value may be obtained by defining

$$B_6 = \begin{cases} b_2^6 \frac{a_6}{a_0} = b^2 A_6 & (0.4475 \leq x_2 \leq 0.6547), \\ b_1^6 \frac{a_6}{a_0} = \frac{A_6}{b^4} & (0.6547 \leq x_2 \leq 0.7651), \end{cases} \quad (A8)$$

where $b=1.0$ at $x_2=0.6547$. Then the minimum value of B_6 occurs at $x_2=0.6547$.

Four Coil System

For $n=2, 4,$ and 6 from Equation 4, with $b=b_2/b_1$ and $I=N_2I_2/N_1I_1$:

$$b^3 (1 - x_1^2) P'_3(x_1) + I (1 - x_2^2) P'_3(x_2) = 0, \quad (A9)$$

$$b^5 (1 - x_1^2) P'_5(x_1) + I (1 - x_2^2) P'_5(x_2) = 0, \quad (A10)$$

$$b^7 (1 - x_1^2) P'_7(x_1) + I (1 - x_2^2) P'_7(x_2) = 0, \quad (A11)$$

where

$$P'_7(x_1) = \frac{7}{16} (429x_1^6 - 495x_1^4 + 135x_1^2 - 5),$$

and $P'_3(x_2)$ and $P'_5(x_2)$ are given after Equation A2. From Equations A9–A11

$$\frac{I (1 - x_2^2)}{b^3 (1 - x_1^2)} = -\frac{P'_3(x_1)}{P'_3(x_2)} = -\frac{b^2 P'_5(x_1)}{P'_5(x_2)} = -\frac{b^4 P'_7(x_1)}{P'_7(x_2)}. \quad (A12)$$

Then

$$b^2 = \frac{P'_3(x_1) P'_5(x_2)}{P'_3(x_2) P'_5(x_1)} = \frac{P'_5(x_1) P'_7(x_2)}{P'_5(x_2) P'_7(x_1)}. \quad (A13)$$

From this equation:

$$\frac{P'_3(x_1) P'_7(x_1)}{[P'_5(x_1)]^2} = \frac{P'_3(x_2) P'_7(x_2)}{[P'_5(x_2)]^2} = f(x_2) \quad (A14)$$

or

$$P'_3(x_1) P'_7(x_1) - f(x_2) [P'_5(x_1)]^2 = 0. \quad (A15)$$

For a given x_2 , Equation A15 is solved for x_1 and the ratio of the radii b is obtained from Equation A13. The ampere-turn ratio I may be obtained by the use of Equation A12:

$$A_8 = \frac{P'_3(x_1) P'_5(x_2) P'_7(x_2) P'_9(x_1) - P'_3(x_2) P'_5(x_1) P'_7(x_1) P'_9(x_2)}{P'_3(x_2) P'_5(x_2) P'_7(x_1) - P'_3(x_1) P'_5(x_1) P'_7(x_2)}, \quad (A18)$$

where

$$P'_9(x) = \frac{45}{128} (2431x^8 - 4004x^6 + 2002x^4 - 308x^2 + 7).$$

$$I = -\frac{b^3 (1 - x_1^2) P'_3(x_1)}{(1 - x_2^2) P'_3(x_2)}. \quad (A16)$$

The use of a computer allows numerous solutions to be obtained over the range for x_2 from 0.4472 which is a root of $P'_3(x)=0$ to 0.8717 which is a root of $P'_7(x)=0$. The corresponding range for x_1 is from 0.2036 to 0.4472. For these values I is positive. I may be negative when both x_1 and x_2 lie in the range between 0.4472 and 0.5917 where 0.5917 is a root of $P'_7(x)=0$. The values for negative I were not calculated because the more useful values were in the range where I is positive.

For the a_8 term of Equation 3 let

$$A_8 = b_1^4 b_2^4 \frac{a_8}{a_0}, \quad (A17)$$

where a_8 and a_0 are given by Equation 2. Then by the use of Equations 4, A13, and A16

The minimum value of $|A_8|$ is 2.21989 which occurs at $x_1=0.27505$ and $x_2=0.75209$. If the sum of both the a_8 and a_{10} terms is minimized the minimum occurs at $x_1=0.27235$ and $x_2=0.74841$. Again, another minimum may be defined by taking

$$B_8 = \begin{cases} b_2^8 \frac{a_8}{a_0} = b^4 A_8 & (x_2 \leq 0.76505) \\ b_1^8 \frac{a_8}{a_0} = \frac{A_8}{b^4} & (x_2 \geq 0.76505) \end{cases} \quad (A19)$$

The minimum value of B_8 occurs at $x_1=0.28523$ and $x_2=0.76505$. The next term in the series has

$$A_{10} = b_1^5 b_2^5 \frac{a_{10}}{a_0}, \quad (A20)$$

where a_{10} and a_0 are given by Equation 2.

If solutions are needed with greater accuracy than those available in the curves and tables presented, they may be obtained by numerical solution of the equations given in this appendix by using the present solutions as a starting point.

ADAPTATION OF AN MoS₂ "IN SITU" PROCESS FOR LUBRICATING SPACECRAFT MECHANICAL COMPONENTS*

CHARLES E. VEST

This paper presents one approach to overcome some of the difficulties of using MoS₂ as a space lubricant. These difficulties occur in obtaining a tenacious bond of MoS₂ film to the substrate material, and in devising a method of application which is subject to good quality control so that a known film thickness (within micro-inches) can be consistently placed on specimens. The procedure described in this technical note is known as the *in situ* method of applying MoS₂ lubricating film. It is accomplished by electrodeposition of an MoO₃ complex film onto the substrate material and converting this film to MoS₂ by subjecting the coated specimen to H₂S gas at elevated temperature and pressure. This procedure is an adaptation of a patented IBM process for lubricating mild steel components.

Results from this investigation have shown that film thicknesses from approximately 50 to 250 micro-inches can be successfully applied to within ± 50 micro-inches to a number of materials used in spacecraft components. Materials that have been coated successfully are: stainless steels—303, 304, 316, 416, and 440C; aluminum alloys—2024, 6061, and 7075; mild steels; and M-10 tool steel. Test specimens were coated and then subjected to standard basic friction and wear tests in air and vacuum. The results show that (1) the *in situ* film has a coefficient of friction between 0.025 and 0.05, (2) film wear life is slightly better than it is for a dry MoS₂ powder film, much better than it is for inorganic bonded MoS₂ powder film, and not quite as good as it is for organic bonded MoS₂ film, and (3) quality control and reproducibility of the film are very good.

In addition, the *in situ* film has been successfully applied to instrument ball bearing races and retainers and to instrument gears. Test results of these components operated in a vacuum environment substantiate the results of the basic friction and wear tests.

INTRODUCTION

Exposure to space has proven to be a particularly harsh condition for mechanical elements that must operate with a very high degree of reliability. Where relative motion must take place and the replenishment of protective oxide layers happens very slowly or not at all, a vacuum, in particular, poses severe operating problems from the standpoint of friction and wear.

Normal liquid and semi-solid grease lubricants have vapor pressures higher than the ambient pressures in outer space (estimated at 10^{-13} torr interplanetary and 10^{-16} torr interstellar); hence these lubricants evaporate rapidly. In addition to the low pressure problem, space lubricants must not be adversely affected by 10^9 rads of ionization radiation.

A literature survey of the space lubricant field shows that, basically, five approaches have been taken to reduce the problem of friction and wear in space. These involve the use of:

1. Special low-vapor-pressure oils and greases.
2. Metallic films.
3. Ceramics or other extremely hard surfaces.
4. Plastics with low vapor pressures and low coefficients of friction.
5. Laminar solids.

The desirable space lubricant must, in addition to satisfying the vacuum and radiation requirements, meet the following criteria: no outgassing to contaminate optical devices, low coefficient of friction to conserve available power, and low wear rate for extended lifetime.

An analysis of the five basic approaches for these conditions listed above shows that none of them will meet all of the requirements because of

*Published as NASA Technical Note D-2288, May 1964.

(1) high vapor pressure, (2) high coefficient of friction, (3) short wear life, and (4) contamination. The approach that meets the majority of the requirements involves the use of laminar solids, in particular MoS₂. Therefore, this material was chosen as the space lubricant to be investigated.

A literature search revealed that there are a number of methods (References 1 and 2) of bonding MoS₂ onto a substrate. The most widely used methods are inorganic—sodium silicate and graphite, organic—phenolic and epoxies, and MoS₂ powder—burnished or sprayed onto a surface. Each of these has its strong and weak points. In wear life comparison under identical conditions these bonded MoS₂ lubricants rate as follows: an epoxy bond is better than a burnished powder bond which is better than a sodium silicate bond. In vacuum stability, the epoxies outgas more than the other binders. The control of the film thickness in the organic or inorganic bonded methods is ± 100 micro-inches, at best. In some of these films, 300 micro-inches is required to cover the surface completely. It is very difficult to measure the film thickness of burnished MoS₂ powder and this was not attempted.

With the undesirable points as prime considerations—short wear life, little control of film thickness, and vacuum instability—an *in situ* MoS₂ technique was considered and tested as the possible solution to elimination of these undesirable features.

The literature states that natural MoS₂ powders adhere tenaciously to a metal because of the attraction of the sulphur atoms to the substrate. This attraction is not sufficient to produce a long lasting film under test conditions. As one of the goals in this investigation is to assure a long wear life, it was felt that this natural bond could be strengthened by the ionic forces produced in electrodeposition and by the chemically clean surface produced by surface activation treatment. Also, as electrodeposition is a definitely controllable process, it was felt that good control of film thickness could be attained. Furthermore, since no binder material is used, the problem of outgassing and contamination would be alleviated.

The particular process investigated is an adaptation of the IBM-patented MoS₂ *in situ* process (References 3 and 4). The process consists of surface activation, electrodeposition of an MoO₃

complex ion onto the substrate surface, and converting this film to MoS₂ in an atmosphere of H₂S gas at 400 psig and 195°C with an exposure period of 4 to 8 hr.

APPROACH

The approach taken was to examine the processing conditions shown in the IBM patent and, by experimentation, modify these conditions to produce an MoS₂ *in situ* film suitable for spacecraft components.

As MoS₂ has a natural tendency to adhere to a metallic substrate through the attraction of the sulphur atom to the metal, the first step was to obtain a chemically clean surface to assist in assuring or improving this natural tendency. The second step was to investigate the composition and temperature of the bath and electrodeposition time. The third step was to assure the conversion of the electrodeposited film to MoS₂, which was accomplished by investigating the conversion time, temperature, and pressure. This approach required the investigation of surface preparation, composition and temperature of the electrodeposition bath, time and current employed in electrodeposition, and time and pressure required for conversion.

To determine the merits of the various processing modifications, a fourth step or the test step was established which made maximum use of "calibrated" test apparatus. Experts in the lubricant testing field were enlisted to work on this step. Included were tests on the Alpha LFW-1 equipment,* and vacuum tests on the friction and wear test equipment at the NASA Lewis Research Center, Cleveland, Ohio (Reference 5). In addition, actual space components such as bearings, gears, and sliding devices were processed and tested under simulated space conditions at the Goddard Space Flight Center.

As the H₂S atmosphere is corrosive to most materials used in spacecraft construction, the fifth step included a study of the microstructure materials subjected to this atmosphere.

In summary, the approach included the investigation of the following:

1. Surface preparation.
2. Electrodeposition.

*Personal communication with Mr. A. DiSapio, Alpha Molykote Corp., Stamford, Connecticut.

3. Conversion.
4. Testing.
5. Substrate deterioration.

APPARATUS AND PROCEDURE

The apparatus used in the experimental work is shown in Figures 1-3. The cleaning, surface activation, and electrodeposition apparatus consist of an ultrasonic cleaner, acid baths (glassware), a magnetic stirrer and hot plate, an electrodeposition bath with a platinum-plated titanium wire mesh anode, rinse tanks (glassware), and a power supply.

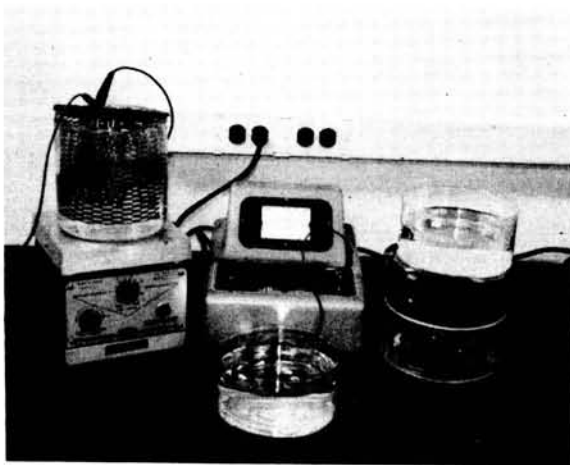


FIGURE 1.—The surface activation and electrodeposition apparatus.

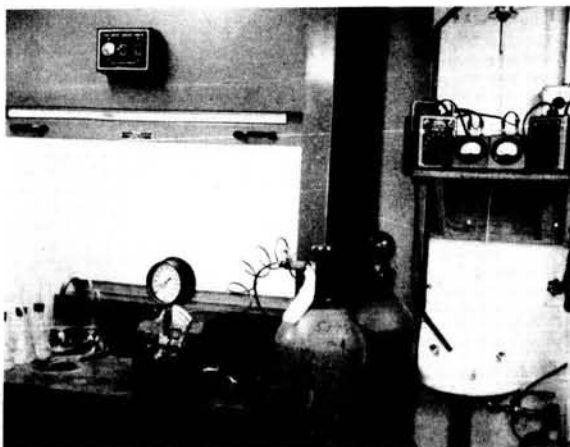


FIGURE 2.—The conversion apparatus—pressure vessel, KOH bottles, H₂S gas cylinder, and heater controls.

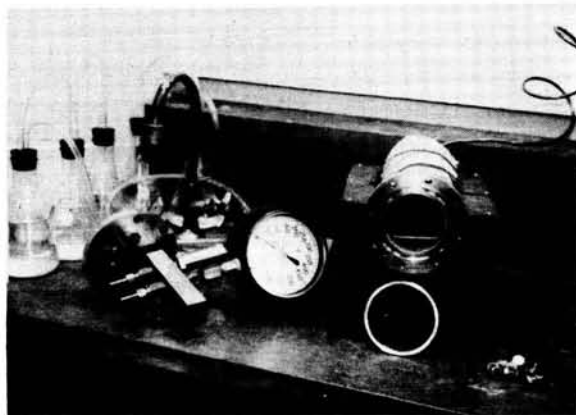


FIGURE 3.—The conversion vessel, 1100 aluminum alloy seal ring, pressure gauge, and pressure relief valve.

The conversion apparatus consists of a conversion vessel fabricated from 304SS pipe with the back closure welded and the front closure bolted in position. A 304SS manifold is threaded into the front closure which contains (1) a pressure gauge reading 0-600 psi, (2) a pressure relief valve set at 450 psig, (3) two outlet valves (one for H₂S and one for vacuum and purging). As H₂S gas is poisonous, the spent gas is bled through a series of five bottles containing 20v/o KOH; also the conversion vessel and the KOH bottles are operated under a chemical exhaust hood. The vessel is wrapped with heating tapes controlled by variacs. The front closure is sealed with an 1100 aluminum alloy ring. This closure also contains two thermocouple lead-throughs for determining temperatures at two places in the vessel.

Before experimentation began, a discussion was held with the IBM Corporation on the background of the process and their present production line. From this meeting, processing conditions in the patent were narrowed down to more specific values, i.e., a current density of 12 ma/in.², a conversion time of 8 hr., and a range of deposition times. A limited number of experimental specimens (1/4" × 1" × 1 1/2" mild steel) were examined metallographically, and the lifetime vs. current density data were evaluated. The results showed that the 12 ma/in.² current density is suitable for their process and a film thickness of approximately 45 micro-inches is produced in the operation. IBM also processed for evaluation a number of 440C and M-10 tool steel specimens (1/4" ×

1"×1½") through their production line with varying time in the electrodeposition step and with their standard 8 hour conversion time. A metallographic examination of these specimens showed that for film thicknesses of 45, 90, and 170 micro-inches, the electrodeposition times were 6, 12, and 18 min., respectively. The film appeared converted and adherent for all specimens (Figure 4).

The developmental work at the Goddard Space Flight Center for adaptation of this process to space components started with the design and set-up of the laboratory *in situ* processing apparatus. After the minimum amount of apparatus was set up, work began on surface preparation and electrodeposition for a study of their effects of film thickness and adherence. Various modifications were made to these two steps in the process before a suitable film was produced. Table 1 is a summary of the solutions used and results obtained.

After the conversion vessel and its ancillary apparatus were set up and tested for leaks, specimens with electrodeposited films were processed under various temperatures, pressures, and times. The results are also shown in Table 1. After the developmental work was completed, a number of specimens were processed for the various tests and substrate reaction examination.

A summary of the major differences between the IBM-patented process and the adapted process are shown in Table 2.

TESTS AND RESULTS

To evaluate the merits of the *in situ* method of depositing MoS₂ onto a space component, a laboratory test program was established. This program covered the following areas:

1. Determination of the coefficient of friction.
2. Wear life in air and vacuum.
3. Life test of instrument gears coated *in situ*.
4. Life test of R-2 size instrument bearings coated *in situ*.
5. Effect of H₂S on substrate materials.
6. Film thickness vs. time in plating bath film adherence.
7. X-ray diffraction analysis of *in situ* film.

Where possible, tests were conducted on calibrated equipment under the control of experts in the various fields of lubrication. In this manner



(a) 45 micro-inches



(b) 90 micro-inches



(c) 170 micro-inches

FIGURE 4.—Photomicrographs showing the IBM *in situ* film with the thicknesses of 45, 90, and 170 micro-inches. Because of pullout of the MoS₂ layer during specimen preparation, the focus is on the MoS₂ and the plane of the steel is out of focus. Mag. 1000X, as polished. All photomicrographs have been reduced for printing.

Table 1
Summary of Trials.

Trial	Procedure	Deposition Bath	Material	Remarks	Conversion
I	Ultrasonically cleaned in trichloroethylene for 5 min., air dried	12.6 gm MoO ₃ + 54.6 gm NH ₄ COOH + 1000 ml distilled H ₂ O at 60°C, bath agitated	6061A1	Film nonuniform and flaky	Not converted
II	Hot detergent wash + I, polished through 400 grit paper	Same as above	6061A1	Film nonuniform and flaky	Not converted
III	II + I + 10% NaOH at boiling temperature for 5 min.	Same as above	6061A1	Film nonuniform and flaky	Not converted
IV	I + 2 min. immersion in 2V/o HF + 20V/o HCl	Same as above	440C SS	Film nonuniform and flaky	Not converted
V	I + IV + a rinse of 0.1V/o HCl + 1.0V/o H ₂ SO ₄	Same as above	M-10 tool steel	Film nonuniform and flaky	Not converted
VI	I - 2 min. immersion in 2V/o HF + 20V/o HCl	12.6 gm (85V/o) MoO ₃ + 54.6 gm NH ₄ COOH + 1000 ml distilled H ₂ O at 60°C, bath agitated	M-10 tool steel	Film black, uniform, and adherent	Converted in 6 hr at 397 psig, 140°C, film dull gray.
VII	I + hot detergent wash + 20V/o H ₂ SO ₄ at 82°C (anodic) + a rinse in 0.1V/o HCl and 1.0V/o H ₂ SO ₄	Same as above	M-10 tool steel	Film black, uniform, and adherent	Not converted
VIII	Same as above	12.6 gm (85V/o) MoO ₃ + 54.6 gm NH ₄ COOH + 1000 ml distilled H ₂ O, bath at room temperature	Four 416 SS gears	Film gray, grainy, and uniform	The two thin gears were shiny and uniform. The two thick gears were dull. It was converted in 6 hr at 400 psig
IX	Same as above	Same as above	Three R-2 440C bearing components	Film non-adherent in spots	Film dull and adherent, converted in 6 hr at 400 psig
X	I + hot detergent + wash + a rinse in 0.1V/o HCl and 1.0V/o H ₂ SO ₄	Same as above	440C SS	Film adherent and black	Film shiny and adherent, converted in 10 hr at 400 psig
XI	I + hot detergent wash + 20V/o H ₂ SO ₄ at 82°C (anodic) + a rinse in 0.1V/o H ₂ SO ₄ and 50V/o HCl (anodic)	Same as above	M-10	Film glossy and black	Film dull, satin, and adherent, converted in 18 hr at 400 psig
XII	Same above	Same as above	416 SS gear	Film flaky, dull and black	Excessive peeling of film, converted in 12 hr at 400 psig
XIII	Hot detergent wash + trichloroethylene (both ultrasonically for 3 min.) + immersion in hot 20V/o H ₂ SO ₄ for 1/2 min.	Same as above	416 SS gear	Film uniform and adherent	Film not completely uniform, flaky in spots, converted in 8 hr at 400 psig
XIV	Same as above	Same as above	440C SS (test specimen for Alpha LFW-1 test.)	Film black and adherent	Film uniform, black and adherent, converted in 4 hr at 400 psig

Table 2
Comparison of the Processes.

Condition	IBM Patent	Adapted
Surface cleaning	Not specified	Chemical reactants
Plating time	Not specified	specific time—3 to 10 min.—determined by thickness desired
Conversion time	30 min. to 110 hr.	4 to 8 hr.
Conversion temperature	60° to 125°C	195°C
Conversion pressure	Atmosphere to 400 psi	395 psig

the results obtained could be correlated having MoS₂ as the lubricant. Also, when possible, actual space components were used as the test specimens; for example, instrument bearings.

The facilities of the Naval Research Laboratory were utilized to determine the coefficient of friction. The equipment used was a standard Bowden-Leben friction and wear machine. The test conditions were a 1/2 in. diameter hardened steel ball, a 2 cm/sec stroke, and an 800 gm load on the ball. The three test specimens were M-10 steel, one with no coating other than atmospheric contaminants and a 32 rms surface finish, one with 150 micro-inches of MoO₃ complex film, and one with 150 micro-inches *in situ* MoS₂ film. The results are: for the coefficient of friction for bare metal, 0.32 average, for MoO₃ complex, 0.12 average, and for MoS₂ *in situ*, 0.05 average. The latter coefficient of friction is approximately the same as that for natural MoS₂ powders. The same coefficient of friction was obtained in the wear life tests.

Two different tests were performed to determine the wear life of the *in situ* film, one performed in air on a Alpha LFW-1 machine and one performed in vacuum using a disc and rider set-up. The first test specimens were Timken rings coated with 200 micro-inches of *in situ* MoS₂ and the friction and wear test was performed per military specification MIL-L-25504A Lubricant, Solid Film. This test includes controlled humidity (50 percent), 72 rpm, test block at R₆₀, and a step-wise loading of the block from 30 to 630 lb. The average Hertz stress was of the order of 80,000 psi.

The wear life data shows an average cycles-to-failure for the MoS₂ *in situ* film of 120,000, compared with 100,000 for burnished MoS₂ powder film, 60,000 for sodium silicate bonded film, and 300,000 for epoxy-bonded MoS₂ film.

The second wear life test was conducted by Lewis Research Center on their vacuum friction and wear test equipment. The conditions of the test are 10⁻⁹ torr pressure, a 396 ft/min. disc speed, and a rider load of 200 to 1000 gm applied in increments. The rider is hemispherical in shape with a 3/16 in. radius. Both disc and rider are of 440C SS. The disc had an *in situ* film thickness of 200 micro-inches. The results* of this test showed the wear life of the film to agree with the results found on the Alpha LFW-1 machine in air.

To determine how this film would function in actual application, space components (instrument gears and bearings) were coated with the *in situ* film and subjected to life tests. The instrument gears were Pic precision class 3, of 48 diametrical pitch and 416SS, heat treated to R₃₀. Four gears were coated with 250 micro-inches and loaded with 20 oz.-in. of torque in a four space gear test fixture (Figure 5 and 6). The gears operated for 3.4 × 10⁶ cycles with a load of 8.7 lb./in. of tooth width before failure. Tests are now underway to compare this film with sodium silicate bonded burnished MoS₂ powder and epoxy bonded films.

The second space component to be coated with *in situ* MoS₂ and tested were R-2 size instrument

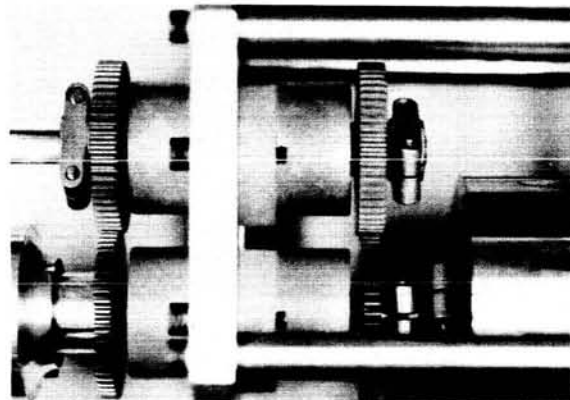


FIGURE 5.—The *in situ* coated gears assembled in the four square gear tester, before the test.

*Personal communication with Mr. D. H. Buckley, Lewis Research Center, NASA, Cleveland, Ohio.

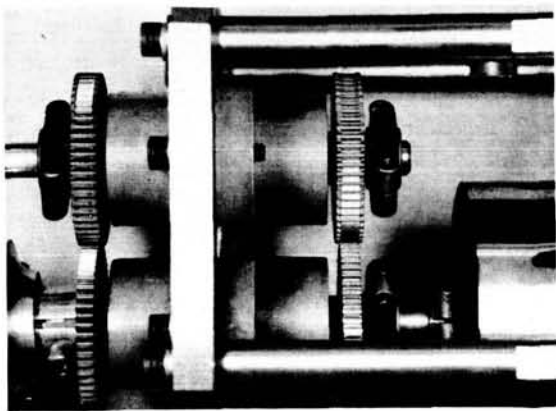


FIGURE 6.—The *in situ* coated gears after the test. Note the wear path of the narrow gear in the larger gear.

ball bearings. The outer and inner races and the machined retainers were processed with 150 micro-inches of *in situ* MoS₂. Two bearings were set up in a bearing tester with a 5 oz. radial load, at 600 rpm and in vacuum. They have operated over 3200 hr. in a vacuum of 10⁻⁹ torr with no indication of failure. A comparison of this film with the other MoS₂ films will be made when the test fails. Figure 7 shows the coated bearings and Figure 8 shows typical coated specimens.

As H₂S gas is very reactive to most materials, and since the materials are subjected to this gas at 195°C and 400 psig for 4 to 8 hours during the conversion step, a program was set up and completed to determine the effect, if any, of the pro-

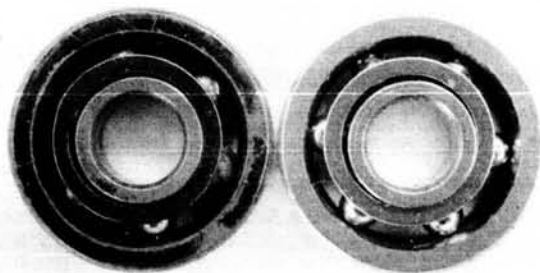


FIGURE 7.—The R-2 size instrument bearing coated with the *in situ* film. The bearing at left has the coating. Note the difference in color. Mag. 4X; reduced for printing.

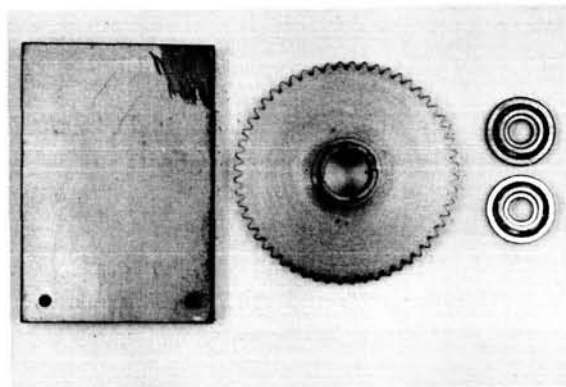


FIGURE 8.—Typical specimens coated with the *in situ* film. Mag. 1X; reduced for printing.

cessing conditions on the substrate materials. A number of specimens—2024, 6061, and 7075 aluminum alloys, 303, 316, 416, and 440C stainless steels, mild steel, and M-10 tool steel—were processed with an excessively thick film at 250° to 300°C for 8 hr., under H₂S gas to 400 psig. The high temperature was necessary to preserve pressure because of a leak in the gas system, and the longer conversion time was due to a desire to accelerate the possible H₂S attack. The specimens were prepared for metallographic examination. For examination at up to 1250 magnification, under bright light, polarized light, and sensitive tint, no indication of attack on the microstructure could be detected (Figures 9–12).

A study of the heat treating practices for the above materials shows that at 195°C the temper



FIGURE 9.—Photomicrograph showing the absence of any reaction of the *in situ* processing to 1020 mild steel. Mag. 360X, 4 percent Nital etch.



FIGURE 10.—Photomicrograph showing the absence of any reaction of the *in situ* processing to 416SS. Mag. 360X, hydrofluoric picral etch.

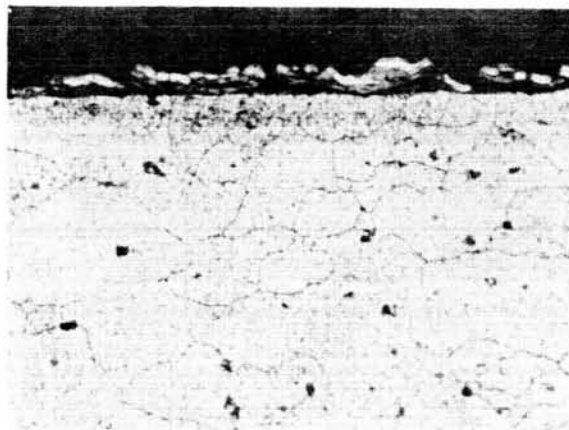


FIGURE 12.—Photomicrograph showing the absence of any reaction of the *in situ* process to 2024 aluminum alloy. Mag. 360X, Kellers etch.



FIGURE 11.—Photomicrograph showing the absence of any reaction of the *in situ* process to 316SS. Mag. 360X, Marbles etch.

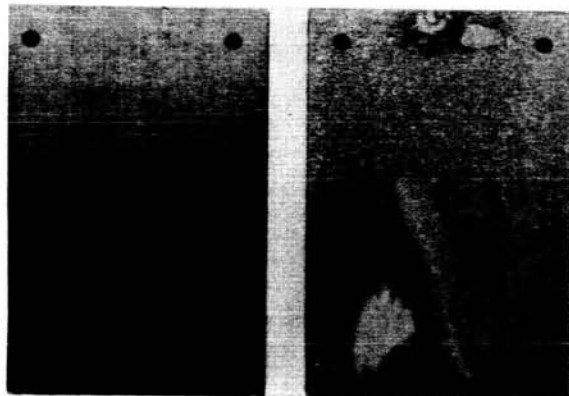


FIGURE 13.—An adherent *in situ* film (150 micro-inches left-hand specimen) and a nonadherent film (300 micro-inches, right-hand specimen).

of the steels will not be affected and that slight over-aging of the aluminum alloys will take place. As this temperature is not excessively over the normal aging temperature, the physical properties will not be adversely affected.

Adherence of the film was a very important factor in deciding the preconversion conditions, such a surface activation and electrodeposition time. The adapted procedure produced a good adherent film. The adherency of the film was determined by shaking or dropping the specimen onto a hard surface and placing a $1" \times \frac{1}{2}"$ piece of cellophane tape onto the specimen under medium thumb pressure and swiftly peeling off the tape.

A good adherent film did not flake or appear powdery (Figure 13). Also, a good film did not peel off with the tape, although a trace of the MoS_2 did show up on the tape.

A number of specimens were processed to determine film thickness vs. time in the electrodeposition bath. The film was measured with a Magne with a Magne Gage, a Dermitron, a Federal Precision Master Comparator, and a microscope. Correlation was good between the Magne Gage and the microscope. Figures 14 and 15 show thickness vs. deposition time. Figure 16 shows film thicknesses.

During the phase wherein conversion time was being determined, samples were submitted for X-ray diffraction analysis to determine if all of the

MoO₃ complex had been converted to MoS₂. Preliminary results showed an amorphorous pattern which indicated that the MoS₂ crystallites produced in this process were too small to be determined with the X-ray diffraction equipment as it was set up. After contacting IBM and discussing this point with them, it was decided that additional analysis would not be performed since IBM had found the same amorphorous pattern in their study. Also, IBM had carried the analysis further and by methods of crystal growth had concluded that the MoO₃ had been converted to

MoS₂. The tests that were made to determine the coefficient of friction and the metallographic examination of the specimen at high magnification also showed that the MoO₃ had been converted to MoS₂.

DISCUSSION

After discussions with IBM personnel and metallographic examination of various coated spec-

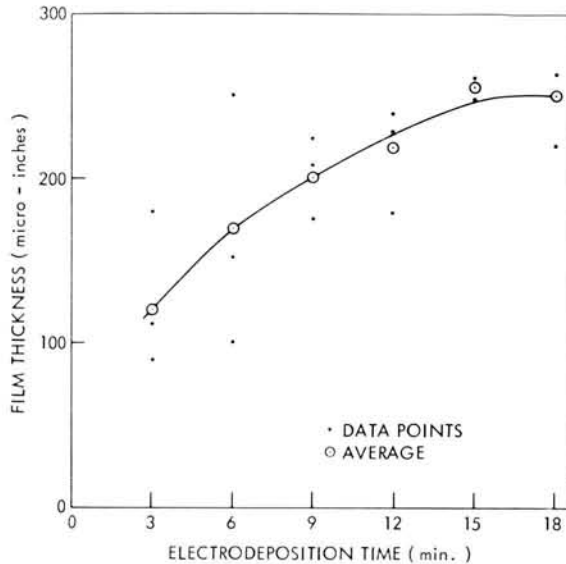


FIGURE 14.—Electrodeposited *in situ* film thickness vs. electrodeposition time.

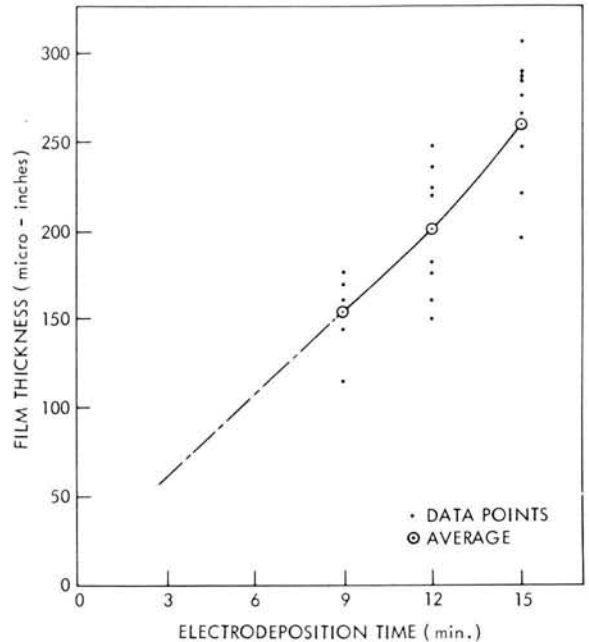


FIGURE 15.—Converted *in situ* film thickness vs. electrodeposition time.

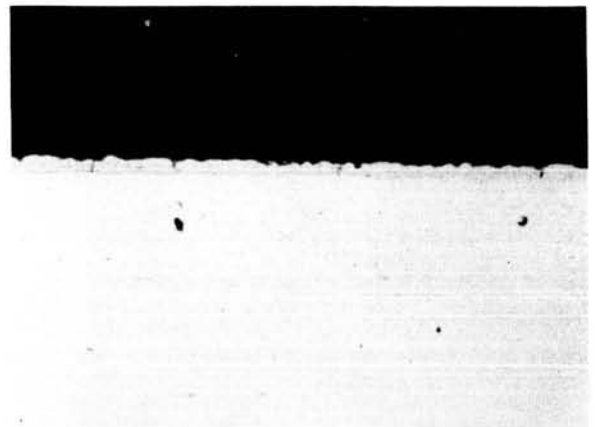
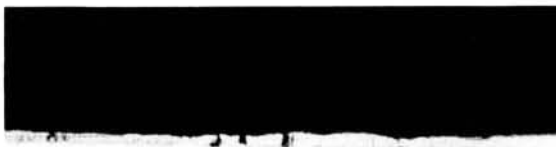


FIGURE 16.—Photomicrographs showing steel specimens coated with *in situ* MoS₂ prepared for microscopic measurement. The thickness is 250 micro-inches Mag. 300X, as polished.

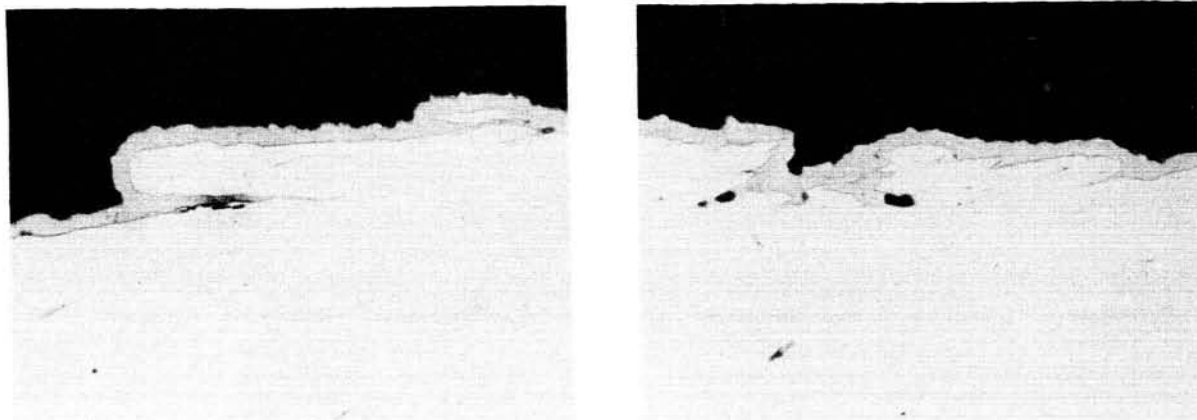


FIGURE 17.—Photomicrographs showing the ability of the *in situ* film to fill cracks, seams, and laps. Mag. 360X, as polished.

imens, it was determined that 120 ma/in.² current density was suitable for the developmental work.

During the metallographic examination of a number of specimens, it was noticed that the film was deposited into all laps, seams, cracks, and micro-fissures (Figure 17). This indicated that this process may lend itself to lubricating hard-to-reach places, such as bearing outer races.

As shown earlier, the wear life of this film is not quite as good as space conditions dictate. Therefore, additional work will be performed on the process in attempt to improve wear life. Present work will be carried out to investigate incorporating the *in situ* process with a reservoir technique. Also, further adjustment to the process, such as changing Mo valence, current density, etc., may produce a longer wear life.

The effects of the temperature of the deposition bath on the film adherence proved undetectable and the elevated temperature caused evaporation of the bath. Hence, the bath was used at room temperature.

Optimum conversion time was determined by using the coefficient of friction and the adherence test. Times of 18, 12, 10, 8, 6, and 4 hr. were used and the results show that 6 hr. is sufficient to convert a 250 micro-inch film.

CONCLUSIONS

An evaluation of the test data shows that the following conclusions can be drawn:

1. The *in situ* process for depositing MoS₂ onto space components as a suitable lubricant is feasible and practical.

2. The procedure for depositing MoS₂ *in situ* has been worked out, an adaptation of an IBM-patented process.
3. Adherence tests show that 250 micro-inches is the maximum film thickness advisable under the processing conditions outlined.
4. The film thickness can be controlled within ± 50 micro-inches.
5. The average coefficient of friction of this film is 0.05 or less and is comparable with or the same as that of MoS₂ powder and lower than those of bonded MoS₂ films.
6. The film can be easily and safely deposited onto a number of common spacecraft materials, including 2024 A1, 6061 A1, 7075 A1, 303SS, 316SS, 416SS, 440C SS, mild steel, and M-10 tool steel.
7. The film has a better wear life than sodium silicate bonded MoS₂, slightly better wear life than a burnished MoS₂ powder, and a somewhat poorer wear life than epoxy bonded MoS₂.
8. The film follows the surface contour and fills up the smallest crack, lap, seam, or indentation and therefore makes it possible to place a controlled amount of MoS₂ on the hard-to-reach surfaces, such as the outer races of ball bearings.

ACKNOWLEDGMENTS

The author wishes to express his appreciation to the following people who have helped on this project: Mr. A. DiSapio of Alpha Molykote Corp.,

Stamford, Connecticut and Mr. Donald H. Buckley, NASA Lewis Research Center. The author also wants to thank personnel of the Commercial Product Division, Development Laboratories, IBM Corporation, Endicott, New York.

REFERENCES

1. DEVINE, M. J., LAMSON, E. R., and BOWEN, J. H. JR., "The Lubrication of Ball Bearings with Solid Films," Amer. Soc. Mech. Engineers paper No. 61-LUBS-11, New York.
2. GODFREY, D., and BISSON, E. E., "Bonding of Molybdenum Disulfide to Various Materials to Form a Solid Lubricating Film," NACA Technical Note 2628, February 1952.
3. BROPHY, J. E., and INGRAM, R. W., "Application of Solid Lubricant Coatings to Surfaces," U. S. Patent 2,902,417 (assigned to the IBM Corp.), September 19, 1956.
4. BROPHY, J. E., INGRAM, R. W., and PITTERS, C., "Low Friction Coatings," *Materials in Design Engineering*, September 1961.
5. BUCKLEY, D. H., SWIKERT, M., and JOHNSON, R. L., "Friction, Wear, and Evaporation Rates of Various

Materials in Vacuum to 10^{-7} mm Hg.," *ASLE Trans.* 5(1):8-22, April 1962.

BIBLIOGRAPHY

- BOWDEN, F. P., and ROWE, G. W., "Lubrication with Molybdenum Disulfide Formed from the Gas Phase," *Engineer*, 204(N5311):667, November 8, 1957.
- GODFREY, D., and BISSON, E., "Bonding of Molybdenum Disulfide to Various Materials to Form a Solid Lubricating Film. II-Friction and Endurance Characteristics of Films Bonded by Practical Methods." NACA Technical Note 2802, October 1952.
- LAFFOON, C. M., JR., "Application of Solid Lubricant Coatings to Surfaces," U. S. Patent 2,420,886, May 20, 1947.
- PETERSON, M. B., and JOHNSON, R. L., "Friction and Wear Investigation of Molybdenum Disulfide. I-Effect of Moisture," NACA Technical Note 3055, December, 1953.
- PETERSON, M. B., and JOHNSON, R. L., "Friction and Wear Investigation of Molybdenum Disulfide. II-Effects of Contaminants and Method of Application," NACA Technical Note 3111, March, 1954.
- PETERSON, M. B., and JOHNSON, R. L., "Friction of Possible Solid Lubricants with Various Crystal Structures," NACA Technical Note 3334, December 1954.

AN INTEGRATED SYSTEM FOR COLLECTING AND ANALYZING ENVIRONMENTAL TEST DATA*

PHILIP YAFFEE, FRED STARBUCK, JAMES W. BAILEY, AND CLIFFORD W. SCOTT

AND

AUGUSTUS C. JOHNSON, C. FRANK RILEY, JR., WILLIAM E. FLOWERS, AND RICHARD K. HOVEY

Booz-Allen Applied Research, Inc.

The simultaneous evaluation of several spacecraft in simulated space environments results in a large volume of data relating to the performance of the spacecraft and to the generation and measurement of the environment. The timely and effective collection, processing, and reduction of these data can only be achieved by means of an automatic data-handling system. This system must be versatile, accurate, and reliable. Such a data system has been designed and installed at Goddard Space Flight Center. It organizes and reduces the data generated, both as an aid to effective control of test operations and for optimum analysis and storage of the test data. The experimenters and test personnel are released from routine data gathering and reduction and permitted to concentrate on data analysis and interpretation.

INTRODUCTION

Goddard Space Flight Center has a facility for testing and evaluating the performance of instrumented satellites and probes in simulated space environments (Figure 1). This Spacecraft Test Facility includes thermal-vacuum chambers, temperature-humidity chambers, optical test chambers, vibration test systems, centrifuges, moment of inertia machines, spin and balance equipment, and tension-compression test equipment. Data must be collected and analyzed for use in controlling the spacecraft environment and the stimuli applied to satellite sensors, for determining the outputs of the sensors through the satellite telemetry systems, and for observing the internal functioning of a satellite at critical system points.

During the planning for this facility, it was decided that great advantages in speed and convenience could be realized by providing a largely automated data collection system for the entire facility. The system as installed consists of portable data collection modules, a permanent transmission cable system, a section known as Data

Central (Figure 2) where data are assembled, recorded, and analyzed, and an Operations Center for monitoring spacecraft tests.

The capability of the data system can best be visualized by considering its application in the test of an observatory. This is a satellite which may contain a large number of separate experiments-instrument packages designed to study separate physical parameters in the orbital environment. These experiments use a common vehicle, power supply, and transmission system, but each experiment may be designed and fabricated by a separate group. Various government agencies, universities, and industrial laboratories may build experiments for a single observatory.

When such a satellite is undergoing a test in a simulated space environment, literally hundreds of different data channels are needed to acquire and organize the information required in evaluating the observatory's performance. Data must be collected and displayed for the use of the experimenter, those who control the environments, and those who evaluate the overall performance of the observatory. Each of these requirements may have its own time frame and format. The complexity of the design requirements for this data

*Published as NASA Technical Note D-2322, June 1964.

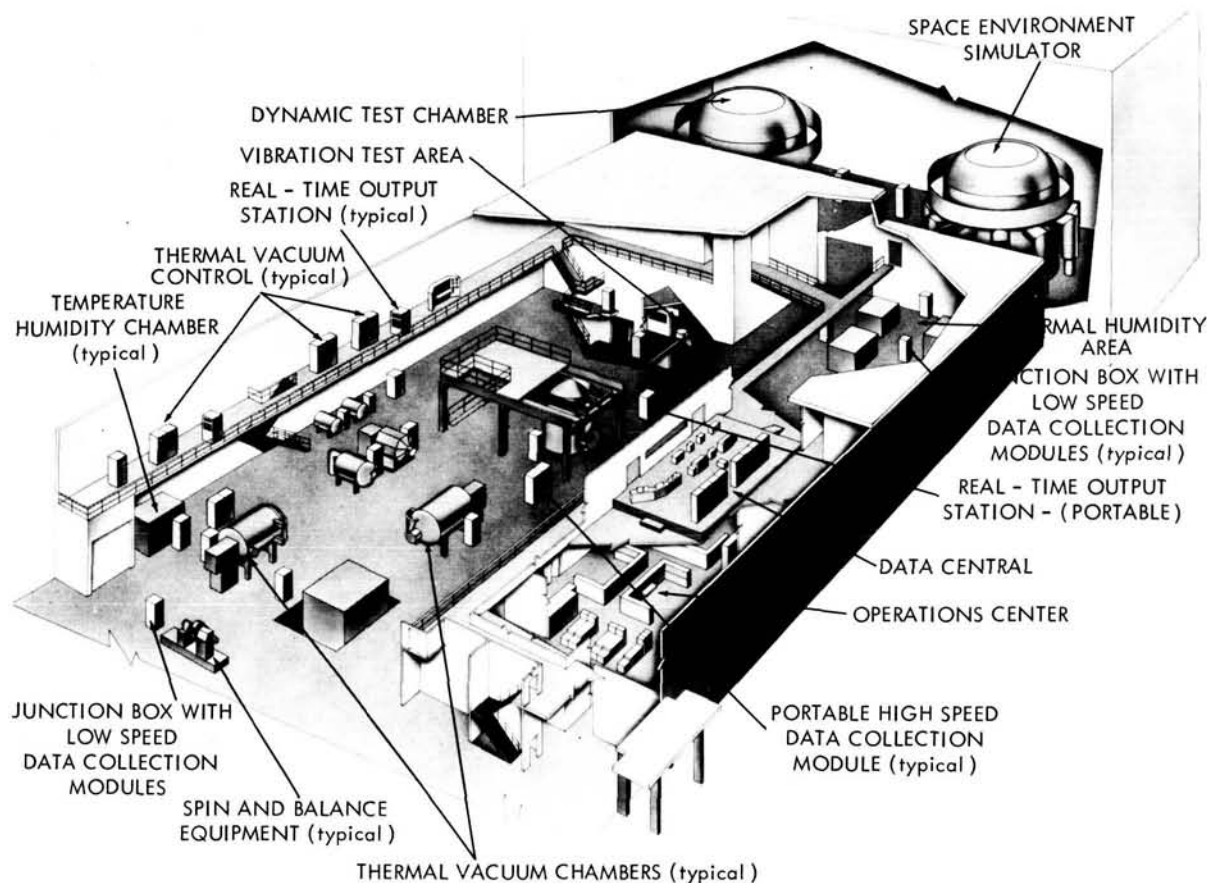


FIGURE 1.—Spacecraft Test Facility, buildings 7 and 10, Goddard Space Flight Center.

system may be appreciated by considering that three major tests and a number of simpler tests may be conducted in the facility simultaneously.

SYSTEM CHARACTERISTICS

The principal characteristics of the system are its versatility, its accuracy, and the careful control which is maintained over the data collected.

Data can be collected from approximately 3500 different points, selected from more than 10,000 points distributed in some 23 different physical locations in the 60,000 ft² building complex. Sampling rates range from the order of a few samples per hour to 15,000 samples per second. A wide variety of signal conditioning is possible, including thermal referencing, bridge balancing, amplification, and attenuation. Setups and changes can be made conveniently and quickly without interrupting other tests.

The system in operation handles data well within the specified span accuracy of 0.25 percent. This accuracy was achieved by careful engineering



FIGURE 2.—Data Central.

of the entire system, including the long runs of shielded analog cables. It also required an engineered instrument grounding system for the building, and a rigidly enforced shielding and grounding discipline. Anything short of the most meticulous care in this regard could cause extreme signal degradation.

The unifying element of the system is the flexibly assigned but rigorously controlled format, into which the overall system data are assembled. This format contains 7500 eleven-bit-plus-sign words per second. The format is enforced by the master system clock, which provides the timing for all sampling, conversion, and assembly in the system except for telemetered data which is asynchronously controlled by clocks on the satellites. The result is a stream of data, each item precisely identified and placed in time. A space in this format is reserved for each data channel, whether or not it is presently in use.

The flexibility of assignment is achieved by using a removable patch panel at the test site and control switches in Data Central. The patch panel permits the pretest assignment of individual channels to individual format spaces, including supercommutation and subcommutation. The switches in Data Central select the block location of each of the high speed data collection modules in the data format. These switches can supercommutate the collection modules within the format, thereby increasing the sampling rates by a factor of 2, 5, or 10 over the rates selected by the patch panel. An on-line computer, under program control, selects data words from the format for reduction and display.

The philosophy of system operation is that the outputs are to be kept to a minimum by planning in advance the information to be delivered to each output station. This precludes tying up the system with masses of raw output, and permits the computer to produce sophisticated results, eliminating the need for further manual reduction. Thus, technical personnel can devote their efforts to data analysis and interpretation, rather than to routine data collection and manual reduction.

DATA COLLECTION

Low Speed Data

Slowly varying parameters such as temperatures, pressures, and relatively stable voltage and

current levels are measured by a low speed subsystem. This subsystem takes 25 samples per second and has a nominal capacity of 2500 separate data channels. The data from this subsystem are handled in two parallel ways. First, all the data are recorded on a low speed magnetic tape for later off-line processing. Second, the low speed data are intermingled with the high speed data in the overall system data format; then they are sent to an on-line computer which provides instantaneous readings in engineering units for the direct monitoring of facilities and experiments.

High Speed Data

Electrical measurements can be made at a system rate of 7500 samples per second on a maximum of 1000 separate channels. The measurements which can be made include:

1. Current and voltage levels on satellite power circuits.
2. Radio frequency power.
3. Voltages, currents, and transducer outputs in several ranges from ± 8 millivolts to ± 320 volts full-scale, with a system bandpass of 500 cps.

The measurements taken by this subsystem, with the low speed data intermingled, are handled in two parallel modes. In the normal mode, a burst of 6 seconds of data is recorded on magnetic tape with a 5 minute interval between bursts. Simultaneously, all the data are recorded in "ping-pong" fashion on two tape transports operating alternately. Each transport records 5 minutes of data and then rewinds and waits while the other transport records for 5 minutes. If no emergency of malfunction occurs, the data are erased through re-use of the tape. If an emergency is detected, all the data recorded during the previous 10 minutes can be analyzed. In addition to these two recording modes, a limited amount of real-time data processing by the on-line computer is possible.

Extra-High-Speed Data

To handle rapidly fluctuating data, a special arrangement can be made by which 100 channels of electrical measurements of the same types as those described above can be taken at a system

rate of 15,000 samples per second. These measurements are recorded on a high speed tape transport holding approximately $3\frac{1}{2}$ minutes of data. The data are handled in parallel with but entirely separate from the data within the overall system format.

Data Recorded Outside the System

Data can be introduced into the system which have been recorded elsewhere in the form of punched tape, punched cards, or magnetic tape, either digital or analog. These data can then be converted to other forms or analyzed and compared with data collected in the test facility.

Telemetered Data

Data coded and multiplexed in the telemetry systems of satellites being tested within the facility can be introduced into the computer subsystem and recorded on tape for future processing at a maximum aggregate data rate of 14,000 twelve-bit words per second. When real-time processing of these data is desired, a second (off-line) computer is brought on-line through the central memory.

Running-Time Data

The running times of up to 230 major items of equipment within the facility are recorded by the running-time subsystem. This subsystem records on punched paper tape the clock time, the channel designation, and the on or off status of each device. Fifty of the samples are taken once every minute and the remaining 180 devices once every 5 minutes. A reel of tape holds 4 days of data produced by this subsystem. Tape is removed biweekly and processed by the computer to produce updated machine-maintenance records and required maintenance reports.

DATA OUTPUTS

Real-time output stations (RTOS), either permanent or portable, are located at convenient spots near the test facilities (Figure 3). Under computer program control each station responds to an inquiry by displaying the reading of a designated channel on an in-line in-plane display, by plotting on a strip chart, or by printing out in engineering units the readings of a number of designated channels. Primarily provided for the

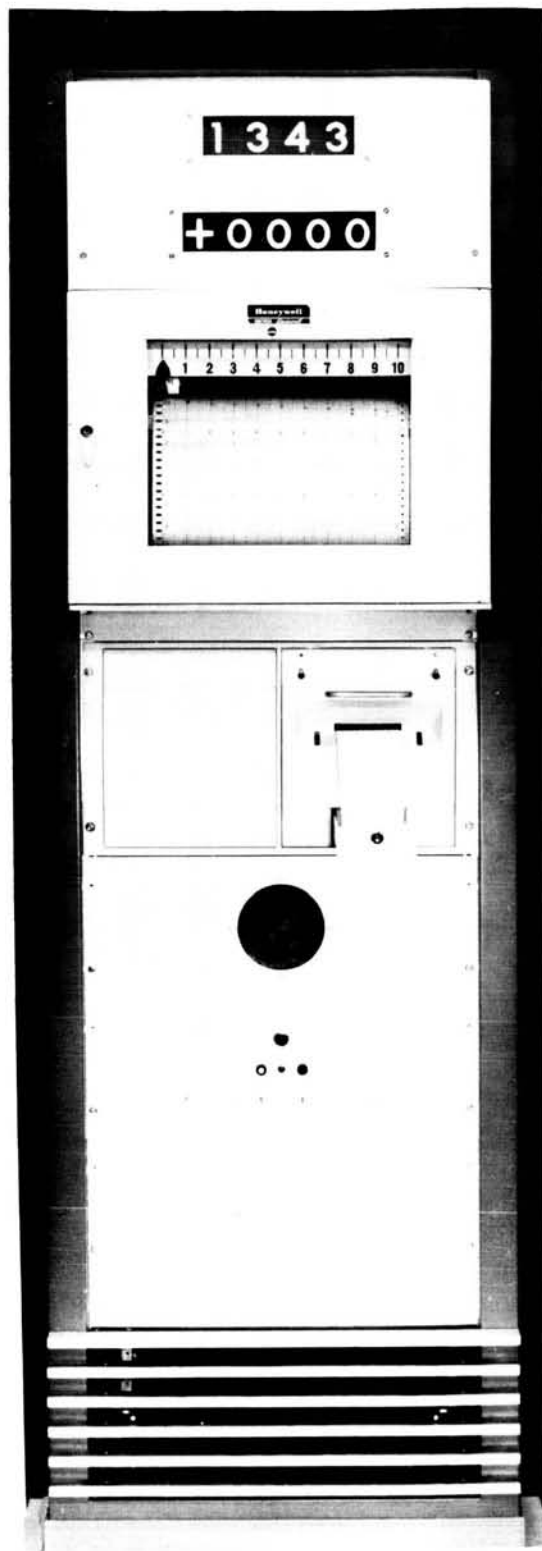


FIGURE 3.—Real-Time output station (RTCS).

display of environmental data, these stations can be programmed to display the output of any channel of the on-line system.

Four cathode-ray bar-graph displays are located in the Operations Center. Each display shows the instantaneous readings on as many as 150 channels. A reading is displayed as a dot whose height above the base is proportional to the scaled reading on the designated channel. A negative voltage is designated by a dot below the horizontal base line. The horizontal scale can be expanded so that any 10 of the 150 channels may be examined closely. These displays are for the use of the experimenters and test personnel. They permit channel readings to be monitored in real time. Thus malfunctions or gradual changes in satellite system operating characteristics can be examined and the effect of sensor exercising can be observed. A malfunction indicated here is the normal reason for reviewing the data recorded on the "ping-pong" transports.

The satellite battery-power currents and voltages, also collected by the combined systems, are displayed on ammeters and voltmeters, and the RF power is displayed on wattmeters, on the collection modules located near the test stations.

Data recorded on tape by any part of the system, or stored in the computer memory in real time, are reduced and analyzed by the computer subsystem. The outputs are magnetic or punched tape, cards, typed reports, printouts, and X-Y plots. Considerable effort is devoted to ensure that these outputs are in final form for inclusion with test reports, or in convenient form for engineering or scientific interpretation.

SYSTEM CONFIGURATION AND DATA FLOW

The setup of the various components of the data-handling system and the flow of data from these components will now be discussed (Figure 4). The system configuration contains junction boxes and high speed and low speed data collection modules. The data are carried via various means into the Master Assembler and then to bar graph displays, tape transports, and the computer.

Junction Boxes

Junction boxes are permanently located at convenient spots near the major test facilities (Figure 4). They serve as termination points for the

cables from the Operations Center and Data Central. These cables transmit data into Data Central and provide control connections for the experimenters. The junction boxes also house the low speed data collection modules (DCML's) described below, and provide quick-connect plugs for the high speed data collection modules (DCMH's), also described below.

Eleven test areas are served by major junction boxes, and the remaining twelve by minor junction boxes. A major junction box will accommodate an observatory, providing the system capability of connection to as many as six DCMH's and four DCML's, thus handling about 600 high speed and 200 low speed measurements simultaneously. In addition to the measurement channels, each major junction box contains the terminals of 60 analog control circuits for sensor exercising and 60 switching circuits for satellite control.

A minor junction box accommodates two DCMH's and two DCML's, providing about 200 high speed and 100 low speed measurement channels. In this case, 30 analog control circuits and 10 switching circuits are available.

Wires from the satellite and from the test facility enter the junction box from the top or bottom rear. Cables to the Operations Center and Data Central emerge through the bottom of the junction box and pass through the floor into the cable trays suspended from the ceiling below. The trays conduct them to a cable vault near the Operations Center for distribution.

High Speed and Low Speed Data Collection Modules

The high speed data collection modules (DCMH's, Figure 4) are mobile. They have the same width as two standard relay racks, are 78 inches high, and are equipped with wheels (Figure 5). In practice, they are positioned next to the junction box to which a test is wired. Through quick-connect plugs the modules are connected to the setup wiring and to the building cables, both of which terminate in the junction box.

A DCMH can handle 100 channels of data. It provides power, voltage and current measurement and indication, transducer excitation power supplies, bridge completion units, amplification, and attenuation. Highly flexible provisions are made

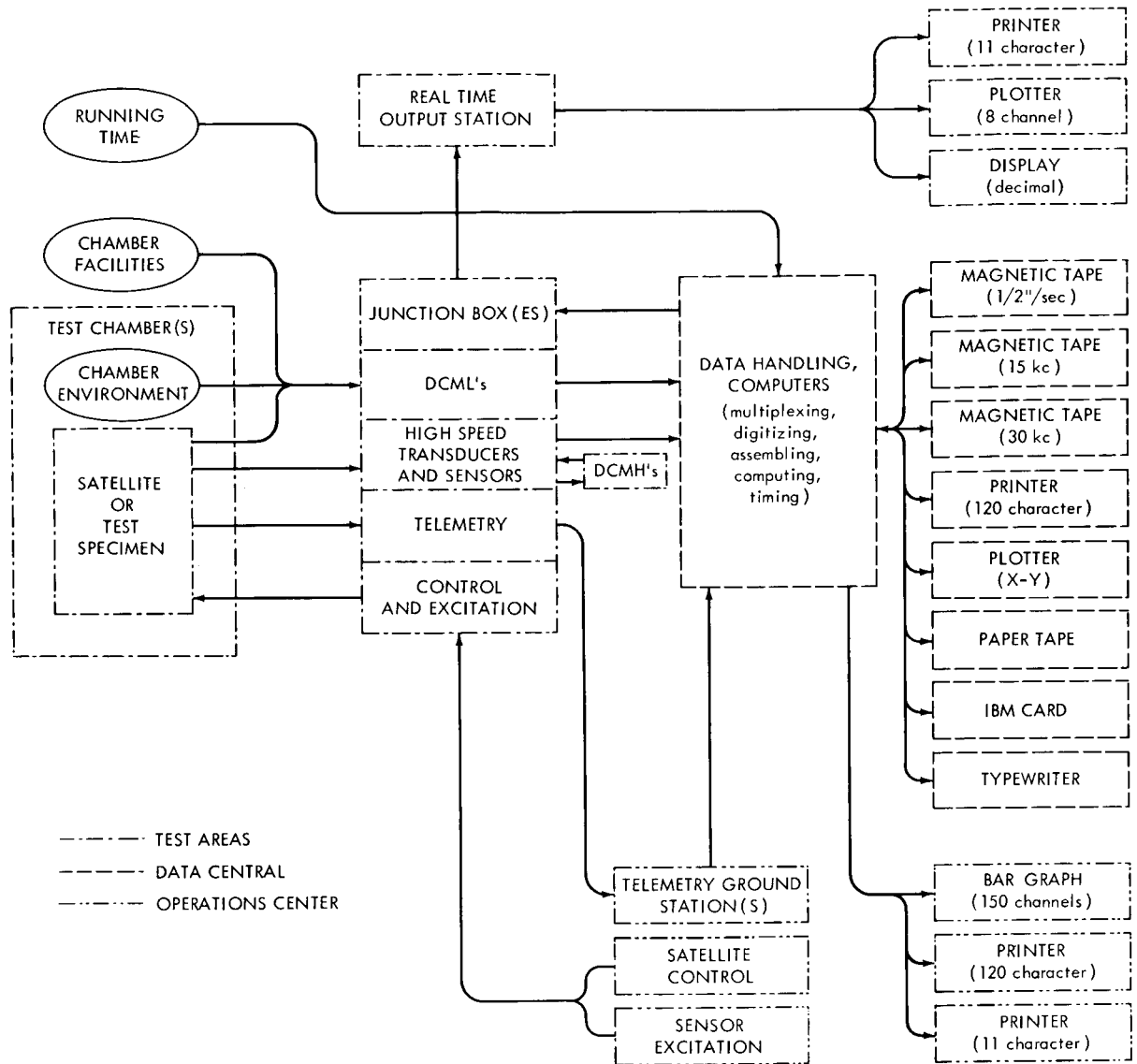


FIGURE 4.—System configuration and data flow.

for channel selection, supercommutation, and subcommutation by means of a removable digital patch board which can be prewired in preparation for a test and quickly inserted in the DCMH. In the DCMH these various input signals are conditioned and amplified or attenuated as necessary to a full-scale value of ± 5 volts. They are transmitted to Data Central through the permanent cabling as a pulsed analog signal at this level.

The low speed data collection modules (DCML's, Figure 4) are semiportable (Figure 6).

When in use they are housed in junction boxes most convenient to the points of use. Each module is capable of receiving inputs from approximately 50 separate data pickups. A module provides thermocouple reference junctions, power supplies, and bridge completion and bridge balancing units for conditioning the data on these 50 channels. It receives the data through the wiring connecting it to the test facility or specimen, multiplexes it by means of a 50-point stepping switch, and sends it through the permanent

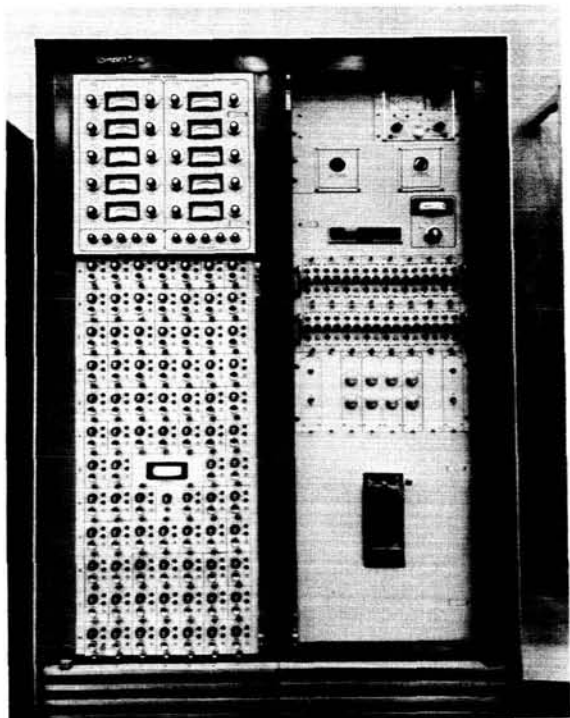


FIGURE 5.—High speed data collection module (DCMH).

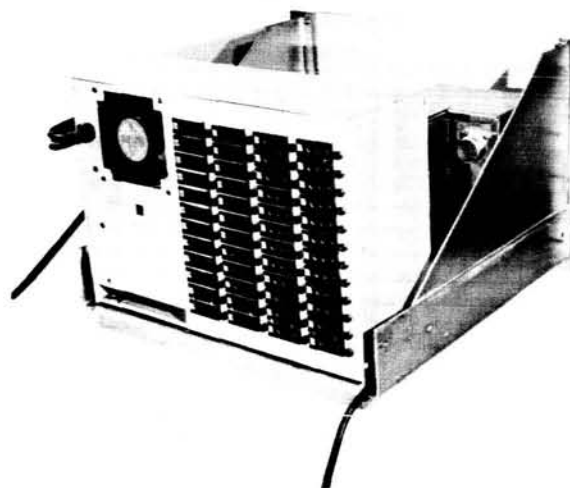


FIGURE 6.—Low speed data collection module (DCML).

building cables to the cable vault in the form of a pulsed, low-level analog signal of ± 10 millivolts full-scale.

Data Transmission

The DCML data received through the permanent cables at the cable vault are multiplexed by

a crossbar switch, amplified to a full-scale value of ± 5 volts, and converted to digital form. The data are then driven through two parallel circuits. One circuit records them on the low speed data tape in Data Central; the other sends them to the Master Assembler in Data Central to be merged with the DCMH data.

In addition to this normal mode of operation, it is possible to switch the DCML data, in high-level pulsed analog form, into the Master Commutator of the high speed system. This emergency circuit takes care of a low speed converter failure.

The DCMH data are transmitted through the building cables and a patch panel to the Master Commutator in Data Central. After analog-to-digital conversion, they are sent to the Master Assembler. Here the DCMH data and DCML data are merged into the overall system data format.

Analog-to-Digital Conversion

The low speed system for DCML data, as noted above, has its own analog-to-digital converter. The high-speed data are converted by a converter which immediately follows the Master Commutator. The converters operate on a programmed successive approximation principle, converting each analog signal into digital form as an 11-binary-bit-plus-sign number. This gives a digital scale from -2000 to $+2000$ counts, corresponding to a transmitted analog signal of from -5 to $+5$ volts, and to full-scale negative and positive measurements at the respective data transducers. Digital output is limited to a maximum of ± 2016 counts.

The converters operate up to 7500 conversions per second, being driven by the same clock that drives the commutators, to provide absolute synchronism. The digital outputs of each converter appear on 12 parallel lines; these outputs set the positions of 12 flip-flops in the Master Assembler.

Overall Data Format

The 7500 eleven-bit-plus-sign words per second in the data format include 25 data points from the low speed subsystem, 25 identification words for low speed data, 25 identifications and control words, and 7425 data points from the high speed subsystem.

In addition to the data received from the analog-to-digital converters, the Master Assembler receives the digital identification and control data and inserts them in the proper points in the data flow. This includes such information as the job run, station, test run numbers, test conditions read from switches at the test control stations (beginning or end of test, calibration, etc.), and time information read from the system clock (day, hour, minute, second, and tenth of second). Also, switches in Data Central permit data to be flagged as setup data, calibration data, or actual test-run data. These flags, introduced at the Master Assembler, serve as instructions to the computer for the subsequent manipulation of the data.

The extra-high-speed rate of 15,000 data points per second is handled according to a different scheme. One DCMH (100 channels) is assigned, when necessary, to this special use. Two multiplexers, two transmission lines, and two analog-to-digital converters operate alternately, each at 7500 cps. A double-sized assembler receives the data from the alternating converters and it is recorded on a 30,000 character-per-second tape transport.

From the Master Assembler, assembled data are switched to the bar graph displays, to the tape transports, and on-line to the computer, where they are used as determined by the test setup.

Computation

The computer subsystem performs as an on-line device; it is an integral part of the data collection system. Engineering information is accepted, processed, and output to printers, plotters, and other display devices, in real time, for the use of those monitoring the test specimen and facilities.

This subsystem can compare a large number of channel readings with preset upper and lower limits and give an alarm when a channel exceeds its limits. It selects the predetermined channels to be displayed by the real-time output stations, converts their readings to engineering units, and transmits the readings to the display units.

The computer subsystem consists of two CDC 160-A computers linked through a shared memory. One computer will normally be connected to the Master Assembler of the data system and

perform the on-line functions described above. The second computer serves as a spare for the first, and also is available for the off-line analysis and output of data fed to it by any of the output media of the data system or stored in the joint memory. It can also be used as a completely separate computer for routine computing tasks.

The off-line functions include processing high speed data previously recorded on tape, processing telemetry data, preparing maintenance reports from data generated by the running-time subsystem, scheduling the test facilities, budget computation, and processing scientific data in support of research.

Data from the telemetry channels of a satellite are demodulated, formatted, digitized where required, and delivered by a ground station in Operations Center into a buffer register. The computer commands the buffer register to "dump" its contents in 12 parallel lines. The resulting word can be stored in the common central memory of the computer subsystem without going through or being controlled by either computer mainframe. When the telemetry data are in memory, they can be called up and manipulated by either computer and the results can be put out in any appropriate form.

CALIBRATION AND MAINTENANCE

Several types of calibration are provided: At each DCML, the first two of the fifty inputs are internal test points, either 0.0 or 8.00 millivolts. Each DCMH provides calibration inputs of ± 4 , ± 2 , 0, -2 , and -4 volts, as well as input voltages derived from each power supply in the DCMH. These are automatically sequenced and sampled by the DCMH at 3-record intervals.

A calibrate switch at the Data Central console initiates the following operation: A reference voltage is applied to each of the eighty conditioning modules in each DCMH; the voltage indicates the setting of each attenuator switch. The multiplexers are disconnected and a reference voltage is fed directly into each buffer amplifier. Each strain gage channel is calibrated by a 2-point linearity check ($\frac{1}{4}$ and $\frac{3}{4}$ full scale).

Voltages at critical points within the system are automatically sampled and these data are processed by the computer, which compares the resulting digital reading with an expected reading

stored in its memory. The pattern of the channels selected to carry these diagnostic inputs is designed so that a logical process permits the computer to isolate a malfunction by elimination, and to print out the location of the trouble or the instructions for eliminating it.

Clearly, it is impractical to perform this automatic diagnosis below a certain system level, and beneath this level a simplified manual diagnostic technique is used. This combination of computer and manual diagnosis is known as the PILE system (*Programmable Identification of Logic Elements*). In this data system, the cutoff is at the major device level; that is, the computer isolates the malfunction to a particular amplifier, analog-to-digital converter, or commutator, with the manual technique taking over for lower level diagnosis.

The manual technique is designed so that relatively inexperienced technicians may operate it satisfactorily. Each cabinet contains a booklet clearly locating the test points and showing the exact wave forms to be expected at each one. The sequence of testing is indicated by writing pseudo-logic equations for each logical circuit, in which each circuit element and function is identified by a short pronounceable word.

The functions schematically described above are actually performed by three complete sets of analog-to-digital converters, assemblers, and commutators located in Data Central. More than one of each type of instrument will be in use simultaneously only when such use is deliberately scheduled. Through convenient patch facilities, the equipment is easily interchangeable when a malfunctioning device must be replaced.

VII. TRACKING SYSTEMS

SOLAR CELL RADIATION DAMAGE STUDIES WITH 1 MEV ELECTRONS AND 4.6 MEV PROTONS*

WILLIAM R. CHERRY AND LUTHER W. SLIFER

The degradation of several types of solar cells irradiated with 1 Mev electrons in stages up to $10^{16}/\text{cm}^2$ was measured under standard tungsten illumination and also under air-mass-one sunlight illumination. The degradation of similar cells irradiated with 4.6 Mev protons in stages up to $3 \times 10^{11}/\text{cm}^2$ was measured under tungsten illumination. The solar cell types included P/N cells, N/P cells with various base resistivities from less than 1 ohm-cm to 25 ohm-cm, and drift field N/P cells.

The results showed that the N/P cells are conclusively more resistant to either type of radiation than the P/N cells and that, within the ranges of the experiments the N/P cells with higher base resistivity are more radiation resistant than those with lower base resistivity. The latter result is most pronounced with the electron-irradiated cells. The results also gave positive indication that the radiation resistance of solar cells is enhanced by the drift field construction.

INTRODUCTION

After the July 1962 Atomic Energy Commission Pacific tests, erratic behavior developed in several U.S. satellites in orbit. Subsequent analysis of data has revealed an enhanced radiation belt consisting primarily of high energy electrons at altitudes between 500 and 4,000 nautical miles.

The need for more radiation-resistant solar cells for NASA spacecraft frequenting the new man-made and the natural belts became apparent. A series of 1 ohm-cm N/P silicon solar cells which been irradiated with 1 Mev electrons at doses ranging from zero to 10^{16} electrons/ cm^2 was used to establish a damage rate specification (GSFC Spec 63-106).

During October and November 1962, state-of-the-art samples of N/P silicon solar cells were purchased from eight sources within the United States. A quantity of P/N silicon solar cells was also purchased for comparative purposes.

EXPERIMENT

In conjunction with personnel at the Naval Research Laboratory, an elaborate experiment was planned using 1 Mev electron and 4.6 Mev proton irradiations. A standard light source

described in Appendix A was used to select 30 cells from each group with median current voltage characteristics of the lot. The initial current, voltage, spectral response, and diffusion length characteristics of each cell were measured. The cells were then irradiated to the various dosage levels shown in Table 1, taking selected current, voltage, spectral response, and diffusion length measurements.

A summary of all individual short current measurements, average short circuit currents per group, percentage change in average short circuit current, and percent change in average maximum power is presented in Appendixes B and C for both electron and proton bombardments. These values were obtained using the standard tungsten light described in Appendix A.

Caution must be exercised in using this data. The cells were state-of-the-art cells; thus some were representative of production in sizable quantities while others were representative of pilot-line production. After the preliminary meeting held at Goddard on January 23, 1963, when change in I_{sc} as a function of 1 Mev electron dosage was discussed, several of the manufacturers changed the base material used in their cells; thus, results would be quite different today. Furthermore, the cell performance was measured in tungsten light,

*Published as NASA Technical Note D-2098, February 1964.

Table 1
Radiation Test Scheme for Solar Cell Degradation.

1 Mev Electrons			4.6 Mev Protons		
Dosage (e/cm ²)	Number of Cells	Number of Cells Withdrawn at Dosage Level	Dosage (p/cm ²)	Number of Cells	Number of Cells Withdrawn at Dosage Level
0	15	1	0	6	0
10 ¹¹	14	1	10 ¹⁰	6	0
10 ¹²	13	1	3 × 10 ¹⁰	6	0
10 ¹³	12	1	10 ¹¹	6	0
10 ¹⁴	11	1	3 × 10 ¹¹	6	6*
10 ¹⁵	10	1			
10 ¹⁶	9	9			

*Except as indicated in Appendix B.

which will emphasize the damage to a cell more than air mass one or air mass zero sunlight. From the data presented, it is not possible to select the best cell for sunlight use. The experiment was to show the relative damage of a cell to an accepted reference; and it accomplished this very well. Clearly, as seen in Figure 1, the change in I_{sc} for 1 Mev electron bombarded cells fell into four groups, which were a function of the base resistivity and drift field. From the data in Appendixes B even cells with base resistivities less than 1 ohm-cm were apparent.

The cells which were irradiated with 1 Mev electrons to 10¹⁶/cm² were measured for percentage change in I_{sc} under tungsten light and sunlight at Table Mountain, California. Also the percentage change in I_{sc} was measured in tungsten light for the cells irradiated with 4.6 Mev protons. The results are shown in Table 2.

Table 2 shows that within a 1 percent change of I_{sc} , the cells maintained their same order of damage resistance in both tungsten light and sunlight. It also shows that with 4.6 Mev protons, the distinction between the various base resistivities is not as pronounced as with 1 Mev electrons, but the proton studies were not carried to the equivalent orbital dosage as the electrons.

When percentage change of maximum power caused by 10¹⁶ electrons/cm² is used as the index, the relative order is again the same as that for the short circuit current (See Table 3). The base resistivity again established distinct groups of less than 1 ohm-cm, 1 ohm-cm, 10 ohm-cm, and 25 ohm-cm, cells. The drift field cells are still

superior. There is not the clear distinction between base resistivities at a 4.6 Mev proton dosage of 3 × 10¹¹/cm² as with a 1 Mev electron dosage of 10¹⁶/cm², but generally the higher base resistivity is more radiation-resistant.

Graphs showing the percentage degradation of maximum power as a function of proton and electron dosage are presented in Figures 2-12. These data show that with a 4.6 Mev proton dosage of 3 × 10¹¹/cm², there is at least 10 times the radiation resistance for N/P cells using a minimum of 1 ohm-cm base resistivity material than for the P/N cell under the tungsten light source. Base

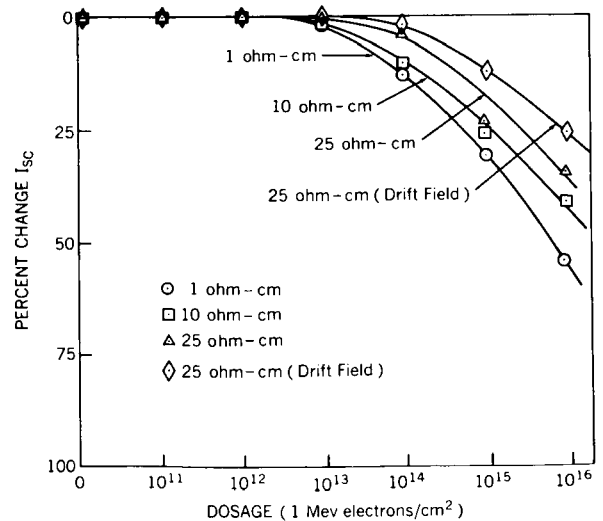


FIGURE 1.—Comparison of base resistivity with percent change in short circuit current caused by 1 Mev electrons.

Table 2
Short-Circuit Current Degradation of Irradiated Cells.

Base Resistivity (ohm-cm)	Cell Group	10^{16} Electrons/cm ²				3×10^{11} Protons/cm ²	
		Change in I_{sc} with Tungsten Light (percent)	Order of Radiation Resistance	Change in I_{sc} at Table Mountain (percent)	Order of Radiation Resistance	Change in I_{sc} with Tungsten Light (percent)	Order of Radiation Resistance
graded	J _{D/F}	22.0	*	No Cells	*	No Cells	*
graded	E _{D/F}	24.5	*	13	*	No Cells	*
25	E	34.5	1	25	1	30.5	2
10	F	40.3	2	28	2	29.8	1
10	G	41.1	3	31	3	32.0	3
1	B	50.6	4	33	4	35.7	4
1	C	51.8	5	39	5	37.0	6
1	D	52.5	6	40	7	37.5	7
<1	K	53.6	7	39	6	36.5	5
<1	H	56.4	8	44	8	40.8	8
1	A(P/N)	77.9	9	66	9	57.1	9

*The drift field cells are not listed since they were not completely tested.

Table 3
Peak Power Degradation of Irradiated Cells.

Base Resistivity (ohm-cm)	Cell Group	10^{16} Electrons/cm ²		3×10^{11} Protons/cm ²	
		Change in P_{mp} with Tungsten Light (percent)	Order of Radiation Resistance	Change in P_{mp} with Tungsten Light (percent)	Order of Radiation Resistance
graded	E _{D/F}	42.6	*	No Cells	*
graded	J _{D/F}	43.2	**	No Cells	*
25	E	48.1	1	37.9	2
10	F	52.1	2	35.2	1
10	G	52.6	3	39.3	3
1	C	58.7	4	41.3	5
1	D	59.2	5	42.2	6
1	B	59.7	6	41.0	4
<1	K	63.1	7	45.2	7
<1	H	63.6	8	46.9	8
1	A(P/N)	85.6	9	65.7	9

*The drift field cells are not listed since they were not completely tested.

resistivity is not as predominant a factor at 3×10^{11} protons/cm² as it is at 10^{16} electrons/cm². Perhaps an equivalent proton dosage, which would be about 10^{13} /cm², might show a correlation of base resistivity versus radiation damage. However, base resistivities of less than 1 ohm-cm are below the Group D reference all the way, while the 25 ohm-cm cells show less degradation of maximum power.

CONCLUSIONS

For cells at room temperature, we can conclude the following:

1. The 1 ohm-cm N/P silicon solar cell is conclusively more radiation-resistant than 1 ohm-cm P/N cells to 1 Mev electrons and/or 4.6 Mev protons in percentage change in both I_{sc} and maximum power output. This applies for both the 2800°K tungsten light and sunlight.

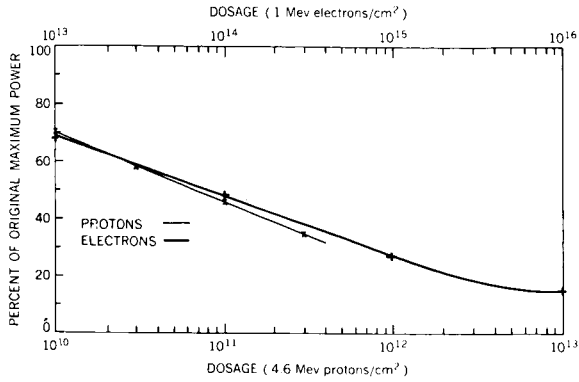


FIGURE 2.—Maximum power versus particle dosage for Group A solar cells.

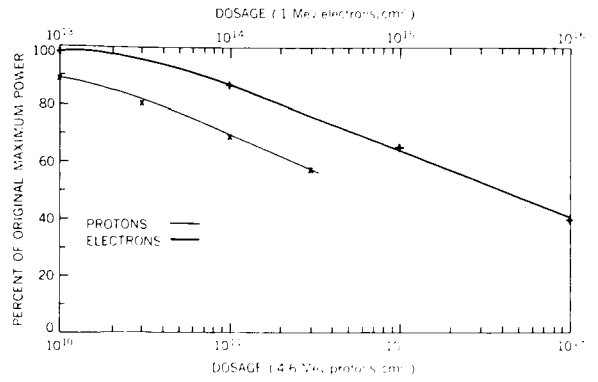


FIGURE 5.—Maximum power versus particle dosage for Group D solar cells.

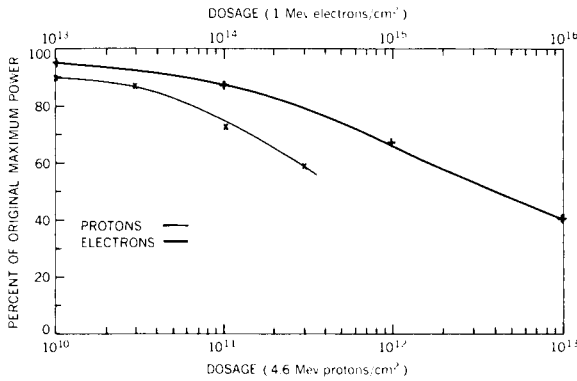


FIGURE 3.—Maximum power versus particle dosage for Group B solar cells.

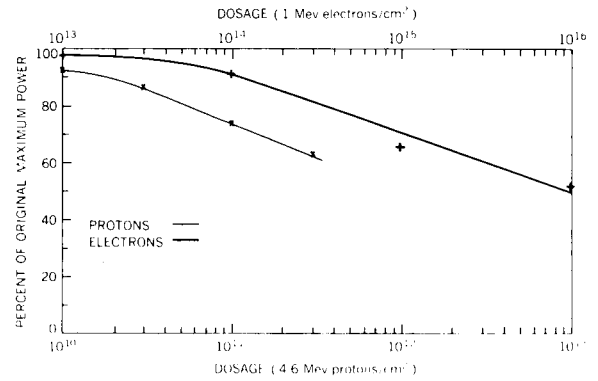


FIGURE 6.—Maximum power versus particle dosage for Group E solar cells.

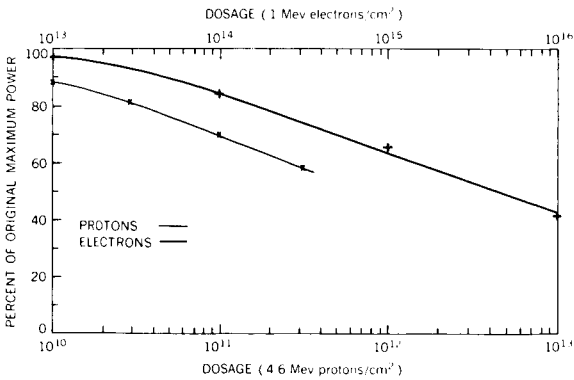


FIGURE 4.—Maximum power versus particle dosage for Group C solar cells.

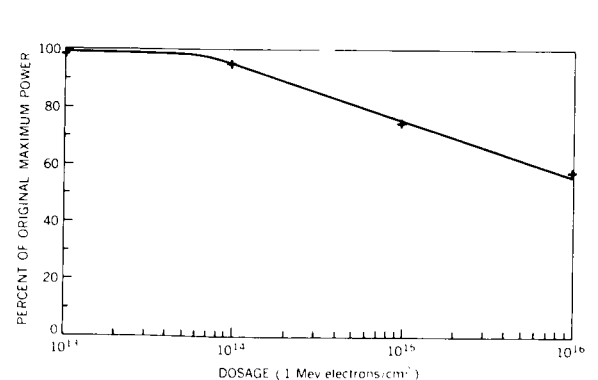


FIGURE 7.—Maximum power versus particle dosage for Group ED,F solar cells.

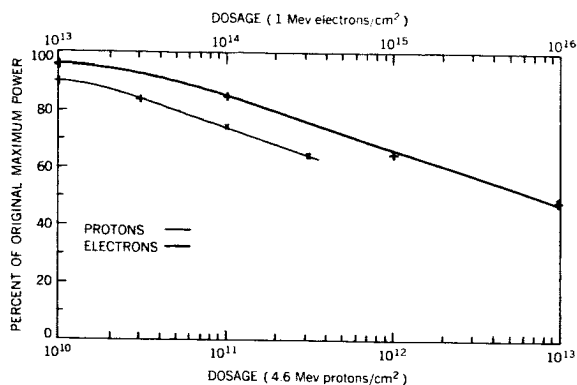


FIGURE 8.—Maximum power versus particle dosage for Group F solar cells.

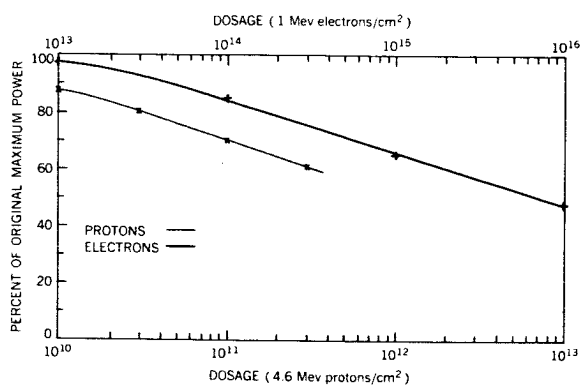


FIGURE 9.—Maximum power versus particle dosage for Group G solar cells.

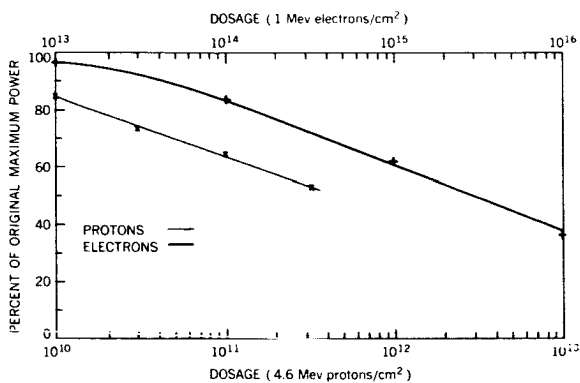


FIGURE 10.—Maximum power versus particle dosage for Group H solar cells.

2. The 1 ohm-cm N/P silicon solar cell is definitely more radiation-resistant to 4.6 Mev protons than cells made of lower base-resistivity material. There is an indication that the 1 ohm-

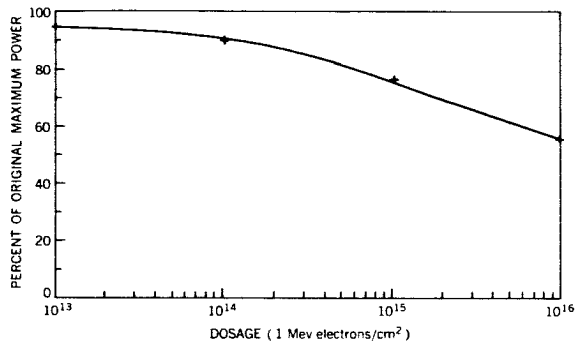


FIGURE 11.—Maximum power versus particle dosage for Group J_{D/F} solar cells.

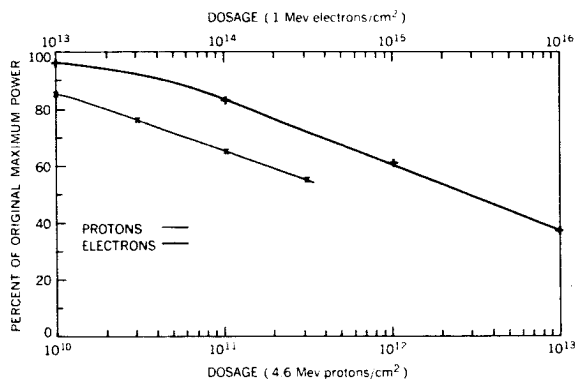


FIGURE 12.—Maximum power versus particle dosage for Group K solar cells.

cm cells are also more radiation-resistant to 1 Mev electrons than lower base-resistivity cells.

3. As base resistivity of N/P cells is increased beyond 1 ohm-cm, the radiation resistance to 1 Mev electrons definitely improves. This trend continues to at least 25 ohm-cm material. Radiation resistance of N/P cells to 4.6 Mev protons seems to improve to at least 10 ohm-cm base resistivity, but beyond that showed little or no improvement.

4. There is a positive indication that a silicon solar cell containing a built-in drift field will have greater radiation resistant to 1 Mev electrons than a silicon cell without such a field. No drift field cells were available for the proton experiment.

ACKNOWLEDGMENTS

The experiments were performed at the Naval Research Laboratory under NASA sponsorship. In particular, the outstanding work of Messrs. E. Brancato, R. Statler, J. Weller, and R. Lambert is acknowledged.

APPENDIX A

LIGHT SOURCE USED TO DETERMINE RADIATION DAMAGE TO SILICON SOLAR CELLS

A reflector flood tungsten bulb of 300 watts, 120 volt rating is color-temperature calibrated at 2800°K. The bulb is mounted so that the light path to the cell passes through a filter consisting of $\frac{1}{2}$ " (total thickness) of plexiglas plus 3 cm of de-ionized water. The filter size is at least 1 foot in diameter in order to reduce edge effects. All air bubbles adhering to the surfaces are removed so as to avoid local distortions.

Ten silicon solar cells, typical of the batch being evaluated, are calibrated in natural sunlight under acceptable atmospheric conditions:

1. Sun in clear sky and horizontal visibility at least 5 miles;
2. Measurements made between 11 am and 1 pm Local Standard Time, March 1 to October 30;
3. Temperature $25 \pm 2^\circ\text{C}$.

The light source is operated at the voltage necessary to attain a 2800°K color temperature at all times. The light intensity of 100 mw/cm² of equivalent natural sunlight is established by moving calibrated cells nearer or farther from the filter until I_{sc} conditions indicate the proper intensity. Measurements on all test cells will be made from this point. Precautions are taken to maintain cell temperature at $25^\circ\text{C} \pm 2^\circ$.

APPENDIX B

A COMPARATIVE STUDY OF 1 MEV ELECTRON-BOMBARDED SILICON SOLAR CELLS

INTRODUCTION

At the request of Goddard Space Flight Center, the Naval Research Laboratory conducted an experiment to determine the effects of 1 Mev electron radiation on the photon conversion properties of N/P silicon solar cells. The solar cells tested were from 10 different groups of recent production by undisclosed sources, which were purchased by GSFC and given to NRL without manufacturer identification. The radiation damage was studied after each radiation with 1 Mev electrons at doses of 10^{11} , 10^{12} , 10^{13} , 10^{14} , 10^{15} , and 10^{16} electrons/cm². Among the parameters examined were the minority carrier diffusion length, spectral response, output power, and short circuit current, although only the latter measurements will be reported herein.

INSTRUMENTATION

To assure uniformity of integrated electron flux to all cells, they were placed on an aluminum wheel rotating at 10 rpm under the beam tube of the Van de Graaff generator. The circular path of every cell was the same and passed through the axis of the beam tube about 6 inches below the exit window. Irradiation was done at different flux densities, and the time to reach each decade of dose varied from 2 minutes at 10^{11} e/cm² to 14 hours at 10^{16} e/cm². The deviation in uniformity of integrated flux from one cell to another was no greater than 0.25 percent at 10^{13} /cm² and was much less at the larger doses. The flux was calibrated by measuring the output of a radiation-damaged solar cell relative to a Faraday cup, and throughout the radiation the flux was monitored by a Faraday cup positioned near the wheel. Air was blown across the surface of the wheel for cooling the cells. The wheel temperature did not go above 30°C at any time during the irradiation.

The short circuit current of the solar cells was determined by measuring the voltage drop across a 1-ohm precision resistor loading the cell when it was illuminated. The light source was a 300-watt reflector flood bulb with a calibrated color temperature of 2800°K with a heat-absorbing filter consisting of 3 cm of distilled water contained in a plexiglas holder. The light intensity was adjusted by placing the cell holder at a distance from the lamp which corresponded to a 100 mw/cm² intensity as determined by a set of 10 Standard N/P solar cells which has been calibrated by GSFC and supplied to NRL for this purpose. The voltage across the 1-ohm loading resistor was measured by a voltage-to-frequency converter and displayed on a frequency counter. The accuracy of the short circuit current measurement is ± 0.1 ma for all cells except Group B, in which some contact resistance was evident. For these cells the accuracy is probably ± 0.3 ma.

RESULTS

Tables B1-B5 list all the measured values of short circuit current following the indicated electron dose. The average value of short circuit current for each group of cells at the indicated dose is shown in Table B6. It should be noted that not all cells were measured at every dose, so that the average current at each dose is for a different set of cells of the group. The percentage change in average short-circuit current is shown in Table B7. This change is computed from the average change in current for the cells measured at a given dose level relative to the initial average current of that same set of cells.

A more complete report describing this and other phases of the experiment in more detail is in preparation.

Table B1
Groups A and B Solar Cell Short-Circuit Current After 1 Mev Electron Radiation.

Cell Number	Initial Current (ma)	Current After Dosage (ma)					
		10^{11} (electrons/cm ²)	10^{12} (electrons/cm ²)	10^{13} (electrons/cm ²)	10^{14} (electrons/cm ²)	10^{15} (electrons/cm ²)	10^{16} (electrons/cm ²)
A-1-1	48.1						
A-1-2	47.3	47.3					
A-1-3	47.2		44.6				
A-1-4	48.6		45.4	37.0			
A-1-5	48.4		44.0	34.7	24.8		
A-1-6	47.4			34.3	24.1	15.0	
A-1-7	48.1			35.8	26.2	16.6	10.3
A-1-8	47.5			34.7	24.8	15.3	9.4
A-1-9	47.7			38.1	28.5	18.7	11.7
A-1-10	47.9			36.8	26.9	17.3	11.0
A-1-11	48.0			36.5	26.9	17.4	11.0
A-1-12	47.9			35.0	25.1	15.8	9.8
A-1-13	47.8			36.0	26.2	16.7	10.2
A-1-14	48.6			37.3	27.5	18.0	11.1
A-1-15	47.8			37.6	27.4	17.9	11.1
B-1-1	70.4						
B-1-2	78.6	78.9					
B-1-3	79.5		78.5				
B-1-4	76.4		76.4	74.6			
B-1-5	75.2		75.1	74.2	70.1		
B-1-6	81.6			79.9	74.5	60.4	
B-1-7	76.4			75.7	71.1	57.5	38.5
B-1-8	77.2			76.6	71.0	57.5	38.2
B-1-9	85.0			81.9	75.9	60.2	39.8
B-1-10	78.2			77.6	72.1	58.8	39.6
B-1-11	80.1			79.0	73.1	59.1	41.5
B-1-12	78.3			77.3	70.4	55.2	35.6
B-1-13	80.5			79.9	74.2	60.2	41.0
B-1-14	83.0			81.7	76.0	61.6	42.9
B-1-15	82.1			79.9	73.5	58.2	39.1

Table B2

Groups C and D Solar Cell Short-Circuit Current After 1 Mev Electron Radiation.

Cell Number	Initial Current (ma)	Current After Dosage (ma)										
		10^{11} (electrons/cm ²)	10^{12} (electrons/cm ²)	10^{13} (electrons/cm ²)	10^{14} (electrons/cm ²)	10^{15} (electrons/cm ²)	10^{16} (electrons/cm ²)					
C-1-1	51.5	51.3										
C-1-2	51.4											
C-1-3	51.4							50.9				
C-1-4	52.0							51.2	50.8			
C-1-5	50.5							50.3	49.4	43.9		
C-1-6	51.6								51.0	45.5	36.4	
C-1-7	51.9								51.1	45.3	36.1	24.8
C-1-8	52.2								51.3	45.2	36.0	24.9
C-1-9	51.4								50.7	45.2	36.1	25.1
C-1-10	51.2								50.3	44.9	35.7	24.6
C-1-11	52.0								50.5	45.0	35.9	25.1
C-1-12	50.7								50.0	44.8	35.8	24.8
C-1-13	51.1								50.4	44.9	35.8	24.4
C-1-14	51.0								49.6	44.6	35.3	25.0
C-1-15	51.2								50.1	44.6	35.6	25.0
D-1-1	45.1	45.4										
D-1-2	45.6											
D-1-3	45.0							45.0				
D-1-4	45.0							44.9	44.5			
D-1-5	44.4							44.5	44.1	39.6		
D-1-6	45.1								44.4	40.8	32.4	
D-1-7	44.4								44.0	39.6	31.3	21.0
D-1-8	45.4								45.2	40.9	32.5	22.0
D-1-9	45.6								45.2	40.4	32.0	21.3
D-1-10	45.1								44.5	40.8	32.6	22.2
D-1-11	45.1								44.9	40.4	32.1	21.6
D-1-12	45.2								44.7	40.1	31.9	21.3
D-1-13	45.1								44.4	40.0	31.5	21.2
D-1-14	44.4								43.9	39.7	31.2	20.9
D-1-15	45.2								44.8	40.5	31.9	21.5

Table B3

Groups E, E_{D/F} and J_{D/F} Solar Cell Short-Circuit Current After 1 Mev Electron Radiation.

Cell Number	Initial Current (ma)	Current After Dosage (ma)					
		10 ¹¹ (electrons/cm ²)	10 ¹² (electrons/cm ²)	10 ¹³ (electrons/cm ²)	10 ¹⁴ (electrons/cm ²)	10 ¹⁵ (electrons/cm ²)	10 ¹⁶ (electrons/cm ²)
E-1-1	22.0						
E-1-4	22.6	22.7					
E-1-5	22.0		22.2				
E-1-8	21.5		21.6	21.7			
E-1-10	21.9		22.0	22.0	21.0		
E-1-11	21.8			21.9	21.0	17.0	
E-1-13	21.9			21.8	20.8	16.9	14.3
E-1-15	22.4			22.5	21.4	17.4	14.8
E-1-18	21.3			21.3	20.4	16.5	13.9
E-1-19	22.1			22.2	21.1	17.1	14.5
E _{D/F} -1-2	20.2						
E _{D/F} -1-3	19.2	19.2					
E _{D/F} -1-6	21.7		21.8				
E _{D/F} -1-7	20.5		20.7	20.7			
E _{D/F} -1-9	19.4		19.5	19.7	19.2		
E _{D/F} -1-12	19.8			19.9	19.6	17.9	
E _{D/F} -1-14	21.0			21.1	20.7	18.8	15.7
E _{D/F} -1-16	21.5			21.6	21.3	18.9	16.2
E _{D/F} -1-17	20.5			20.7	20.3	18.4	15.3
E _{D/F} -1-20	19.8			19.8	19.4	17.6	14.4
J _{D/F} -1-1	43.9						
J _{D/F} -1-2	44.0		44.1				
J _{D/F} -1-3	44.5		44.6	44.5			
J _{D/F} -1-4	46.1		46.2	45.9	45.8		
J _{D/F} -1-5	42.7		42.9	42.7	42.3	37.5	
J _{D/F} -1-6	45.5			45.8	45.0	40.9	34.7

Table B4

Groups F and G Solar Cell Short-Circuit Current After 1 Mev Electron Radiation.

Cell Number	Initial Current (ma)	Current After Dosage (ma)					
		10^{11} (electrons/cm ²)	10^{12} (electrons/cm ²)	10^{13} (electrons/cm ²)	10^{14} (electrons/cm ²)	10^{15} (electrons/cm ²)	10^{16} (electrons/cm ²)
F-1-1	52.5						
F-1-2	50.7	50.8					
F-1-3	52.1		52.0				
F-1-4	52.3		52.1	51.4			
F-1-5	52.2		52.1	51.2	47.3		
F-1-6	52.0			51.2	47.1	38.8	
F-1-7	51.5			50.1	45.9	37.1	28.0
F-1-8	51.9			50.7	46.9	39.0	31.7
F-1-9	51.5			50.5	47.1	39.0	31.8
F-1-10	52.4			51.7	47.9	39.4	32.2
F-1-11	51.8			51.0	47.3	38.7	31.2
F-1-12	52.7			51.8	48.0	39.6	31.9
F-1-13	52.4			51.8	47.8	40.1	32.3
F-1-14	51.6			50.8	46.8	39.5	31.7
F-1-15	50.0			48.9	44.5	36.7	27.7
G-1-1	50.3						
G-1-2	50.2	50.4					
G-1-3	49.7		49.7				
G-1-4	47.7		47.5	46.6			
G-1-5	46.8		46.6	45.8	41.5		
G-1-6	50.6			49.6	44.9	37.1	
G-1-7	48.5			47.4	43.2	35.8	28.7
G-1-8	49.1			48.0	44.1	36.4	29.2
G-1-9	48.2			47.2	43.1	35.8	28.1
G-1-10	49.5			48.5	44.2	36.4	29.0
G-1-11	49.5			48.5	44.1	36.6	29.7
G-1-12	50.0			49.0	44.3	37.0	29.7
G-1-13	48.5			47.4	42.9	35.7	27.8
G-1-14	50.1			49.4	44.5	37.2	30.0
G-1-15	49.4			48.3	44.3	36.4	28.5

Table B5

Groups H and K Solar Cell Short-Circuit Current After 1 Mev Electron Radiation.

Cell Number	Initial Current (ma)	Current After Dosage (ma)					
		10^{11} (electrons/cm ²)	10^{12} (electrons/cm ²)	10^{13} (electrons/cm ²)	10^{14} (electrons/cm ²)	10^{15} (electrons/cm ²)	10^{16} (electrons/cm ²)
H-1-1	50.1						
H-1-2	46.8	47.0					
H-1-3	48.5		48.0				
H-1-4	50.3		50.1	48.5			
H-1-5	48.0		47.5	46.0	41.3		
H-1-6	48.3			47.0	42.1	32.9	
H-1-7	50.0			48.5	42.8	33.5	21.7
H-1-8	49.2			48.1	42.4	33.0	21.3
H-1-9	49.5			48.3	42.7	33.3	21.7
H-1-10	49.0			47.6	42.3	32.8	21.3
H-1-11	47.5			46.3	40.8	31.7	21.0
H-1-12	48.0			47.0	42.0	32.8	21.5
H-1-13	50.1			48.5	42.8	33.0	21.6
H-1-14	48.8			47.4	42.1	32.6	21.4
H-1-15	49.9			48.1	42.4	32.8	21.2
K-1-1	47.9						
K-1-2	49.3	49.7					
K-1-3	47.6		47.4				
K-1-4	47.7		47.6	46.5			
K-1-5	47.4		47.4	46.5	41.3		
K-1-6	48.0			46.8	41.2	32.4	
K-1-7	49.1			47.7	42.3	33.2	22.6
K-1-8	48.7			47.4	41.7	32.5	22.1
K-1-9	45.6			44.2	38.8	30.0	19.6
K-1-10	49.3			47.7	41.9	32.8	22.2
K-1-11	48.1			46.8	41.4	32.6	22.4
K-1-12	47.7			46.3	41.0	32.4	22.1
K-1-13	49.4			48.8	43.9	35.7	27.1
K-1-14	46.6			45.5	40.2	30.9	20.3
K-1-15	48.4			47.1	41.4	32.4	22.0

Table B6
Group Average Short-Circuit Current of Solar Cells After 1 Mev Electron Radiation.

Cell Group	Initial Current (ma)	Current After Dosage (ma)			
		10^{13} (electrons/cm ²)	10^{14} (electrons/cm ²)	10^{15} (electrons/cm ²)	10^{16} (electrons/cm ²)
A	47.9	36.2	26.2	16.9	10.6
B	78.8	78.2	72.9	58.9	39.6
C	51.4	50.4	44.9	35.9	24.8
D	45.0	44.6	40.3	31.9	21.4
E*	22.0	21.9	21.0	17.0	14.4
E _{D/F} *	20.4	20.5	20.1	18.3	15.4
F	51.8	50.9	47.0	38.8	30.9
G	49.2	48.0	43.7	36.4	29.0
H	48.9	47.6	42.2	32.8	21.4
J _{D/F} *	44.5	44.7	44.4	39.2	34.7
K	48.1	46.8	41.4	32.5	22.3

*The E, E_{D/F} and J_{D/F} group averages are based on a smaller number of cells, the number of which decreases with increasing radiation dose (see Table B3).

Table B7
Percentage Change in Average Short-Circuit Current of Solar Cells After 1 Mev Electron Radiation.

Cell Group	Change After Dosage (percent)			
	10^{13} (electrons/cm ²)	10^{14} (electrons/cm ²)	10^{15} (electrons/cm ²)	10^{16} (electrons/cm ²)
A	24.6	45.3	64.7	77.9
B	1.6	8.6	25.6	50.6
C	1.9	12.5	30.2	51.8
D	0.9	10.4	29.3	52.5
E*	0.5	4.5	22.6	34.5
E _{D/F} *	0.0	1.5	10.3	24.5
F	1.9	9.3	25.1	40.3
G	2.0	11.0	26.2	41.1
H	3.1	13.7	33.1	56.4
J _{D/F} *	0.0	0.2	12.0	22.0
K	2.5	13.7	32.4	53.6

*The E, E_{D/F} and J_{D/F} group averages are based on a smaller number of cells, the number of which decreases with increasing radiation dose (see Table B3).

Table B8

Percentage Change in Average Maximum Power of Solar Cells After 1 Mev Electron Radiation.

Cell Group	Change After Dosage (percent)			
	10^{13} (electrons/cm ²)	10^{14} (electrons/cm ²)	10^{15} (electrons/cm ²)	10^{16} (electrons/cm ²)
A	29.7	52.3	72.4	85.6
B	5.2	12.4	33.1	59.7
C	3.0	15.2	35.4	58.7
D	1.4	12.9	34.8	59.2
E*	1.8	8.6	34.2	48.1
E _{D/F} *	0.4	4.1	25.0	42.6
F	3.3	14.1	34.2	52.1
G	2.0	15.1	35.2	52.6
H	3.4	16.5	38.6	63.6
J _{D/F} *	5.0	9.4	23.2	43.2
K	3.5	17.3	39.3	63.1

*The E, E_{D/F} and J_{D/F} group averages are based on a smaller number of cells, the number of which decreases with increasing radiation dose (see Table B3).

APPENDIX C

A COMPARATIVE STUDY OF 4.6 MEV PROTON-BOMBARDED SILICON SOLAR CELLS

INTRODUCTION

A comparative study was made of the radiation damage rates of silicon solar cells from 4.6 Mev protons as a companion experiment to a 1 Mev electron damage study recently done at the Naval Research Laboratory.* The solar cells used in both experiments were supplied by Goddard Space Flight Center from lots of experimental cells and typical production quality cells of late 1962, which had been procured for this purpose. The proton energy of 4.6 Mev was low enough to produce a high damage rate throughout the bulk material of the cell to a depth exceeding the minority carrier diffusion length before irradiation. The radiation damage was measured after each proton bombardment at doses of 10^{10} , 3×10^{10} , 10^{11} , and 3×10^{11} protons/cm². Included in the measured characteristics were the minority carrier diffusion length, the spectral response, and the voltage-current characteristics of the cells under filtered tungsten light. Only the results of the short circuit current and maximum power measurements will be reported herein.

RADIATION

The NRL 5 Mev Van de Graaff accelerator was used for the source of protons. The beam was scattered by a .0004-inch gold foil placed about 18 feet from the end of the drift tube where the solar cells were placed. The foil served the dual purpose of reducing the beam intensity and providing a uniform flux distribution at the cells. The proton energy incident on the foil was measured as 5.127 Mev, while the energy incident on the solar cells was calculated to be 4.657 Mev, because of energy absorption by the foil. The flux was measured by collecting all of the beam current passing through a defining aperture into the insulated end section of the tube, which acted

as a Faraday cup. The flux density was about 1.4×10^8 protons/cm². The flux variation over the end of the beam tube where the solar cells were placed was ± 1 percent. The variation in integrated flux from one radiation to another was within ± 1.3 percent. There were 9 groups of 6 cells each which were irradiated 4 at a time by permuting the cells types and positions for each exposure. The pressure in the tube was about 5×10^{-5} mm-Hg.

MEASUREMENTS

The voltage-current characteristics of the cells were measured under the same light source used for the electron damage study.* This was a 300-watt reflector flood bulb calibrated to a color temperature of 2800°K with a heat absorbing filter consisting of 3 cm of deionized water in a plexiglas holder interposed before the sample cells. The light intensity was adjusted by moving the cell holder to a position from the lamp which gave a short-circuit current from a standard cell equivalent to 100 mw/cm² solar intensity at air mass one. This was determined by a set of 10 Standard N/P solar cells which had been calibrated by the Goddard Space Flight Center for this purpose. The short-circuit current was measured by precision resistors and voltage-to-frequency conversion with an overall accuracy of ± 0.1 ma for all cells except in Group B, where the accuracy is probably ± 0.3 ma (see Appendix B).

RESULTS AND CONCLUSIONS

Tables C1-C3 present all the measured values of short-circuit current following the indicated proton dose. The average value of short-circuit current for each group of cells at the indicated dose is shown in Table C4. The percentage change in the average short-circuit current is given in Table C5. Table C6 gives the percentage change in maximum power output of each group as a function of proton dose.

*Radiation Damage Symposium, Goddard Space Flight Center, January 23, 1963.

Table C1
Groups A, B, C Solar Cell Short-Circuit Current After 4.6 Mev Proton Radiation.

Cell Number	Initial Current (ma)	Current After Dosage (ma)			
		10^{10} (protons/cm ²)	3×10^{10} (protons/cm ²)	10^{11} (protons/cm ²)	3×10^{11} (protons/cm ²)
A-1-16	47.5	36.9	31.8	26.0	20.9
A-1-18	49.1	36.9	32.1	26.0	20.9
A-1-20	48.1	35.1	30.6	24.5	19.6
A-1-22	48.6	37.0	32.6	27.0	21.4
A-1-23	47.6	35.7	31.1	25.4	20.0
A-1-24	47.3	36.2	31.7	25.9	20.4
B-1-16	81.7	77.5	70.4	61.6	51.7
B-1-17	82.8	78.8	72.0	63.7	54.2
B-1-18	81.6	76.9	71.6	62.6	53.5
B-1-19	78.6	74.5	69.0	59.4	
B-1-20	83.4	78.4	71.9	62.3	52.7
B-1-21	79.6	75.4	70.0	61.0	51.7
C-1-16	50.7	45.5	41.6	36.7	31.9
C-1-19	51.5	45.2	41.7	36.6	31.5
C-1-20	49.2	43.3	39.8	35.1	30.2
C-1-21	50.0	44.6	41.2	36.8	31.5
C-1-22	47.2	42.9	40.0	35.6	30.5
C-1-24	49.0	44.8	42.0	37.3	31.8

The radiation damage rates, indicated by the percentage change in short-circuit current, follow the same trend as shown in the 1 Mev electron damage study; the higher bulk-resistivity cells have the lowest damage rate, and the P/N cells have the highest rate. There are two variations in the order of damage rates from electrons compared to protons. Group E, which has a significantly lower damage rate under electron radiation, had the same rate as the 10 ohm-cm cells under protons. Group K has a lower damage rate from protons than groups C and D. However, this amounts to only one percentage point difference and could be attributed to small variations in the average initial properties of the cell group. It is conceivable, for example, that one group of 6 cells from a particular lot may show

a difference of 1 percent in radiation resistance compared to another group of 6 from the same lot. The data were analyzed by applying the "Students-t-Test", and difference between means was found to be significant to better than .01 level; which means that there is less than 1 chance out of 100 that the differences are due to random errors. The validity of using the filtered tungsten light for measuring I_{sc} was established by measuring the electron-damaged cells at Table Mountain and under a solar simulator. The relative order of radiation damage rate was unchanged when measuring under these three sources, except for Group K at Table Mountain.*

*"One Mev Electron Damage in Silicon Solar Cells," by R. L. Statler-Photovoltaic Specialists Conference of IEEE, Washington, D.C., April 10, 1963.

Table C2

Groups D, E, F Solar Cell Short-Circuit Current after 4.6 Mev Proton Radiation.

Cell Number	Initial Current (ma)	Current After Dosage (ma)			
		10^{10} (protons/cm ²)	3×10^{10} (protons/cm ²)	10^{11} (protons/cm ²)	3×10^{11} (protons/cm ²)
D-1-16	44.7	40.0	36.4	31.9	27.3
D-1-17	45.9	41.1	38.1	33.4	28.8
D-1-20	44.5	39.9	36.7	32.2	27.7
D-1-21	44.7	40.2	37.2	32.7	27.8
D-1-24	46.0	40.9	37.9	33.4	28.4
D-1-25	44.7	40.4	38.1	33.9	29.2
E-1-21	21.6	20.7	19.2	17.0	15.1
E-1-22	21.7	20.5	19.4	17.1	15.3
E-1-23	21.7	20.4	19.3	17.0	15.1
E-1-24	22.4	21.0	*	17.3	15.3
E-1-25	22.1	20.1	*	17.2	15.3
E-1-26	21.8	20.0	*	17.2	15.1
F-1-16	52.7	48.4	44.7	40.3	36.2
F-1-19	52.1	48.0	45.1	40.7	36.6
F-1-22	52.2	48.0	45.1	41.2	37.3
F-1-23	51.6	46.6	44.0	39.8	35.6
F-1-24	51.8	47.7	45.2	41.4	37.2
F-1-30	52.0	46.7	†	†	†

*Measurements were inadvertently omitted.

†Cell was broken.

Table C3

Groups G, H, K Solar Cell Short-Circuit Current after 4.6 Mev Proton Radiation.

Cell Number	Initial Current (ma)	Current After Dosage (ma)			
		10^{10} (protons/cm ²)	3×10^{10} (protons/cm ²)	10^{11} (protons/cm ²)	3×10^{11} (protons/cm ²)
G-1-18	46.2	41.3	37.7	34.4	30.5
G-1-20	49.1	44.5	41.8	38.0	34.3
G-1-21	50.5	45.7	42.6	38.7	34.6
G-1-22	49.6	45.1	42.3	38.5	34.3
G-1-25	46.9	42.1	39.6	35.8	31.7
G-1-26	48.0	43.2	40.5	36.5	32.3
H-1-16	49.9	43.1	38.5	34.0	28.8
H-1-17	52.6	45.1	41.4	36.0	30.7
H-1-18	47.3	41.7	38.3	33.7	28.6
H-1-19	49.2	43.2	39.9	34.7	29.5
H-1-21	49.5	43.6	40.3	35.2	30.0
H-1-22	49.3	42.8	39.4	34.0	28.8
K-1-16	47.1	42.0	38.2	33.9	29.5
K-1-17	49.9	45.1	41.9	37.6	33.2
K-1-18	49.5	44.6	41.5	37.4	32.9
K-1-20	49.7	44.5	41.6	37.2	32.6
K-1-24	48.2	41.5	38.4	33.6	28.7
K-1-26	47.8	41.9	38.8	34.0	29.0

Table C4
Average Short-Circuit Current of Solar Cells After 4.6 Mev Proton Radiation.

Cell Group	Initial Current (ma)	Current After Dosage (ma)			
		10^{10} (protons/cm ²)	3×10^{10} (protons/cm ²)	10^{11} (protons/cm ²)	3×10^{11} (protons/cm ²)
A	48.0	36.3	31.7	25.8	20.5
B	81.3	76.9	70.8	61.8	52.8*
C	49.6	44.4	41.0	36.4	31.2
D	45.1	40.4	37.4	32.9	28.2
E	21.9	20.5	19.3*	17.1	15.2
F	52.1	47.6	44.8*	40.7*	36.6*
G	48.4	43.7	40.8	37.0	33.0
H	49.6	43.3	39.6	34.6	29.4
K	48.7	43.3	40.1	35.6	31.0

*Less than 6 cells averaged for this value.

Table C5
Percentage Change in Average Short-Circuit Current of Solar Cells After 4.6 Mev Proton Radiation.

Cell Group	Change After Dosage (percent)			
	10^{10} (protons/cm ²)	3×10^{10} (protons/cm ²)	10^{11} (protons/cm ²)	3×10^{11} (protons/cm ²)
A	24.4	34.1	46.0	57.1
B	5.4	12.9	24.0	35.7
C	10.5	17.2	26.7	37.0
D	10.4	17.1	27.0	37.5
E	6.5	10.9	21.6	30.5
F	8.7	13.9	21.7	29.8
G	9.8	15.8	23.6	32.0
H	12.8	20.1	30.3	40.8
K	11.2	17.8	26.9	36.5

Table C6
Percentage Change in Average Maximum Power of Solar Cells After 4.6 Mev Proton Radiation.

Cell Group	Change After Dosage (percent)			
	10^{10} (protons/cm ²)	3×10^{10} (protons/cm ²)	10^{11} (protons/cm ²)	3×10^{11} (protons/cm ²)
A	29.3	41.4	53.9	65.7
B	4.9	12.9	27.2	41.0
C	11.7	18.5	30.2	41.3
D	10.9	19.2	30.3	42.2
E	7.4	13.4	26.5	37.9
F	9.5	15.6	25.7	35.2
G	11.5	19.7	30.1	39.3
H	15.6	23.8	36.2	46.9
K	14.7	23.9	35.0	45.2

CODING AND ANALOG VARIABLE FOR CONSTANT PERCENTAGE ERROR*

RODGER A. CLIFF

Systems are treated which code an analog variable as a sequence of discrete values. Given one of these discrete values, there is necessarily some uncertainty about what value of the analog input variable produced the discrete output since there are only a finite number of distinct outputs. This paper will investigate the nature of this uncertainty or possible error.

It is shown that a logarithmic quantization scheme produces an uncertainty in the output which is constant over the input range which the system is designed to cover. Relationships are derived between the uncertainty which must be tolerated, the width of the input range to be covered, and the number of discrete outputs required. These relationships are presented both analytically and graphically.

INTRODUCTION

An analog variable is one which may take on any value within a continuous range. In order to operate upon such a variable in a discrete device, such as a digital computer or a digital data transmission link, it is necessary to convert the analog variable to discrete form. In other words, the analog variable, which may have any of an infinite number of values, must be mapped onto a finite set of discrete values. This is shown schematically in Figure 1 and graphically in Figure 2. The device which accomplishes the mapping is referred to as an analog to digital converter. It repeatedly samples the analog to digital converter. It repeatedly samples the analog input x and produces a discrete output y for each of these samples.



FIGURE 1.—Conversion of an analog variable to digital form.

*Published as NASA Technical Note D-2257, March 1964.

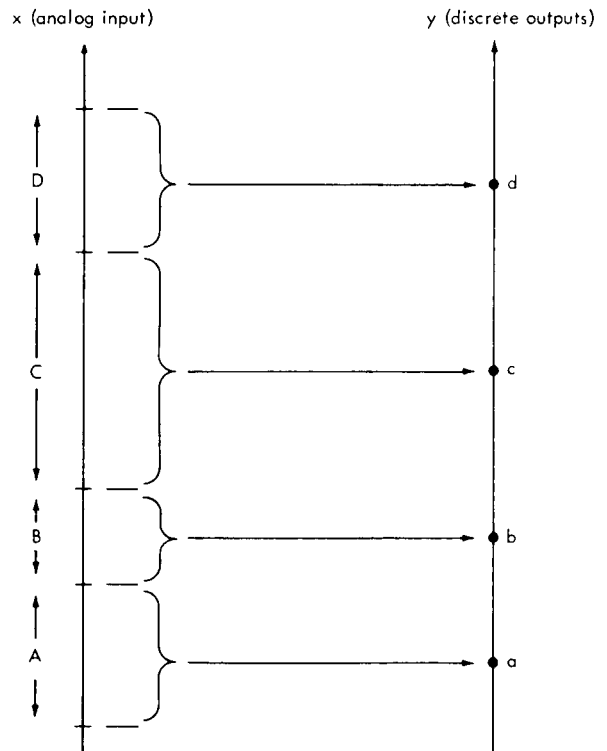


FIGURE 2.—Arbitrary quantization scheme.

The analog variable is divided into intervals denoted A,B,C,... in Figure 2. An input x which falls within a given interval produces an output y corresponding to that interval. For instance,

if the input x has a value within the interval B the analog to digital converter produces an output y having the discrete value b . Similarly, inputs within interval A produce an output a etc.

There are various instances when it is necessary to convert a continuous variable into digital form: for example, analog input to a digital computer, or digital telemetry of continuous information. It is often desired that the possible percentage error associated with each discrete output be less than a certain amount. A logarithmic quantization is used to accomplish this end.

In this paper the problem of logarithmic quantization of a continuous input into a finite number of outputs is discussed. Results are presented showing the necessary number of discrete outputs for a given dynamic range and percentage error. For those interested in binary systems, a discussion of the number of bits required to express with a given accuracy an input variable of a given dynamic range is included.

LOGARITHMIC QUANTIZATION

Derivation of the Basic Relationships

A derivation of the relationship between the number of discrete outputs, the dynamic range, and the percentage error will be given. The relationship between the input x and the output y of a logarithmic quantization is depicted in Figure 3 on linear axes, and in Figure 4 on log axes. Let an arbitrary quantization interval X range from a lower value x^- to an upper value x^+ . The output y_x corresponding to this interval is chosen to have a value between x^- and x^+ . The choice is made such that the possible error is minimized.

The maximum possible positive percentage error may be expressed as

$$Q^+ = 100 \frac{x^+ - y_x}{y_x} \quad (1a)$$

and the maximum possible negative percentage error as

$$Q^- = 100 \frac{y_x - x^-}{y_x} \quad (1b)$$

In order to obtain a symmetrical error bracket about y_x , we will choose y_x such that

$$Q^+ = Q^- = Q \quad (1c)$$

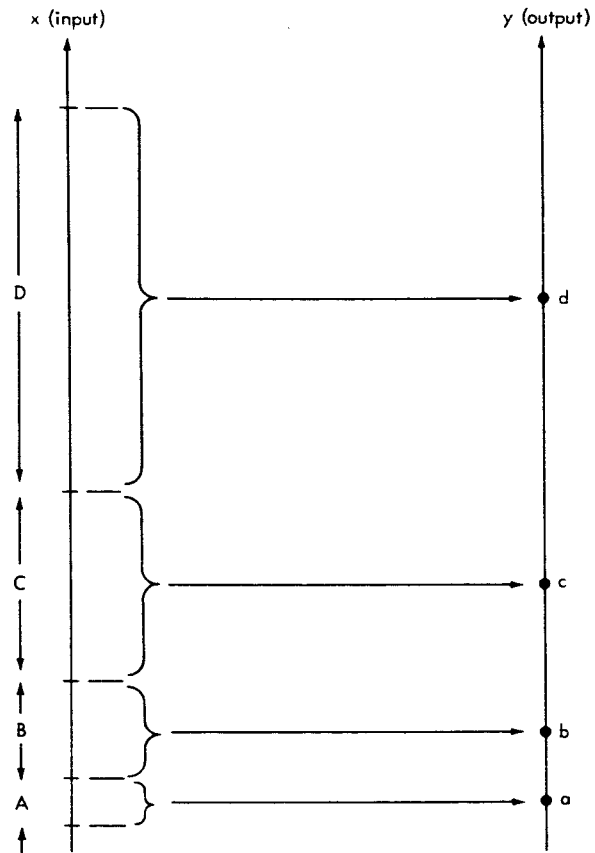


FIGURE 3.—Logarithmic quantization scheme—linear axes.

where Q is the maximum possible \pm percentage error. The three Equations 1 may now be solved for y_x to obtain

$$y_x = \frac{x^+ + x^-}{2} \quad (2)$$

It is not surprising that the optimum value for y_x is found to be halfway between x^- and x^+ ; certainly it is the conclusion we reach intuitively.

Equations 1 may also be solved for x^- and x^+ .

$$x^+ = \left(1 + \frac{Q}{100}\right) y_x \quad (3a)$$

$$x^- = \left(1 - \frac{Q}{100}\right) y_x \quad (3b)$$

Observe that the ratio x^+/x^- of the upper limit to the lower limit of the quantization interval X is

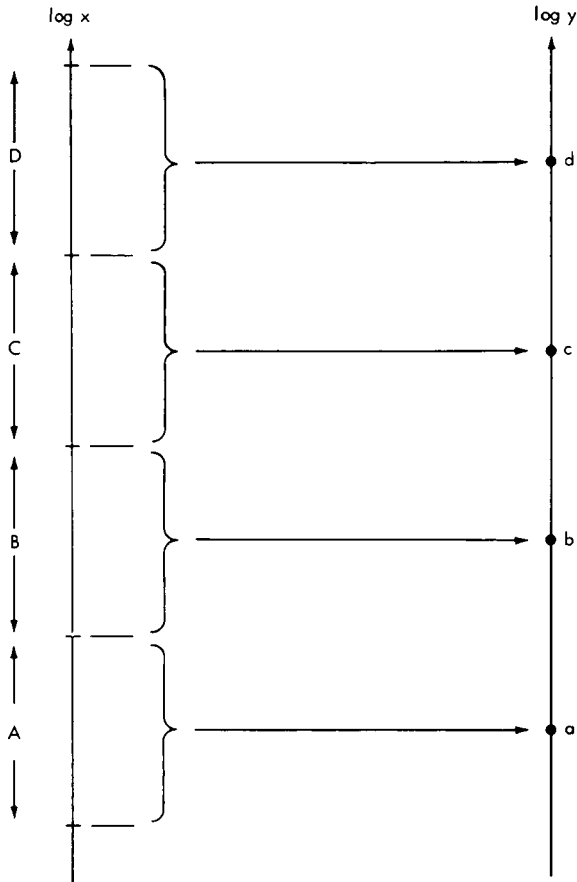


FIGURE 4.—Logarithmic quantization scheme — log axes.

independent of which interval X may be. This ratio shall be denoted by K:

$$K = \frac{1 + \frac{Q}{100}}{1 - \frac{Q}{100}} = \frac{x^+}{x^-} \quad (4)$$

Equation 4 says that for any interval X, $x^+ = Kx^-$; therefore the levels that separate intervals are expressible in terms of the lower limit of the lowest interval, as shown in Figure 5.

The dynamic range R is defined to be the ratio of the largest possible value of the input variable x to the smallest possible value of the input variable x that fall within the range to be quantized. From Figure 5 we see that

$$R = \frac{K^N x_0}{x_0} = K^N \quad (5)$$

where N is the number of quantization intervals.

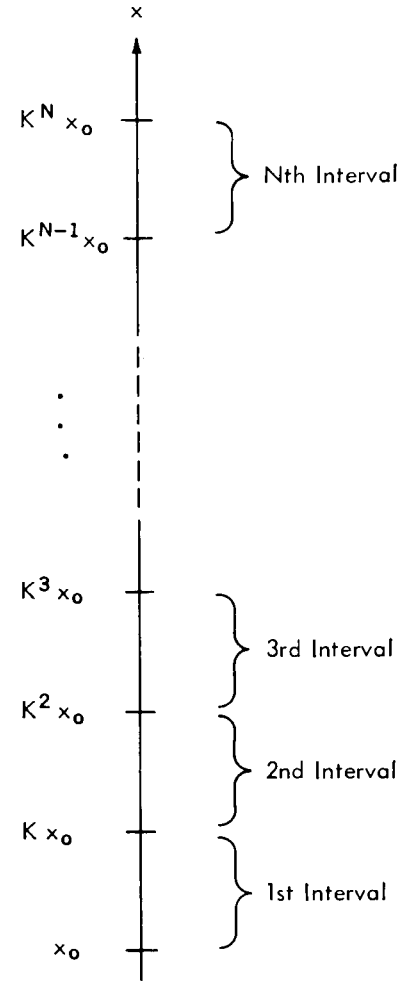


FIGURE 5.—Decision levels for a logarithmic quantization.

To obtain N as a function of K (and therefore as a function of Q), we take the logarithms of both sides of Equation 5:

$$\log R = N \log K \quad (6)$$

or

$$N = \frac{\log R}{\log K} \quad (7)$$

which in terms of Q is

$$N = \frac{\log R}{\log \left(\frac{1 + \frac{Q}{100}}{1 - \frac{Q}{100}} \right)} \quad (8)$$

This is the result we require, relating N (the number of quantization intervals required) to R (the

dynamic range) and Q (the maximum possible \pm percentage error).

The Number of Bits Needed for a Binary System

A binary code of B bits can specify that one out of 2^B possible events has occurred. In the system under discussion, a particular event occurs when the input variable falls within a particular one of the N quantization intervals. Therefore, there must be at least as many characters in the binary code as there are quantization intervals. That is

$$2^B \geq N, \tag{9}$$

so that

$$B \geq \log_2 N. \tag{10}$$

It is advantageous that N be a power of 2; then a binary code may be chosen such that each character is used and full advantage is taken of the capacity of the binary system. In this case there are N quantization intervals and N characters in the binary code. The relationship (Equation 10) will hold as an equality.

APPLICATION OF THE RESULTS OF THE DERIVATION

Graphs of the Basic Equation

To facilitate system design, the basic relationship of the logarithmic quantization scheme (Equation 8) has been plotted. Figure 6 shows the trade-off between the dynamic range and the

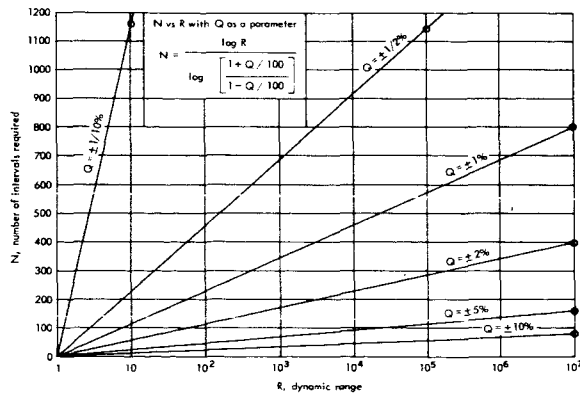


FIGURE 6.—The trade-off between the dynamic range and number of intervals required.

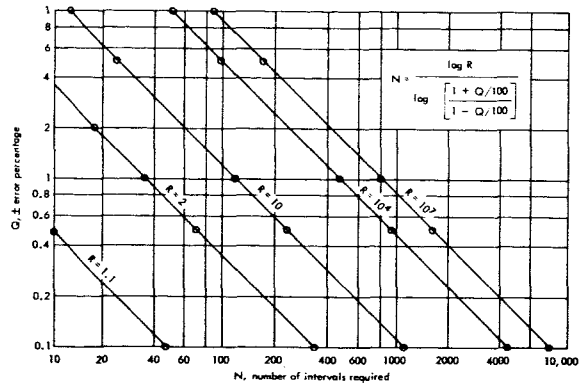


FIGURE 7.—Relationship between number of intervals required and percentage error.

number of intervals required for a given percentage error. Since the number of intervals N varies as the logarithm of the dynamic range R for fixed percentage error Q , the equation plots as a family of straight lines on semi-log coordinates.

Figure 7, on the other hand, emphasizes the relationship between the number of intervals needed and the percentage error. It may be shown that N varies approximately as $1/Q$ (see Appendix A); hence contours of constant R are straight lines on full logarithmic coordinates.

General Remarks

A few general remarks are in order at this point. First, notice that for the condition where the percentage error is greater than about 5 percent, the number of quantization intervals required for a given range is relatively insensitive to the percentage error (see Figure 6). However, as the percentage error approaches zero the number of intervals required increases rapidly. Notice also that the number of intervals needed is much more sensitive to changes in the dynamic range when the dynamic range is small than it is when the dynamic range is large. Figure 7 shows that a change of dynamic range from 2 to 10 (a factor of 5) increases the number of intervals needed in about the same proportion as a change in dynamic range of from 10 to 10,000 (a factor of 1,000).

Application to Binary Systems

Binary system parameters are plotted in Figures 8, 9, and 10. The basic relationship governing a binary system is obtained by combining

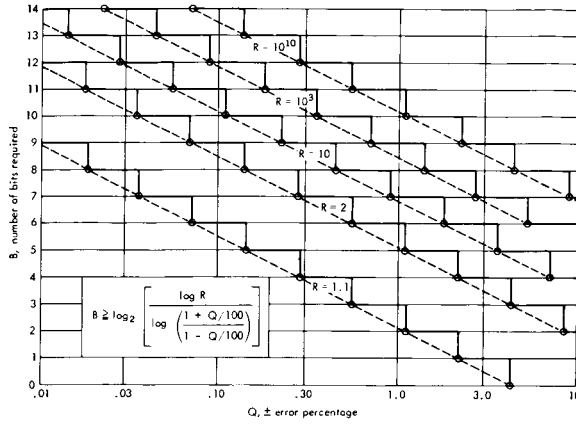


FIGURE 8.—The binary system parameters B and Q.

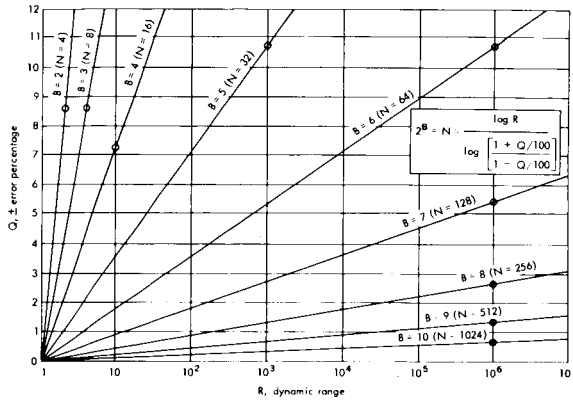


FIGURE 9.—The binary system parameters Q and R.

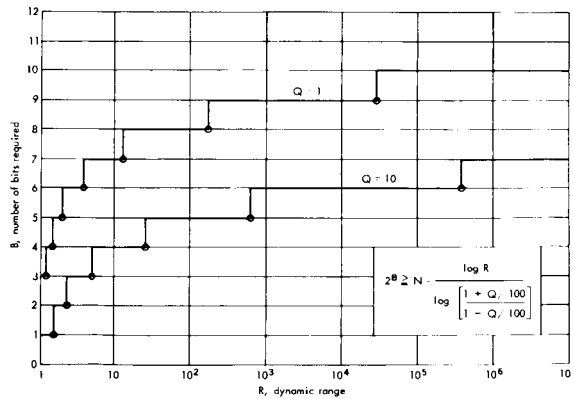


FIGURE 10.—The binary parameters B and R.

Equation 8 with the inequality (Equation 9) to obtain

$$2^B \geq \frac{\log R}{\log \left(\frac{1 + \frac{Q}{100}}{1 - \frac{Q}{100}} \right)} \quad (11)$$

The number of bits required is the smallest integer B such that Equation 11 is satisfied.

Turning our attention to Figure 8, we see the number of bits B versus the error Q for various values of the dynamic range R. This is a particularly useful presentation of system parameters, because the dynamic range is a characteristic of the input variable, and as such is often not under the control of the designer of the coding system.

Since B must be an integer, the curves are discontinuous and have a stair-step appearance. The most desirable points on these curves are the circled lower corner points; at these points Equation 11 holds as an equality and full advantage is being taken of the capacity of the binary system. At these points one obtains the minimum percentage error for any given dynamic range and number of bits.

If, on the other hand, the designer is confronted with a system of fixed binary capacity, and he wishes to know the percentage error versus the dynamic range, Figure 9 will be found useful. In this figure the percentage error Q is plotted versus the dynamic range R for binary systems of from 2 to 10 bits.

For fixed percentage error, the trade-off between the number of bits and the dynamic range is shown in Figure 10. Again, the desirable operating points are the circled lower corner points where full use is made of the capacity of the binary system. It is helpful to use this graph when trying to decide just how wide a dynamic range to cover when the exact range of the input variable is not known and some guard space is to be left at the ends of the dynamic range.

METHODS OF ACHIEVING LOGARITHMIC QUANTIZATION

There are a number of methods which may be used to effect logarithmic quantization. Conceptually, the most appealing method is to use an analog to digital converter which has logarithmically spaced decision levels. This is the sort of device described heretofore. An equivalent result is produced if the logarithm of the input x is taken using analog techniques, and then the logarithm is subsequently quantized by a linear analog-to-digital converter. It is also possible to

convert an analog input to digital form and to then take the logarithm using digital methods.* Unfortunately this approach requires a rather large capacity digital system preceding the point where logarithms are computed. Particularly, there may be problems involved in obtaining a linear analog-to-digital converter which has both the required dynamic range and the required accuracy in the lower portion of the dynamic range.

AN EXAMPLE

A Logarithmic System

It will be instructive to consider an example of a system using logarithmic quantization and to compare this to linear systems that might be used in the same application. We will assume that the input variable has values of interest which range from 1 to 10, and that we wish to obtain ± 1 percent accuracy. In the example then, $R=10$ and $Q=1$. The number of bits that will be required may be determined from Figure 8 by following the $Q=1$ line up from the Q axis until it intersects the plot for $R=10$. The intersection is found to lie on the $B=7$ line; hence 7 bits will be required. Observe that the system will not be optimum in the sense that with 7 bits (128 quantization intervals) either less error or more dynamic range could be obtained. Reference to Figure 8 reveals that in a 7 bit system with a dynamic range of 10, the error need be only $\pm .9$ percent. On the other hand, the $Q=1$ contour in Figure 10 shows that for 7 bits the dynamic range may be as great as 13 without exceeding ± 1 percent error.

In our example, we will cover the 10 to 1 dynamic range with sufficient intervals to give ± 1 percent error. This will require

$$N = \frac{\log R}{\log \left(\frac{1 + \frac{Q}{100}}{1 - \frac{Q}{100}} \right)} = \frac{\log 10}{\log \left(\frac{1.01}{.99} \right)} = 116 \text{ intervals.}$$

Since $2^6=64$ and $2^7=128$, we require 7 bits. Since $128-116=12$, 12 characters of the 7 bit binary code will not be required. These characters can be used to indicate various error conditions, such as an input which was not within the prescribed dynamic range or a malfunction of the electronic circuitry.

Two Linear Systems

If we do not wish to employ a logarithmic system, an obvious alternative is a linear system. Two such systems come immediately to mind. One has a possible error of ± 1 percent at full scale and progressively greater as the value of the input variable decreases. The other will have ± 1 percent error at the lower end of the dynamic range and considerably less at the upper end. These represent extremes between which any system which is roughly equivalent to our logarithmic system will fall.

Consider first the case where the error is to be ± 1 percent at full scale. As before, the dynamic range will be 10. Starting at full scale (which we will take to be 10) the first few quantization intervals and the outputs which correspond to them are shown at the top of Figure 11. At the other end of the dynamic range, the last few quantization intervals are also shown. For the interval $9.8 < x < 10.0$, the percentage error is $Q'_{is} = \pm (.1/9.9) 100 = \pm 1.01$ percent; at the low end of the scale $1.0 < x < 1.2$, $Q'_{is} = \pm (.1/1.1) 100 = 9.1$ percent. Observe that the width of any quantization interval is 0.2; therefore $(10-1)/.2 = 45$ intervals will be required for this system.

Since $2^5=32$ and $2^6=64$, 6 bits will be required, thus $64-45=19$, and 19 characters of the binary code are not used. As was the case with the logarithmic system, these characters may be used to extend the dynamic range or reduce error, or they may serve as indication of malfunctions and error conditions.

The second linear system is shown in Figure 12. Again it will be assumed that the input ranges from 1 to 10. For the interval $1.00 < x < 1.02$, $Q''_{is} = (.01/1.01) 100 = .99$ percent. As the value of the input increases, the error decreases until for the interval $9.98 < x < 10.00$, $Q''_{is} = (.01/9.99) 100 = .10$ percent. Each interval has the width .02; so $(10-1)/.02 = 450$ intervals will be required. $2^8 = 256$ and $2^9 = 512$; therefore 9 bits must be used

*D. H. Schaefer, Goddard Space Flight Center, has developed simple methods of computing logarithms digitally.

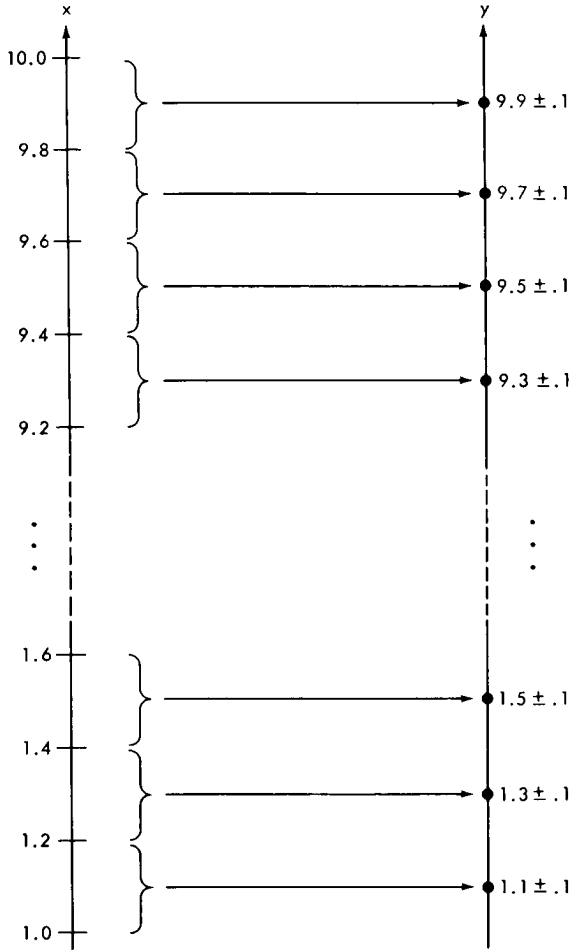


FIGURE 11.—Linear system—Error ± 1 percent of full scale.

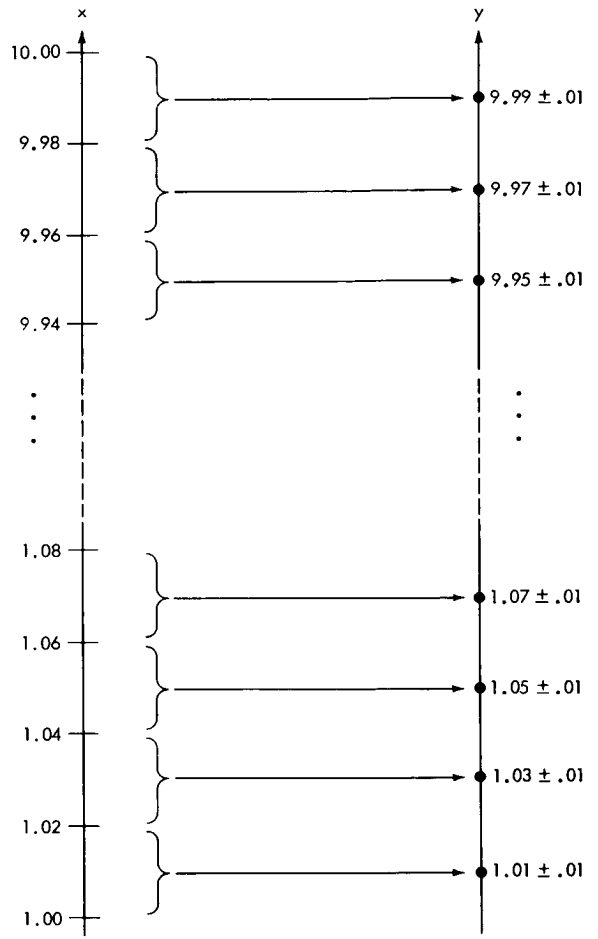


FIGURE 12.—Linear system—Error ± 1 percent of bottom scale.

in the binary system. Thus $512 - 450 = 62$ unused characters. As before, they may be used to extend the dynamic range, reduce the error, or indicate malfunctions and error conditions.

Comparison of Logarithmic and Linear Systems

The quantization intervals of the logarithmic system are shown in Figure 13. Notice that near $x=10$ they resemble those of the first linear system, but near $x=1$ they resemble those of the second linear system.

Table 1 gives a comparison of the logarithmic system and the two linear systems. In order to achieve a percentage error which is everywhere at least as small as that of the logarithmic system, a linear system (System C) requires over three times as many quantization intervals. However,

System C provides more accuracy than is needed throughout most of the dynamic range. This linear system is inefficient because it provides more information than is desired. System A, on the other hand, uses about 1/3 as many quantization intervals as the logarithmic system. Unfortunately it does not meet the error specifications except at full scale. It is desirable because it does not provide enough information.

Although in the example the dynamic range of 10 ran from 1.0 to 10.0, the percentage error and the number of intervals required are a function only of the dynamic range, and not of the absolute value of the limits of the dynamic range. This applies to all three systems which were considered.

It should be remarked that there are situations which require high accuracy in only a portion of

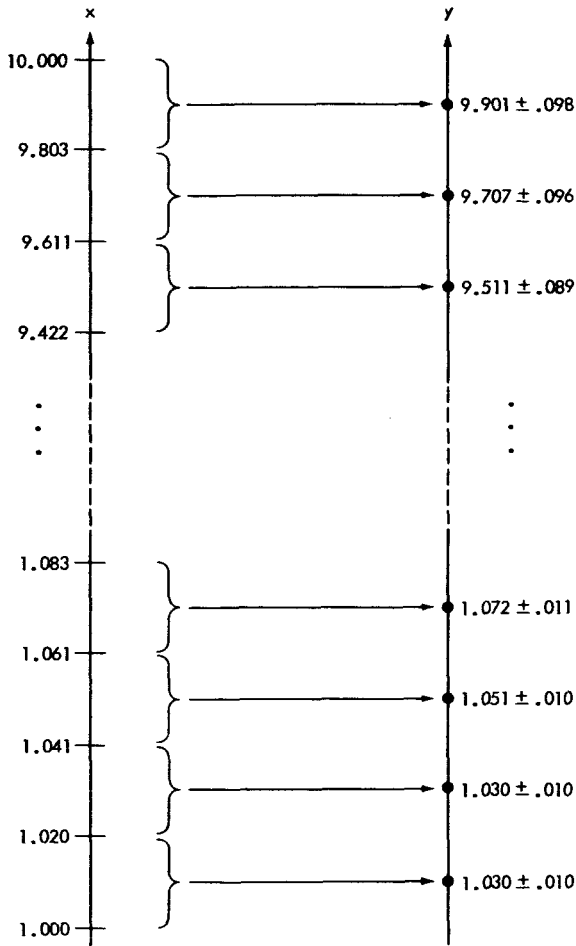


FIGURE 13.—Logarithmic system—Error ± 1 percent of output.

the dynamic range. The remaining part of the dynamic range can be covered by relatively wide quantization intervals. These systems must be tailored to specific applications.

There are other instances where the error is specified to be \pm so many units, rather than as a percentage. In this case a linear system would be used. For example, system A provides an error of ± 1 . The logarithmic system also meets this criterion, but it uses three times as many quantization intervals. Error would be specified in this way if for example the difference of two discrete outputs of near equal magnitude was to be calculated.

CONCLUDING REMARKS

It has been shown that a logarithmic system can be used to advantage to code an analog variable as a sequence of discrete outputs. This coding is such that the possible error associated with each discrete output is the same. An example has shown that if a system is specified as having less than a certain percentage error throughout its dynamic range, that a logarithmic system is considerably more efficient than a linear system.

The basic equations describing the logarithmic system are repeated here for convenience:

$$N = \frac{\log R}{\log \left(\frac{1 + \frac{Q}{100}}{1 - \frac{Q}{100}} \right)} \tag{8}$$

Table 1

The Characteristics of Three Systems.

Characteristic	System A: Linear, $\pm 1\%$ of greatest output	System B: Log, $\pm 1\%$ of output	System C: Linear, $\pm 1\%$ of least output
Greatest Input	10	10	10
Least Input	1	1	1
Dynamic Range	10	10	10
Greatest Output	$9.9 \pm .1$	$9.9 \pm .1$	$9.99 \pm .01$
Least Output	$1.1 \pm .1$	$1.01 \pm .01$	$1.01 \pm .01$
Width of Quantization Intervals	.2	.02y	.02
Number of Quantization Intervals	45	116	450
Number of Bits Required	6	7	9
Unused Characters in Binary Code	19	12	62
Total Characters in Binary Code	64	128	512

$$\frac{x^+}{x^-} = \frac{1 + \frac{Q}{100}}{1 - \frac{Q}{100}} \quad (4)$$

$$y_x = \frac{x^+ + x^-}{2} \quad (2)$$

Appendix B contains the definitions of the symbols used in these equations and elsewhere in this paper.

A logarithmic system may involve more com-

plex equipment for the conversion of an analog variable to discrete form than a linear system would. However, the digital system which uses the discrete outputs of the logarithmic system will have fewer bits upon which to operate (for equal percentage errors) and will consequently be less complex. This feature is particularly attractive in the telemetry business where it is frequently necessary to transmit a maximum of information with a minimum number of bits in order to conserve bandwidth and transmitter power.

APPENDIX A

DERIVATION OF THE 1/Q RELATIONSHIP

It may be shown that for constant R, N varies as 1/Q. In the case of constant R, Equation 9 becomes

$$N = \frac{\log_a R}{\log_a \left(\frac{1 + \frac{Q}{100}}{1 - \frac{Q}{100}} \right)} = \frac{\text{const}}{\log_a \left(\frac{1 + \frac{Q}{100}}{1 - \frac{Q}{100}} \right)} \quad (\text{A1})$$

The logarithms in the right-hand terms may be taken to any convenient base. By the properties of logarithms,

$$\log \left(\frac{1 + \frac{Q}{100}}{1 - \frac{Q}{100}} \right) = \log \left(1 + \frac{Q}{100} \right) - \log \left(1 - \frac{Q}{100} \right) \quad (\text{A2})$$

The left hand terms of Equation A2 may conveniently be expanded in a power series by application of the relation:*

$$\log_e (1 + X) = X - \frac{1}{2} X^2 + \frac{1}{3} X^3 - \frac{1}{4} X^4 + \dots \quad (\text{A3})$$

This yields

$$\log_e \left(1 + \frac{Q}{100} \right) = \frac{Q}{100} - \frac{1}{2} \left(\frac{Q}{100} \right)^2 + \frac{1}{3} \left(\frac{Q}{100} \right)^3 - \frac{1}{4} \left(\frac{Q}{100} \right)^4 + \dots \quad (\text{A4})$$

$$- \log_e \left(1 - \frac{Q}{100} \right) = \frac{Q}{100} + \frac{1}{2} \left(\frac{Q}{100} \right)^2 + \frac{1}{3} \left(\frac{Q}{100} \right)^3 + \frac{1}{4} \left(\frac{Q}{100} \right)^4 + \dots \quad (\text{A5})$$

Hence,

$$\log_e \left(\frac{1 + \frac{Q}{100}}{1 - \frac{Q}{100}} \right) = 2 \left(\frac{Q}{100} \right) + \frac{2}{3} \left(\frac{Q}{100} \right)^3 + \dots \quad (\text{A6})$$

and therefore, if $\frac{Q}{100} \ll 1$,

$$N \approx \frac{\log_e R}{2 \frac{Q}{100}} = \frac{\text{const}}{Q} \quad (\text{A7})$$

*Handbook of Chemistry and Physics, 43rd ed., pg. 323.

By using the relation*

Equation A7 may also be written

$$\log_e x = 2.30 \log_{10} x . \quad (\text{A8})$$

$$N \approx \frac{115 \log_{10} R}{Q} . \quad (\text{A9})$$

**Handbook of Chemistry and Physics*, 43rd ed., pg. 12.

APPENDIX B

SYMBOLS

B = number of bits required for N intervals	Q''_{bs} = bottom of scale percentage error for system C
$K = x^+/x^-$, the dynamic range covered by one quantization interval	Q''_{fs} = full scale percentage error for system C
N = number of quantization intervals required	R = dynamic range over which the input variable is to be quantized
Q^+ = maximum possible positive percentage error	x = analog input variable
Q^- = maximum possible negative percentage error	X = an arbitrary quantization interval
Q = maximum possible \pm percentage error	x^- = lower limit of quantization interval X
Q'_{bs} = bottom of scale percentage error for system A	x^+ = upper limit of quantization interval X
Q'_{fs} = full scale percentage error for system A	y = discrete output variable or system C
	y_x = value of output variable corresponding to quantization interval X

THE ELECTRICAL CHARACTERISTICS OF IRRADIATED SILICON SOLAR CELLS AS A FUNCTION OF TEMPERATURE*

BRIAN T. CUNNINGHAM, ROBERT L. SHARP AND LUTHER W. SLIFER

The electrical characteristics of N/P (less than one ohm-cm to a nominal 10 ohm-cm base resistivity) and P/N (one ohm-cm) silicon solar cells irradiated with 1 Mev electron flux to a level of 10^{16} electrons/cm² have been obtained through a temperature range of -100 degrees Centigrade to $+125$ degrees Centigrade. The information was obtained using a tungsten light source in conjunction with a 3-centimeter water filter. Correction techniques are described to convert the spectral and intensity distributions to the air mass zero operating environment. Curves are presented which relate the power, voltage, temperature and 1 Mev electron flux level in a manner useful to the solar power supply designer.

Satellite solar power supplies are subject to degradation from temperature and hard particle radiation. When the power supply is operating in the constant voltage mode, it is essential to the optimum operational voltage per cell for the space environment.

Post-irradiation electrical measurements were made on N/P (less than 1 ohm-cm to a nominal 10 ohm-cm base resistivity) and P/N (1 ohm-cm base resistivity) state-of-the-art (1962) silicon solar cells. These cells had been irradiated at room temperature with 1 Mev electrons at various flux levels up to 10^{16} electrons/cm². The electrical characteristics of the cells were determined for the entire temperature range between -100°C and $+125^{\circ}\text{C}$, using a tungsten light source with a 3-cm water filter. Spectral and intensity corrections can be made to obtain air mass zero equivalent outputs.

The measurements clearly show the superiority of the 10 ohm-cm N/P cell for all solar power supplies subject to appreciable amounts of irradiation. Families of curves were developed which relate the solar cell power, operational voltage, temperature and 1 Mev flux level. These curves will aid the designer in the selection of the optimum operational voltage per cell. They can be directly applied when the solar power supply is operating in the constant voltage mode and adapted through cross-plotting to a peak power or constant load system. Selection of too high an operational voltage per cell, combined with a low estimate of either solar array temperature or hard particle irradiation, can be disastrous for the power system.

INTRODUCTION

Numerous studies have been conducted to determine the effect of electron irradiation on silicon solar cells. In order to extend the knowledge already gained in this area, an investigation was undertaken to determine the electrical characteristics of irradiated silicon solar cells as a function of temperature. Consultation with individuals familiar with satellite environments indicated that a temperature range of -100°C to $+125^{\circ}\text{C}$ would be appropriate for study.

During October and November 1962, silicon N/P and P/N state-of-the-art solar cells were purchased from nine different manufacturers. All the cells received have a photovoltaic area of approximately 1.8 cm². Since no particular requirement was placed on base resistivity, the cells received from the producers varied from less than 1 ohm-cm to approximately 10 ohm-cm, depending on the manufacturer's production methods. The voltage current (E-I) characteristics of these cells were measured in sunlight and under tungsten illumination at, or near, room temperature. Some cells from each manufacturer were set aside for thermal and mechanical tests and thirty cells from

*Published as *Goddard Space Flight Center Document X-636-64-155*, March 1964.

each manufacturer were forwarded to the Naval Research Laboratory, Washington, D.C. for irradiation. Fifteen of these cells were irradiated with 1 Mev electrons in dosages varying from 10^{11} to 10^{16} electrons/cm². The room temperature characteristics of these irradiated cells were obtained¹ and these same cells were then made available for the temperature effects study described here.

EQUIPMENT

The experimental equipment used to collect the data is shown in Figures 1 and 2. The unit consists of a 9" diameter aluminum temperature-controlled chamber sealed off on one end by a 3 cm thick, circulating water, cooling bath contained by $\frac{1}{4}$ " thick plexiglass walls. The water bath also served as a filter for the tungsten source by cutting out a large amount of the undesirable infrared. The other end of the chamber was fitted with an aluminum temperature control block shown in Figure 2. The unit contains both a 340-watt, nichrome wire heater and a cooling spiral of $\frac{1}{4}$ " aluminum tubing for the transfer of liquid nitrogen. The front of the temperature

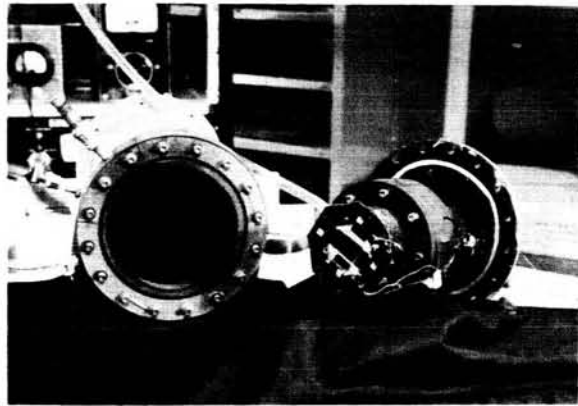


FIGURE 2.—Temperature control block and chamber.

control block is a $\frac{1}{2}$ " thick heat sink and acts as a base for the $\frac{1}{8}$ " thick copper plate on which the cells are mounted.

The light source used in the investigation was a 300-watt tungsten filament, reflector flood bulb. The bulb was calibrated for temperature and was operated at 2800°K in the cylindrical light chamber shown in Figure 1. The entire assembly was constructed so as to prevent stray light from affecting the test. The bulb platform was mounted on a worm screw so that the intensity of illumination on the cells could be adjusted by moving the source closer to or further away from the cells under test. All surfaces exposed to the light source were coated with optical quality black paint to prevent reflections which would disturb the uniformity of illumination. After assembly the entire 6.5 square inch area of the temperature control block was checked for illumination uniformity, using a standard solar cell as a detector and found to be uniform within 2 percent. The uniformity of illumination on the cells is considered better than this because they were mounted in the central 3 square inches of this area.

When in operation, the temperature chamber was evacuated to 35 microns of mercury in order to prevent any frost accumulation on the cells at the low temperatures. Temperatures were measured using three copper-constantan thermocouples calibrated over the temperature range of the tests to an accuracy of $\pm 1^\circ$ C. An attempt was made to determine the cell junction temperature as closely as possible. To this end, the thermocouples were placed on the top (on the



FIGURE 1.—Test equipment.

contact strip) and on the bottom of the cells, first to determine if a temperature gradient existed, and second, to find out what its magnitude might be. The data collected led to somewhat inconclusive results and further work is being performed in this area. The thermocouple outputs were initially measured with a potentiometer. Later a deflection type pyrometer was employed, which gave results nearly as accurate as those from the potentiometer with a great saving of time. At the extremes of the temperature range, an accuracy of $+5^{\circ}\text{C}$, including errors due to temperature gradients on the cells, was obtained. This accuracy improved to within $+1^{\circ}\text{C}$ as room temperature was approached.

The solar cell electrical characteristics were measured by varying the load resistance across the cell from 1 ohm to 2500 ohms. The current flowing with the 1 ohm load across the cell is, by accepted definition, short circuit current, I_{sc} , while the voltage across the 2500 ohms is considered to be open circuit voltage, V_{oc} . The complete $I-V$ characteristics were recorded on an $X-Y$ plotter to obtain a permanent record.

PROCEDURE

$I-V$ characteristics were obtained for all cells, using both sunlight and the filtered tungsten source. For the sunlight measurements, the sun intensity was monitored with a 15-junction normal incidence pyrliometer. The solar cell short circuit current was measured in collimated sunlight and then linearly normalized to the short circuit current for a sunlight intensity of 100 mw/cm^2 . These normalized current readings were compared to the short circuit current readings obtained from measurements made on the cells at room temperature using the tungsten light source. The latter measurements were made with the illumination set at $100\text{ milliwatts/cm}^2$ "sunlight equivalent." To accomplish this, a standard solar cell with a known short circuit current at $100\text{ milliwatts/cm}^2$ sunlight input was used in setting the illumination level from the tungsten light source so that the same short circuit current was obtained.

Mounting of the cells was initially accomplished by heating a pretinned copper plate to the solder melting point. Four solar cells were then placed on the plate and the unit was removed to a cool

heat sink as quickly as possible. Results using this system were poor. First, electrical measurements before and after mounting indicated changes in short circuit current up to 4 percent; and secondly, cells under test would occasionally shatter when undergoing temperature changes. To alleviate this problem, a conductive epoxy was used to bond the cell to the test plate. Curving of the epoxy at 60°C had no measurable effect on the electrical characteristics of the cells; neither was any additional series resistance evident. The epoxy held throughout the test temperature range and no cracking to the cells was observed after institution of its use.

Cells to be tested were selected at random in groups of four and mounted on the copper plate. The cells were then placed in the chamber and the illumination was adjusted to give the "pre-chamber installation" short circuit current value from the cells. The chamber was then evacuated and the temperature lowered to -110°C . Because of the comparatively large heat sink, it was found suitable to simply let the temperature drift upward while readings were taken at the desired temperature intervals. In most cases, all four $I-V$ curves could be obtained with a variation of only $\pm 2^{\circ}\text{C}$ of the nominal measuring point. $I-V$ curves were then analyzed for short circuit current, open circuit voltage, maximum power, and power at specific voltages.

RESULTS AND DISCUSSION

The data obtained from the measurements and its analysis and interpretation will be presented in four parts. The first three parts present individual electrical parameters, i.e., short circuit current, open circuit voltage and maximum power as a function of temperature and irradiation. The final section shows the interrelationship of power, voltage, temperature and flux (P , V , T and Φ) in the manner considered most useful to the solar power supply designer.

It should be noted that the electrical parameters in the ensuing analysis are presented in absolute terms and are therefore subject to criticism from a "representative sampling" viewpoint. In defense of this approach it can be said:

1. All the available irradiated samples were temperature-tested.

2. Wherever more than one sample was available (i.e., nonirradiated cells and those irradiated to 10^{16} electrons/cm²), favorable comparisons were obtained between similar cells throughout the temperature range.
3. There is a favorable comparison between the results obtained during these tests and earlier results obtained with a larger number of samples at room temperature.

For these reasons, it is felt that the results are, in reality, representative of large samples and the information sufficiently reliable to be extended to this application.

SHORT CIRCUIT CURRENT

The initial room temperature measurements showed that for non-irradiated cells under tungsten light, essentially the same readings were obtained as in sunlight regardless of manufacturer. However, for *N/P* solar cells irradiated with dosages greater than 10^{13} electrons/cm², the percent change as seen under tungsten light became significantly different, depending upon manufacturer, from that seen under sunlight. It should be recalled that in *N/P* category several different base resistivity cells were under evaluation, ranging from less than 1 ohm-cm to more than 10 ohm-cm. Results of an earlier investigation¹ show that the higher base resistivity cells are less susceptible to radiation damage from 1 Mev electrons at room temperature. For *P/N* solar cells a significant difference in the results under the two light sources was observed for dosages greater than 10^{11} electrons/cm². The percent change in short circuit current for solar cells due to irradiation of 10^{16} electrons/cm² as measured using tungsten and sunlight is shown in Table 1.

These results show the extent to which the degradation in the long wavelength end of the response of the solar cells due to electron irradiation² is to be considered when analyzing data obtained from measurements on irradiated cells using a source other than sunlight. In the interpretation of the information presented here, it must be emphasized that tungsten operating at 2800°K is a "red" source, that is, the spectrum reaches a maximum near 1.0 micron (in the infrared region), while the sun is "blue" source and peaks near 0.46 micron (in the blue). Since electron irradiation degrades the red end of the solar cell spectral response,

TABLE 1.—Comparison of Sunlight and Tungsten Sources for Measurement of Short Circuit Current on Irradiated Solar Cells

Solar Cell Nominal Base Resistivity	Type	Percent Change in I_{sc} after 10^{16} electrons/cm ²	
		Tungsten Light (100 mw/cm ² sunlight equivalent)	Sunlight (100 mw/cm ²)
10 ohm-cm-----	<i>N/P</i>	37	26
1 ohm-cm-----	<i>N/P</i>	49	37
<1 ohm-cm-----	<i>N/P</i>	57	40
1 ohm-cm-----	<i>P/N</i>	76	63

radiation degradation appears worse under tungsten light than under sunlight.

As shown in Figure 3, the short circuit current does not change linearly with temperature; however, it is worthwhile to compare the change observed over the entire temperature range with respect to irradiation. The average short circuit current coefficient for a non-irradiated cell is 50.8μ A/°C, while for a cell irradiated to 10^{16} electrons/cm² it increases to an average of 69.8μ A/°C over the range -100°C to $+100^\circ\text{C}$. This average is derived from data taken on all manufacturers' cells.

Short circuit current degradation as a function of irradiation has been adequately covered in an earlier report.¹ Since the results of the present study do not add to the data already presented except as previously mentioned, no further discussion need be given on this parameter.

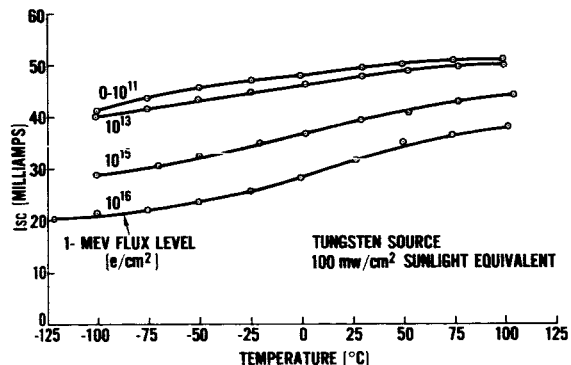


FIGURE 3.—Effect of temperature and irradiation on short circuit current (typical cells).

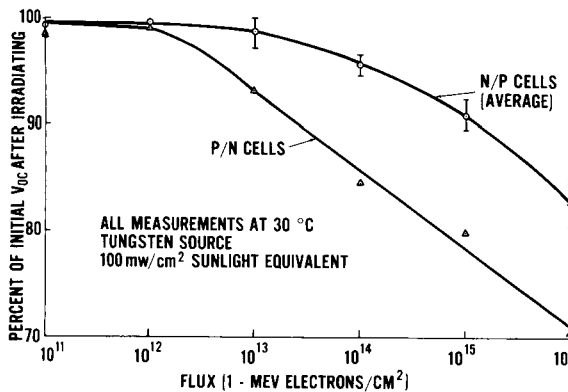


FIGURE 4.—Degradation of open circuit voltage.

OPEN CIRCUIT VOLTAGE

Open circuit voltage degradation as a function of flux level is shown in Figure 4 for both *N/P* and *P/N* solar cells. The *N/P* degradation points at each flux level represent an average of five different cells, while the *P/N* points at each level represent a single cell. The latter data should therefore be treated cautiously. Within the limits of the experiment no appreciable difference in radiation degradation of the open circuit voltage was observed between *N/P* cells of different base resistivities.

As shown in Figure 5, it was found that open circuit voltage, for both irradiated and non-irradiated (control) cells, decreased linearly with temperature over the entire test range. The average coefficient of open circuit voltage change with temperature was -2.34 $\text{mv}/^\circ\text{C}$. The extremes

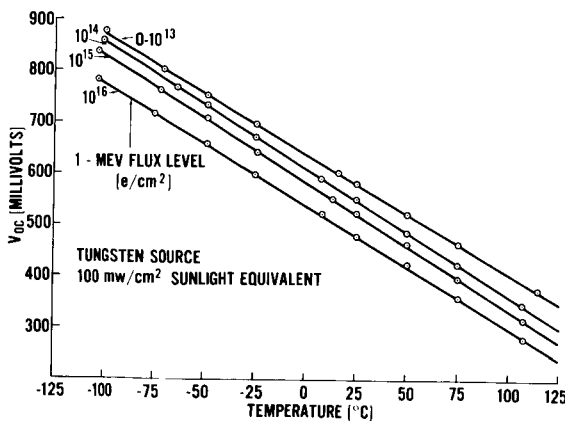


FIGURE 5.—Effect of temperature and irradiation on open circuit voltage (typical cells).

were -2.16 $\text{mv}/^\circ\text{C}$ and -2.60 $\text{mv}/^\circ\text{C}$. Since the variation noted appeared between individual cells of the same manufacturer rather than being manufacturer-dependent, it is probably due in large part to experimental error, but effects of real differences between cells (base resistivity effects, for example) are not considered negligible. For this reason, the average of -2.34 $\text{mv}/^\circ\text{C}$ is considered the appropriate choice for power supply design calculations. The temperature coefficient was also found to be independent of the level of irradiation as shown in Figure 5. Figure 6 depicts the percent change in open circuit voltage as a function of the percent change in short circuit current for both sunlight and tungsten sources at different irradiation levels for a nominal 1 ohm-cm solar cell. Figure 7 is a similar curve except that the cell has a nominal 10 ohm-cm base resistivity. The purpose of presenting these figures is to emphasize the fact that the degradation of

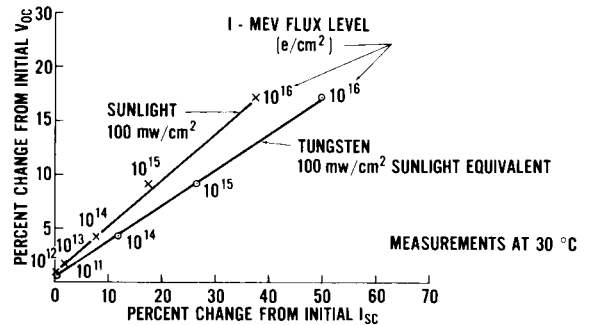


FIGURE 6.—Comparison of open circuit voltage degradation to short circuit current degradation for nominal 1 ohm-cm cells.

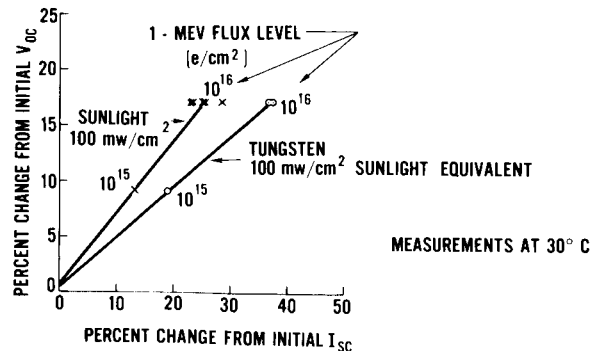


FIGURE 7.—Comparison of open circuit voltage degradation to short circuit current degradation for nominal 10 ohm-cm cells.

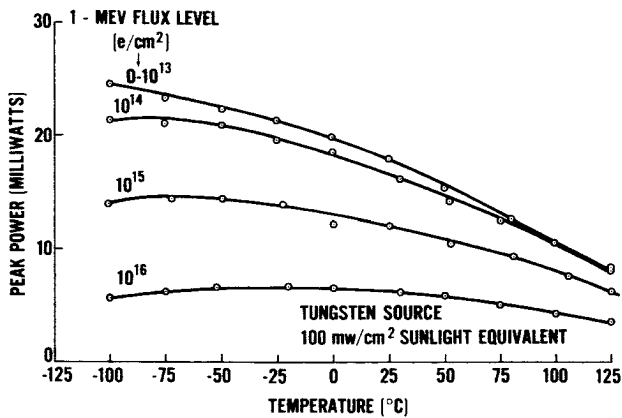


FIGURE 8.—Effect of temperature and irradiation on peak power (typical cells).

open circuit voltage with irradiation must be considered in the design of a solar power supply.

Figure 8 shows the variation in peak power as a function of temperature and irradiation for a typical set of *N/P* cells under investigation. Peak power is somewhat meaningless to the power supply designer without reference to temperature, voltage and irradiation level, so no attempt will be made at this point to present detailed variations from one type of cell to another. In a later section of this report, curves will be presented which will allow the designer to determine peak power, taking into consideration the temperature and irradiation levels of interest. Specific values of radiation degradation of peak power at room temperature have been reported earlier.¹ Comparison between results obtained with tungsten light and sunlight as sources for measurement is given in Table 2.

P, V, T, AND Φ CURVES

In order to relate the appropriate solar cell electrical parameters for the power supply designer, curves have been drawn in the manner of Figure 9. These curves relate power (*P*), "operational" voltage (*V*), temperature (*T*), and flux level (Φ). The origin of this data is the set of *I-V* characteristic curves which were obtained during this investigation for temperatures between -100°C and $+125^{\circ}\text{C}$ in increments of 25°C . The power at each "operational" voltage is determined and the results are then combined to form the family of *P*, *V*, *T*, and Φ curves. In Figures

TABLE 2.—Comparison of Sunlight and Tungsten Sources for Measurement of Peak Power on Irradiated Solar Cells

Solar Cell Nominal Base Resistivity	Type	Percent Change in P_{\max} after 10^{16} Electrons/cm ²	
		Tungsten Light (100 mw/cm ² sunlight equivalent)	Sunlight (100 mw/cm ²)
10 ohm-cm-----	<i>N/P</i>	50.2	41.3
1 ohm-cm-----	<i>N/P</i>	58.2	47.6
<1 ohm-cm-----	<i>N/P</i>	65.2	47.4
1 ohm-----	<i>P/N</i>	85.5	78.0

9, 10, and 11, the 1 Mev electron flux level and the operational voltage have been held constant so that it is possible to evaluate the effect of cell type and base resistivity on power output. As shown in Figure 9, where the cells have not been exposed to any hard particle irradiation, all

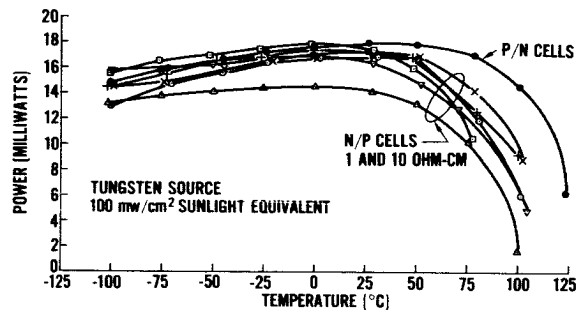


FIGURE 9.—Comparison of power output for non-irradiated solar cells at 0.35 volts.

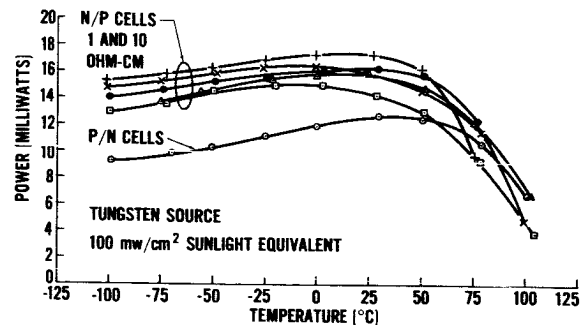


FIGURE 10.—Comparison of power output at 0.35 volts for solar cells irradiated with 1 Mev electrons to 10^{13} electrons/cm².

the manufacturers' cells lie in a relatively narrow range. For the power supply designer, the P/N cell is the most appropriate choice here due to its higher open circuit voltage, which gives it a higher power over a wider temperature range. In Figure 10, irradiation has progressed to 10^{13} electrons/cm² and all cells with the exception of the severely degraded P/N cell lie in a close range of power output over the temperature range. Hence, if hard particle irradiation equivalent to the damage done by 1 Mev electrons at 10^{13} electrons/cm² were anticipated, the designer would be able to use most effectively the highest efficiency (as measured at room temperature) non-irradiated N/P cell available. Some cross-over will be noted at the higher temperatures; however, it is too early, in terms of irradiation, to see any distinct advantage of base resistivity. Progressing to 10^{16} electrons/cm², in Figure 11, there is a distinct advantage to the higher base resistivity cells. This advantage becomes less evident as the operational voltage is increased, and in fact, in *extreme cases* (high temperature) the 1 ohm-cm cell is the more desirable. The reason for this is the higher initial open circuit voltage of the 1 ohm-cm cell relative to the 10 ohm-cm cells among the cells under investigation. However, the solar cell would be operating in an extremely precarious and inefficient region of the $P, V, T,$ and Φ curve in order to achieve any advantage of this characteristic. Figure 12 depicts this extreme situation on an expanded power scale and shows the 10 ohm-cm, 0.40 volt curve

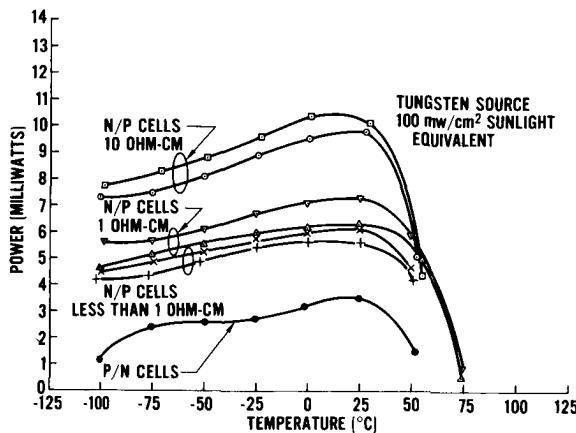


FIGURE 11.—Comparison of power output at 0.35 volts for solar cells irradiated with 1 Mev electrons to 10^{16} electrons/cm².

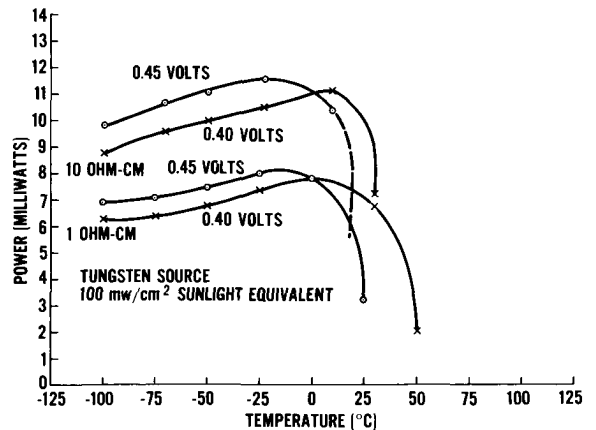


FIGURE 12.—Comparison of 1 and 10 ohm-cm solar cell power output at 0.40 and 0.45 volts after irradiation to 10^{16} electrons/cm².

dropping inside the 0.40 volt, 1 ohm-cm curve. The 0.45 volt curves of this figure would have the relationship if complete data were available for this region and hence are shown as such by extrapolation of the 0.45 volt, 10 ohm-cm curve. It should be kept in mind that choosing too high an operational voltage, combined with a low estimate of operating temperature, would be disastrous for the power supply because the curves have a steep slope relative to increasing temperature after the peak is achieved. In consequence of the foregoing, the next series of figures (13–21) present the available $P, V, T,$ and Φ curves for the $N/P,$ 1 and 10 ohm-cm cells only. (Although the 10 ohm-cm solar cell has been shown to be more radiation resistant and the preferred choice for

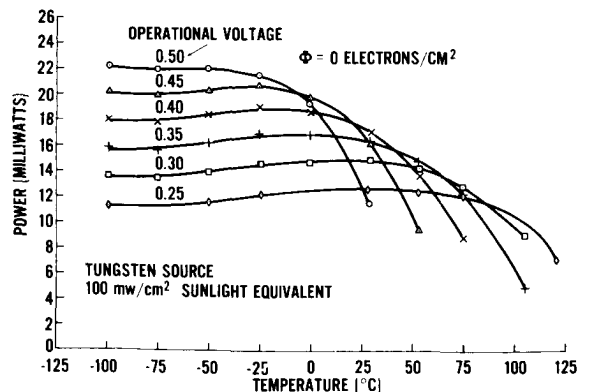


FIGURE 13.— P, V, T, ϕ Curves for a 1 ohm-cm N/P cell with $\phi = 0$ electrons/cm².

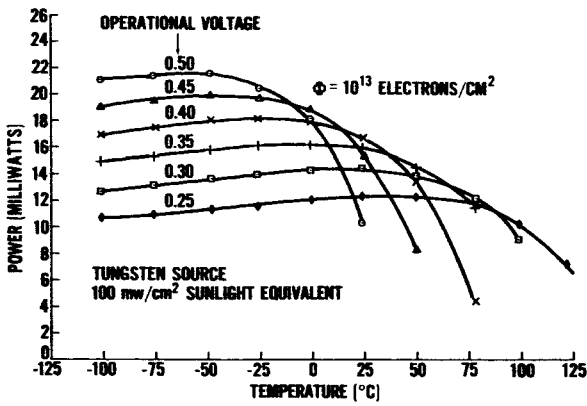


FIGURE 14.— P, V, T, ϕ Curves for a 1 ohm-cm N/P cell with $\phi = 10^{13}$ electrons/cm².

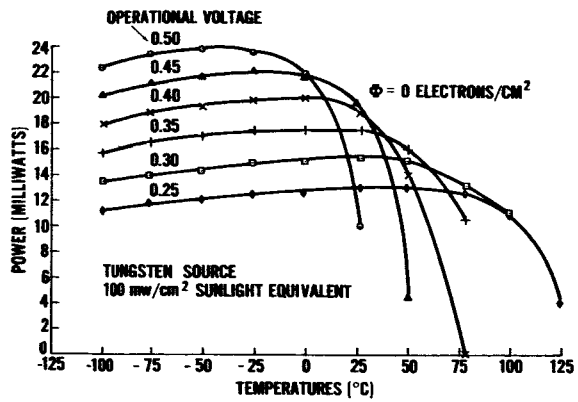


FIGURE 17.— P, V, T, ϕ curves for a 10 ohm-cm N/P cell with $\phi = 0$ electrons/cm².

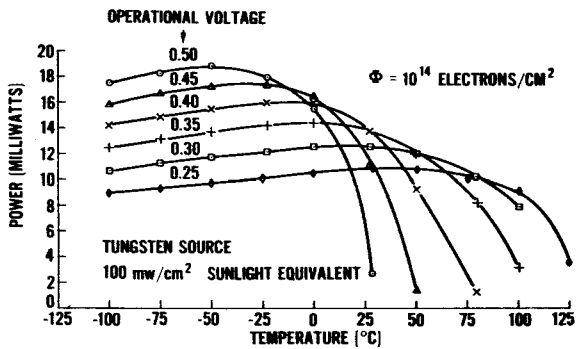


FIGURE 15.— P, V, T, ϕ curves for a 1 ohm-cm N/P cell with $\phi = 10^{14}$ electrons/cm².

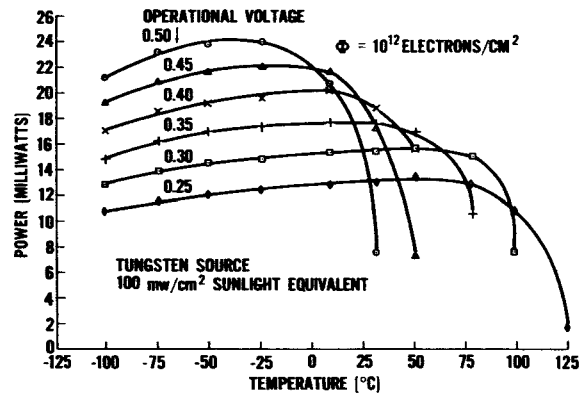


FIGURE 18.— P, V, T, ϕ curves for a 10 ohm-cm N/P cell with $\phi = 10^{12}$ electrons/cm².

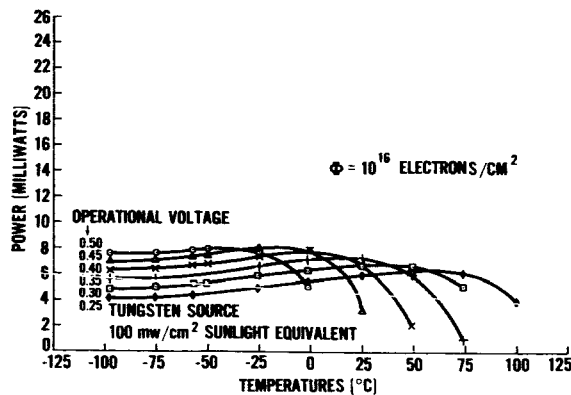


FIGURE 16.— P, V, T, ϕ curves for a 1 ohm-cm N/P cell with $\phi = 10^{16}$ electrons/cm².

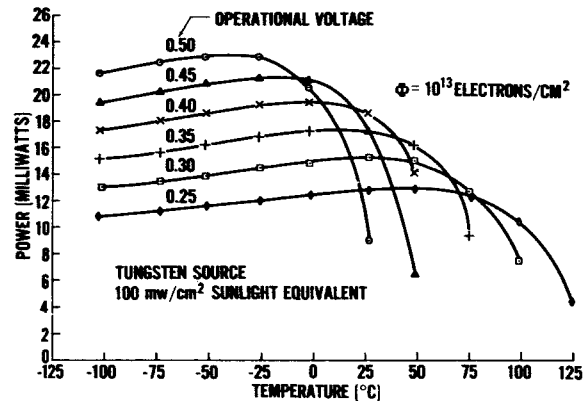


FIGURE 19.— P, V, T, ϕ curves for a 10 ohm-cm N/P cell with $\phi = 10^{13}$ electrons/cm².

the power supply designer in almost all circumstances, the 1 ohm-cm curves are presented because of continued use of this type of cell on solar

power supplies.) These figures will directly aid the power supply designer who is designing a constant voltage supply (the most common in

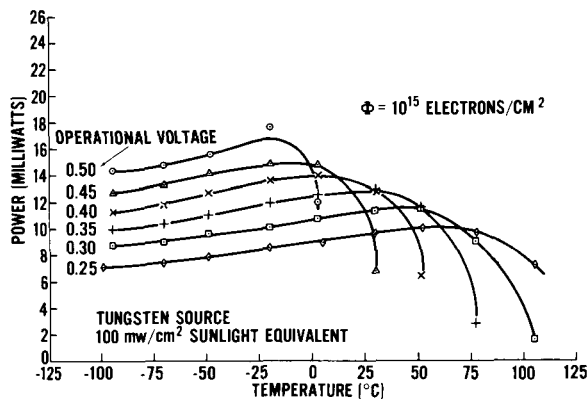


FIGURE 20.— P , V , T , ϕ curves for a 10 ohm-cm N/P cell with $\phi = 10^{15}$ electrons/cm².

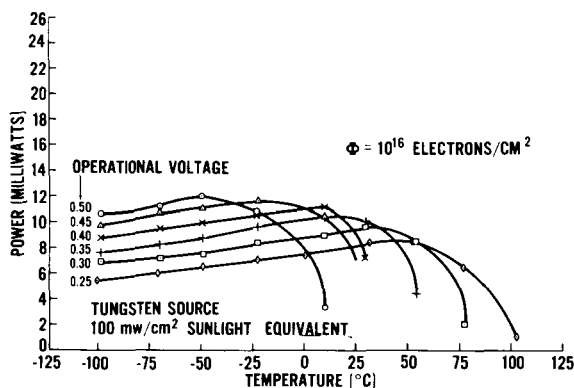


FIGURE 21.— P , V , T , ϕ curves for a 10 ohm-cm N/P cell with $\phi = 10^{16}$ electrons/cm².

satellite use) in determining the most desirable operating voltage once he knows the 1 Mev electron flux equivalent and the operating temperature of the cell. Also, since these curves are derived directly from the I - V curves, they can be used (through cross-plotting) for any other mode of operation. The curves presented for irradiated cells, if used "as is," allow for a considerable margin of safety in that the source used in the investigation was tungsten.

In order to be useful to the spacecraft solar power supply designer, the data obtained during this investigation will have to be corrected to reflect the air mass zero operating environment. Also, extrapolations are necessary to reflect solar cell conversion efficiency improvements made by manufacturers since this project was initiated.

To correct the tungsten air mass 1 equivalent data (i.e., all P , V , T , and ϕ curves in this report) to the air mass zero operating environment, both spectral and intensity adjustments must be made. Tungsten, as was indicated earlier, exaggerates I_{sc} damage seen on irradiated cells. To evaluate this difference, I_{sc} measurements were made under both tungsten light and sunlight on all cells. The results were then normalized to 100 mw/cm² and 30°C. The ratio of the short circuit currents under the two illumination conditions was obtained for each group of cells and is shown in Figure 22 for the 1 and 10 ohm-cm samples. Measurements indicate that this ratio holds, within experimental limits, throughout the temperature range. If the power value, on any of

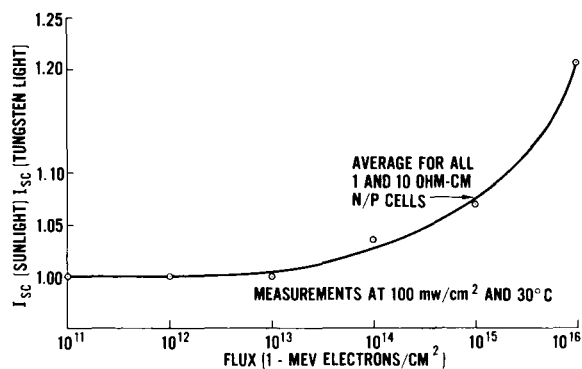


FIGURE 22.—Ratio of short circuit current in sunlight to short circuit current in tungsten light—after degradation.

the P , V , T , and ϕ curves is multiplied by this factor, taken at the appropriate flux level, the result is a sunlight air mass one, 100 mw/cm² value. The air mass zero intensity adjustment is made by obtaining the ratio of energy, in the solar cell response region, of the sunlight air mass zero spectrum to the sunlight air mass one spectrum, weighted according to the spectral response of the cells. This ratio has been determined³ to be 1.17.*

The solar cell efficiency improvement correction is made by obtaining an I - V curve of the new non-irradiated cell at some convenient temperature and illumination condition. Similar infor-

*Different values ranging between 1.17 and 1.23 are in common use. The value 1.17 was chosen here because it represents a properly conservative approach to extrapolations.

mation is extracted from the appropriate (1 or ohm-cm) P , V , T , and Φ curve for the same temperature and illumination conditions. A ratio of the currents of the old and new cell at the constant voltage point of interest is then extracted from this data. All of the power values on the P , V , T , and Φ curves at this particular voltage can then be multiplied by this ratio to obtain a good estimate of the output from the improved cell. This correction may be applied only to cells which are similar, in terms of dopants, base resistivity, etc., to the ones under investigation in this report. Fortunately, most of the cells being used on today's power supplies fall into this category. The result of this correction will still yield a slightly conservative power figure in that additional improvements have been made in both the series resistance and open circuit voltage of the newer cells.

To summarize the above conversions:

$$P_2 = P_1 \left(\frac{I_S}{I_T} \right) \left(\frac{M_0}{M_1} \right) \left(\frac{I_2}{I_1} \right) \quad (1)$$

where

P_2 = Power of the improved solar cell at air mass zero at the given conditions of temperature and irradiation.

P_1 = Power as measured under the tungsten air mass one equivalent conditions. This value is taken directly from the P , V , T , and Φ curves presented herein.

I_S = Sunlight short circuit current after irradiation at a given temperature and 100 mw/cm².

I_T = Tungsten short circuit current after irradiation at the same temperature as I_S (see Figure 22).

M_0 = Air mass zero energy in the solar cell response region weighted by the solar cell response.

M_1 = Air mass one energy in the solar cell response region weighted by the solar cell response.

I_1 = Current of the lower efficiency (old) cell at the desired operational voltage and comparative temperature.

I_2 = Current of the higher efficiency (new) cell at the same operational voltage and temperature and under the same illumination conditions as I_1 .

The use of the P , V , T , and Φ curves and the corrections described above are best illustrated with the aid of a design procedure description.

In this example, the solar array operating conditions have been predicted to be 0°C with a 1 Mev equivalent flux 10¹⁵ electrons/cm² at end of mission.

1. The 10 ohm-cm cell is chosen due to the high Mev flux equivalent. The P , V , T , and Φ curve which fits these conditions is found in Figure 20.
2. The optimum operational voltage is determined from Figure 20. In this case 0.45 volt is nearly optimum. (If more detail is required, power-voltage curves can be constructed from the given data.)
3. Correct the power at this voltage to reflect the air mass zero operating environment and the efficiency improvement on the newer cells using equation (1).

CONCLUSIONS

The following conclusions regarding the characteristics of 1962 state-of-the-art silicon solar cells bombarded under low level illumination and room temperature conditions by 1 Mev electrons have been obtained.

1. Post-irradiation measurements in the temperature range from -100°C to +125°C show N/P solar cells to be decidedly more radiation resistant than the 1 ohm-cm P/N cells.
2. Post-irradiation measurements in the temperature range from -100°C to +125°C show the nominal 10 ohm-cm N/P solar cells to have greater radiation resistance than the lower, 1 ohm-cm and less than 1 ohm-cm nominal base resistivity cells.
3. No single cell type (N/P or P/N ; 1 ohm-cm or 10 ohm-cm base resistivity) can be chosen as the best cell (for power supply use) for the entire range of irradiation from 0 to 10¹⁶ electrons/cm², the entire range of temperature from -100°C to +125°C, and the

entire range of operating voltage. This results primarily from the fact that significant differences exist in the open circuit voltages of the various types of cells. However, the following generalizations can be made:

a. Prior to irradiation, P/N cells have better electrical power generating characteristics than N/P cells and the difference is most significant for higher operating voltages (above 0.35 volt/cell) and higher temperatures (above 50°C). This is primarily because of the higher open circuit voltage of the P/N cells.

b. After nominal irradiation (equivalent to 10^{13} electrons/cm²), the overall higher radiation degradation rate of the P/N cells results in the loss of the advantage at the higher operating voltages and higher temperature conditions and also results in a disadvantage under low voltage and low temperature conditions. The amount of radiation is still not sufficient to produce a distinction between base resistivities in the N/P cells.

c. After extensive irradiation (equivalent to 10^{16} electrons/cm²), the P/N cells are significantly poorer than any of the N/P cells under all conditions of temperature and operating voltage. The difference in radiation sensitivity is sufficient to clearly show the nominal 10 ohm-cm cells to be better than the nominal 1 or less than 1 ohm-cm cells except at the higher temperatures and higher voltage.

4. Degradation of open circuit voltage as a function of both 1 Mev electron flux and increasing temperature has a significant effect on the choice of operating voltage for a solar power supply.

a. The percentage degradation of open circuit voltage in sunlight with irradiation is roughly $\frac{1}{2}$ the percentage degradation of short circuit current.

b. Open circuit voltage varies linearly with respect to temperature between -100°C and +125°C for both irradiated and non-irradiated silicon solar cells.

c. The temperature coefficient of open circuit voltage is constant for any cell but varies

from cell to cell. The average value is -2.34 mv/°C.

d. Irradiation does not appreciably affect the temperature coefficient of open circuit voltage in the temperature range from -100°C to +125°C.

5. Degradation of short circuit current with 1 Mev electron flux is significant, but variation with temperature is relatively unimportant.

a. Short circuit current increases but does not increase linearly over the temperature range from -100 to +100°C.

b. In general, the average temperature coefficient of short circuit current increases with increasing radiation. For non-irradiated 1×2 cm cells, the average coefficient is 50.8μA/°C and for cells irradiated to 10^{16} electrons/cm², it is 69.8μA/°C.

6. Peak power generally decreases linearly above -50°C. Below -50°C, peak power is observed to increase, decrease and not change, depending on the cell. In most cases the change at temperatures below -50°C is small.

7. The P , V , T , and Φ curves represent a convenient aid to the solar power engineer, particularly in the design of constant voltage power supplies, but also in the design of power supplies operating in other modes. These curves show that at less than 10^{13} electrons/cm² the degradation is essentially the same for all N/P cells regardless of base resistivity. They also show that at irradiation levels above 10^{13} electrons/cm², the use of higher base resistivity cells is desirable.

BIBLIOGRAPHY

1. WILLIAM R. CHERRY and LUTHER W. SLIFER, JR., *Solar Cell Radiation Damage Studies with 1 Mev Electrons and 4.6 Mev Protons*, NASA—Goddard Space Flight Center, Greenbelt, Md., May 27, 1963.
2. WILLIAM C. COOLEY and ROBERT J. JANDA, *Handbook of Space Radiation Effects on Solar Cell Power Systems*, Washington, D.C., Office of Scientific and Technical Information, NASA, 1963, p. 28.
3. *Solar Cell Measurements*, Hoffman Electronics Corporation, Semi-conductor Division, El Monte, California, May 1960.

EXPERIMENT.

TIME SYNCHRONIZATION OF REMOTE STATIONS USING SYNCOM SATELLITE*

RAYMOND L. GRANATA, PETER D. ENGLER, AND PAUL F. McCAUL

OBJECTIVE

The primary objective of this experiment will be the accurate transfer of "time" between two remote stations. As our goal, we intend to transfer UT-2 time, corrected Universal Time, from one Syncom station to another remote Syncom station using the Goddard Range and Range Rate equipment. One station will be designated the master and the other the slave. The proposed system will be capable of achieving time synchronization to an accuracy of .5 microsecond or less. Present time synchronization systems, (WWV or VLF), will only allow an accuracy of approximately 1000 microseconds. Included in our experiment is a method of verifying the time synchronization, measuring the system uncertainty, and maintaining accurate synchronization.

TECHNIQUE

In order to achieve this accurate synchronization of time, we have two major areas to cover, (1) the transfer and maintenance of time from the U.S. Naval Observatory to the master, (2) the transfer and maintenance of time between the two Syncom Ground Stations, via the Syncom Satellite.

Time Synchronization for the first area will be accomplished by the use of a portable clock which can easily be carried from point to point. This portable clock will first be calibrated for its drift characteristics and then carried to the U.S. Naval Observatory for the initial time synchronization to UT-2 time; it will then be carried by air to the master station and used to set the station clocks.

The second transfer of time between the two Syncom ground stations, will be the primary

purpose of the experiment. This transfer will be accomplished using the Goddard Range and Range Rate System; this system determines distance by means of measuring and recording delay time so that this system can easily be utilized for transferring time between the two ground stations. The Goddard Range and Range Rate System has been capable of measuring the round trip time delay to the satellite to an accuracy of 0.1 microsecond.

After synchronization of the master station's time standard to NBS the master will at a prearranged time broadcast a pulse to the slave. The slave will then measure the delay, T_D , from the prearranged time at its station until it receives the pulse from the master. Immediately after receiving this pulse the slave will range to the satellite and measure the delay, T_2 . The master will have measured the delay, T_1 , from its own pulse sent at the prearranged time. The value of T_1 will be communicated to the slave by means of a teletype link.

With the delay information the slave can compute (Figure 1) the amount by which it is out of time synchronization with the master and correct its time standard. After the correction is made the synchronization will be verified by reversing the procedure.

It is extremely important that the experiment be performed at periods of minimum satellite radial velocity, and that accurate predictions of range and velocity be provided. A radial velocity component of 50 feet per second will cause an error of 0.1 microsecond per second; so that in order to materially erase the problem of velocity correction, the experiment will be performed when the radial velocity is at a minimum for the slave station.

Once time has been transferred to the station,

*Published as *Goddard Space Flight Center Document X-531-64-7*, January 1964.

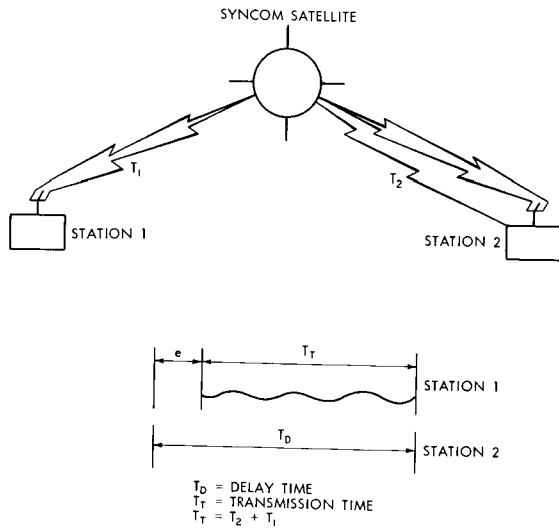


FIGURE 1.—Syncom experiment.

it is necessary to maintain accurate synchronization in order to accomplish the experimental objective. The problem involved is due to the drift of the oscillators. A drift of one microsecond per hour is equal to an oscillator stability of approximately 3×10^{-10} /hour. Oscillator stabilities of this order of magnitude can be achieved; we have lab units which have shown a stability of approximately 1×10^{-11} /day (1 microsecond per day).

Even with this stability oscillators have a drift characteristic which must be measured and corrected for in order to adequately maintain synchronization. In order to make these corrections VLF (Very Low Frequency) techniques have been developed which enable the station to continuously monitor their frequency and maintain a daily error of less than 5×10^{-11} ; if the monitoring is performed for one to two weeks the daily error can be reduced to 1 to 2 parts in 10^{11} . We propose to use this VLF technique at both the Syncom ground stations for maintaining an ac-

curate frequency and time synchronization.

Another important factor is the ability to reset the time of the station clocks accurately. The present Syncom station clocks do not allow the operator to make fine adjustments of the time in order to synchronize their station with another time reference. We propose to furnish, along with other equipment, clocks with the ability to set time to 0.1 microsecond.

OBTAINABLE PRECISION

The measurable time difference using Range and Range Rate is ± 0.1 microsecond, and the measurable time delay due to satellite velocity using Range and Range Rate is ± 1 nanosecond.

The differential phase shift in receivers and transponders due to Doppler and signal level change is expected to be less than 0.1 microsecond.

Using VLF, and the portable time standard the measurable frequency change of the oscillators will be about 1 to 2 parts in 10^{11} per day.

Thus, with provision for the proper peripheral equipment, there is a high probability that two-way time synchronization accuracies of less than .5 microsecond can be achieved, and actual proof of the accuracy obtained can be demonstrated.

EQUIPMENT REQUIRED

The Range and Range Rate System equipment is in operation, at both the stations, as is the Syncom satellite. Also at each station there is an HP-104 A oscillator.

The new equipment consists of 2 auxiliary digital clocks that can be set to 0.1 microsecond, a portable clock, 2 VLF receivers, and two standard oscillators more stable than the HP-104 AR for use with the auxiliary digital clocks.

The auxiliary clocks, portable clock, VLF receivers, and the standard oscillators are already in-house or on purchase order.

AN AUTOMATIC MICROWAVE POLARIZATION ANALYZER

J. A. KAISER AND L. K. HANLIN

INTRODUCTION

A method whereby the polarization characteristics of an incident wave may be instantaneously and continuously measured and indicated is described. It is possible to analyze either modulated or unmodulated signals whose frequencies are within the operating range of the individual system components. The system comprises an antenna which receives and separates the two senses of circular polarization, a microwave phase comparator, and a means for making amplitude measurements. The method described differs from previous methods in that it utilizes the two senses of circular polarization in conjunction with a phase measurement to obtain "tilt" angle rather than using two orthogonal linearly polarized fields in conjunction with a phase shifter.

POLARIZATION RELATIONSHIPS

An arbitrarily polarized wave incident on a receiving antenna, in terms of two orthogonal linear components, may be written in the form

$$E = E\phi + jE\theta = C \cos Wt + jD \sin Wt. \quad (1)$$

As shown by Rumsey,¹ these two orthogonal linearly polarized fields can be represented in terms of two circular fields:

$$\begin{aligned} E_L &= A e^{+jWt} \\ E_R &= B e^{-jWt}. \end{aligned} \quad (2)$$

In a plane normal to the direction of propagation, the circular fields are periodically in-phase, i.e., they add constructively in one direction, twice

during the course of one cycle. In the orthogonal direction, the circular fields add destructively, i.e., they are anti-phase. Figure 1 shows a representation of the two circular fields in a $X-Y$ plane at an instant of time when they are in-phase and so add to produce the major axis of the polarization ellipse. It may be seen that 90 degrees later in time the two fields will be anti-phase because of their opposite senses of rotation, producing the minor axis of the polarization ellipse.

An amplitude envelope of the resultant field is generally an ellipse which has a major and a minor axis. $E\phi$ assumes the direction of the major axis while $E\theta$, which is orthogonal to $E\phi$, is in the direction of the minor axis of the polarization ellipse. In terms of the circular fields, the linear components are:

$$\begin{aligned} E\phi &= \text{Re}(E_L + E_R) = (A + B) \cos Wt \\ \text{that is, } C &= A + B \end{aligned} \quad (3)$$

$$\begin{aligned} \text{and } E\theta &= \text{Im}(E_L + E_R) = (A - B) \sin Wt \\ \text{that is, } D &= A - B. \end{aligned} \quad (4)$$

It may be noted that if $B=0$, then $A=C=D$. In this case (1) reduces to left circular as represented in (2). Similarly for $A=0$, (1) reduces to right circular polarization. When $A=B$, $E\theta$ disappears which is the case for linear polarization.

The ellipticity of a wave may be expressed as a ratio of the maximum value of field to the minimum value of field:

$$E\phi/E\theta = C/D = \frac{A+B}{A-B}. \quad (5)$$

This ratio ranges from infinite for linear polarization to one for circular polarization.

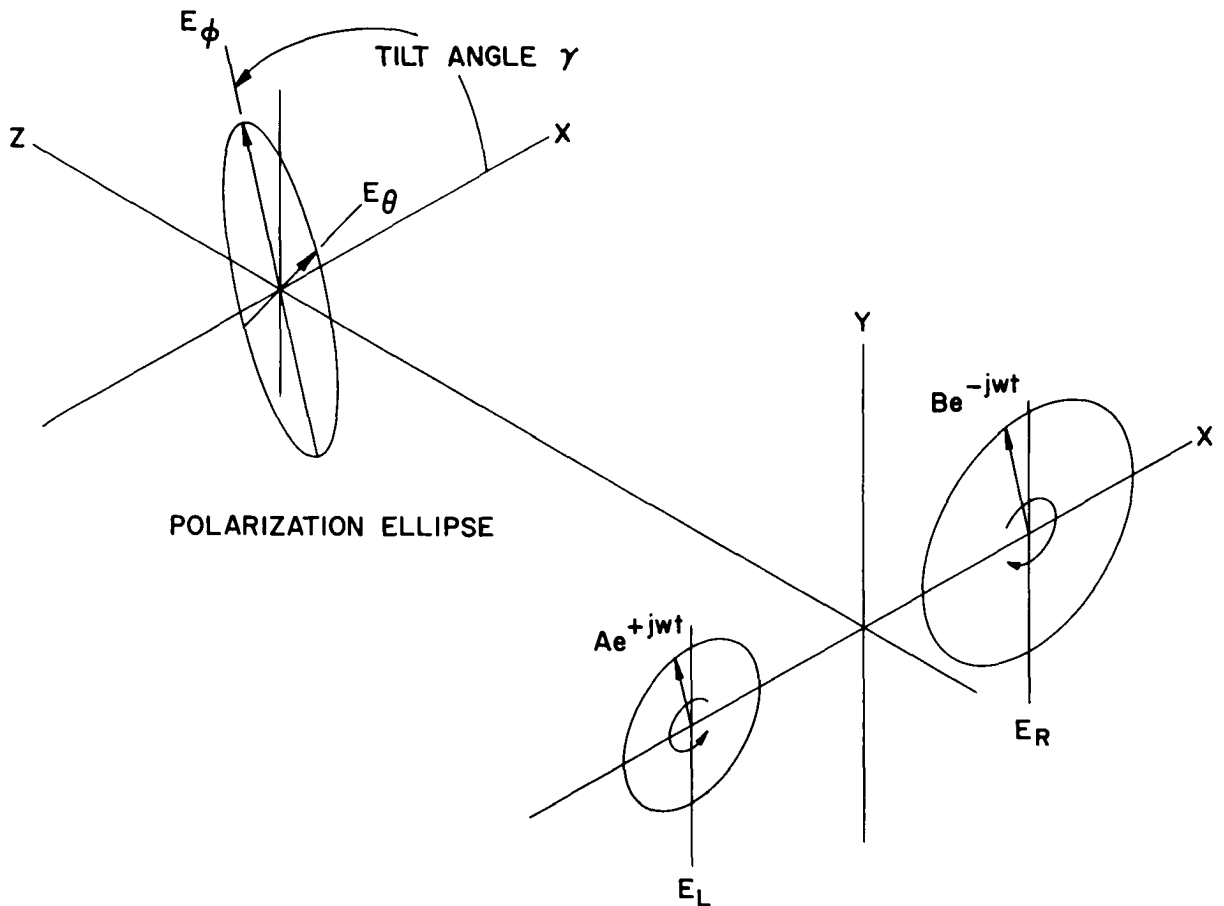


FIGURE 1.—The polarization ellipse and its circular components.

POLARIZATION ANALYZER

Automatic polarization analysis can be accomplished with the aid of the circuitry shown in Figure 2. In order to analyze polarization characteristics it is first necessary to transfer the wave in space into fields in transmission lines. For the purpose of this paper, an antenna such as a turnstile array is selected which separates a received signal into its two circular components. For a given orientation of the receiving antenna, the two fields in the transmission lines will be in-phase for only one orientation of the major axis of the incident polarization ellipse. This direction of the major axis may be established as a reference direction. As the polarization ellipse tilts in space from this reference direction, the phase of the field in one transmission line will be advanced an angle equal to the angle of tilt while

the phase of the field on the other transmission line will be retarded a like amount. The signal intensities in the transmission lines from the antenna may be written:

$$\begin{aligned} E_L &\approx A \cos (\omega t + \gamma) \\ E_R &\approx B \cos (\omega t - \gamma) \end{aligned} \quad (6)$$

where γ is the angle between the major axis of the polarization ellipse and a reference direction.

In Figure 1 the reference direction is shown as the X-axis. It may be noted that the reference direction may be changed at will by an observer who can simply rotate the receiving antenna.

A portion of each field in (6) is detected to provide quantities proportional to component amplitudes. The remainder goes to a phase comparator to be used for determination of the ellipse tilt angle.

RATIO OF AMPLITUDES

The detected amplitudes P_L and P_R shown in Figure 2 are proportional to the squares of the absolute amplitudes of left and right circular components. Recalling that the ellipticity ratio is given by $E\phi/E\theta$, a computer which solves the equation

$$\frac{A+B}{A-B} \approx \frac{\sqrt{P_L} + \sqrt{P_R}}{\sqrt{P_L} - \sqrt{P_R}} \quad (7)$$

provides two factors of interest concerning polarization characteristics of a wave: (1) degree of ellipticity (2) dominant sense of circular polariza-

tion (by noting whether the ratio is positive or negative).

It is somewhat simpler to determine the ratio P_R/P_L and in some applications is sufficient. The degree of ellipticity—for a limited range of values of P_R/P_L —can be determined if required with the aid of Figure 3. The dominant sense of circular polarization is determined by noting whether the ratio is greater than or less than one.

DETERMINATION OF TILT ANGLE

The tilt angle of the polarization ellipse is the remaining unknown and is determined easily by measuring the phase difference between the two senses of circular polarization with a microwave phase comparator.² Assuming the signals of (6) as inputs to the phase comparator shown in Figure 2, the signal intensity at terminal a is proportional to:

$$Ea \approx A \cos(Wt + \gamma) + B \cos(Wt - \gamma) \quad (8)$$

$$\approx (A+B) \cos Wt \cos \gamma + (A-B) \sin Wt \sin \gamma.$$

The detected signal is proportional to the square of the signal intensity:

$$Pa \approx Ea^2 \quad (9)$$

$$\approx (A^2 + B^2 + 2AB) \cos^2 Wt \cos^2 \gamma$$

$$+ (A^2 + B^2 - 2AB) \sin^2 Wt \sin^2 \gamma$$

$$\approx (A^2 + B^2) (\cos^2 Wt \cos^2 \gamma + \sin^2 Wt \sin^2 \gamma)$$

$$+ 2AB (\cos^2 Wt \cos^2 \gamma - \sin^2 Wt \sin^2 \gamma)$$

$$Pa \approx (A^2 + B^2) + 2AB \cos 2\gamma.$$

Similarly, the other detected outputs from the phase comparator are:

$$Pb \approx A^2 + B^2 - 2AB \cos 2\gamma$$

$$Pc \approx A^2 + B^2 + 2AB \sin 2\gamma$$

$$Pd \approx A^2 + B^2 - 2AB \sin 2\gamma. \quad (10)$$

Differencing the outputs from each of the hybrids produces:

$$Pa - Pb \approx 4AB \cos 2\gamma$$

$$Pc - Pd \approx 4AB \sin 2\gamma. \quad (11)$$

Displaying these difference amplitudes orthogonally, i.e., $Pa - Pb$ on the X-axis and $Pc - Pd$ on the Y-axis of an X-Y indicator, which may be

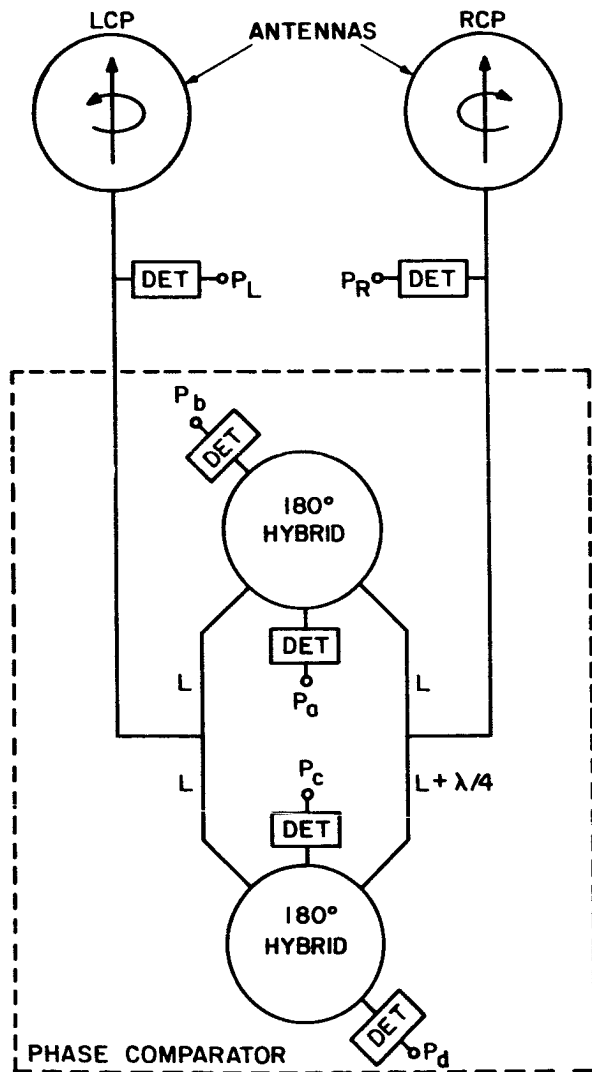


FIGURE 2.—Automatic Polarization analyzer.

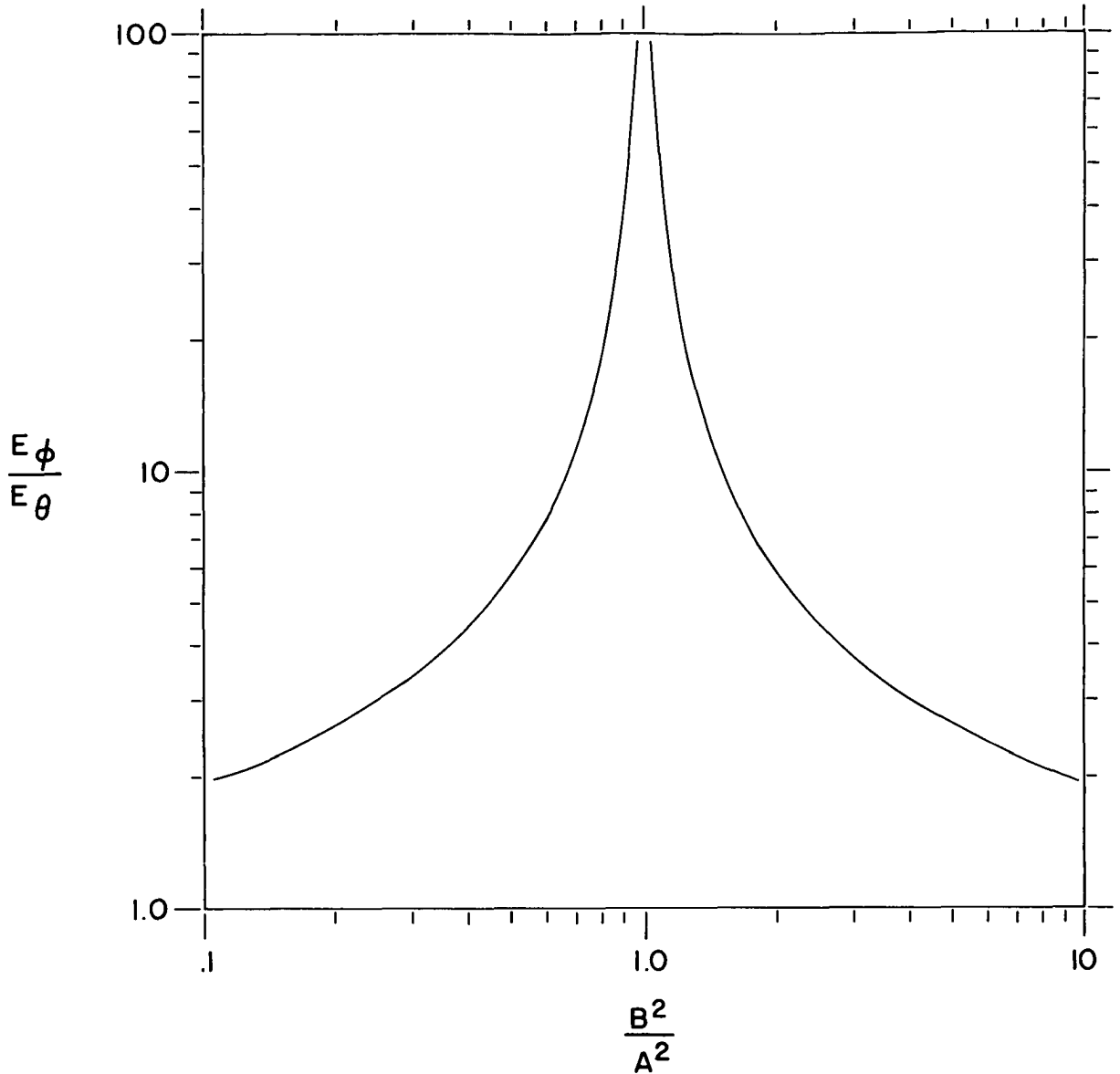


FIGURE 3.—Ellipticity ratio vs. ratio of detected circular fields.

an X-Y recorder or oscilloscope, yields the indication:

$$Pt \approx 4AB (\cos 2\gamma + j \sin 2\gamma) \quad (12)$$

$$4ABe^{j2\gamma}.$$

Thus, the product of the amplitude of the two circular components is the radius to a circle which is described by the phase difference (2γ) between the circular components. See Figure 4. If either A or B is zero, which means that only one

sense of circular polarization is received, the radius of the phase circle goes to zero, showing no elliptical polarization present.

The X-Y indicator here as a computer, taking as inputs the functions of an angle and displaying the angle itself.

It may be observed that as the tilt angle γ ranges through 180 degrees, a complete circle is traced on the indicator, meaning that the relative phase of the fields in the transmission lines has

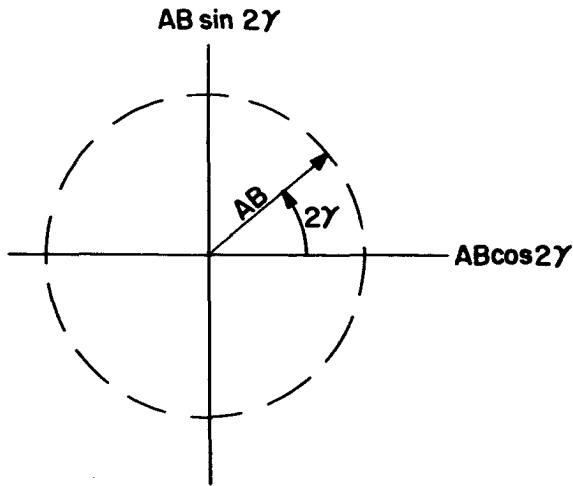


FIGURE 4.—Phase comparator output displayed on X-Y indicator.

undergone a change of 360 degrees. The X-Y indicator thus does not automatically display the tilt angle in the sense that it yields a change in displayed angle of twice the change in tilt angle.

Tilt angle can be automatically displayed if quantities proportional to (11) are applied to a

two-phase motor whose rotor is connected to a pointer through a 2:1 gear reducer.

CONCLUSION

It has been shown how to determine automatically and continuously all of the polarization characteristics of a wave. The degree of ellipticity and dominant sense of rotation are obtained from a ratio derived from the circular component amplitudes and the direction of polarization is obtained from a measurement of the phase difference between the two components. The method described requires no moving parts or active phasing devices.

REFERENCES

1. H. G. BOOKER, V. H. RUMSEY, G. A. DESCHAMPS, M. L. KALES, and J. I. BOHNEPT, "Techniques for Handling Elliptically Polarized Waves with Special Reference to Antennas," Proceedings of the IRE, Vol. 39, No. 5, pp. 533-552, May, 1951.
2. J. A. KAISER, H. B. SMITH, JR., W. H. PEPPER, and J. H. LITTLE, "An Automatic Microwave Phase Comparator," IRE Transactions on MTT, MTT-10(6), pp. 548-550, November, 1962.

PHASE COMPARISON PROPERTIES OF THE DUAL MODE SPIRAL WITH APPLICATION TO DIRECTION FINDING

J. A. KAISER AND J. H. LITTLE

SUMMARY

A spiral antenna excited in any two adjacent radiation modes simultaneously has built-in phase comparison properties useful in direction finding. The information derived from the phase comparison on the spiral allows unambiguous determination of direction of arrival of signals from anywhere in a hemisphere. The phase comparison properties are frequency insensitive when consideration can be confined to the two adjacent modes only.

This paper will examine the phase comparison properties of the dual mode spiral and show how they can be applied to a very simple, passive direction finder system. In addition, the problem of a rotating azimuthal reference plane with change in frequency is treated.

In general, the spiral antenna can be excited in any of an infinite number of odd and even modes (Ref. 1), although in practice the spiral size and construction place practical restrictions on the number of modes possible. The phase difference between any two adjacent modes is in every case 2π radians. Since the change in phase for each mode is uniform with change in the azimuthal angle ϕ , a measure of the phase difference between modes can be related directly to the coordinate angle ϕ . In this paper, attention will be confined to consideration of the first and second radiation modes of the Archimedean spiral, although the argument can readily be extended to any two adjacent modes.

In lieu of exact field equations for the spiral antenna, it may be assumed that, as a first approximation, the first mode radiation field of the spiral can be represented (neglecting time variations) by $Be^{j\phi} \cos \theta$ and the second by $Ae^{j2\phi} \sin \theta$, where A is the peak of the second mode radiation

pattern, B is the peak of the first mode radiation pattern, $0 \leq \phi \leq 2\pi$, and $0 \leq \theta \leq \pi/2$. See Figure 1. An incident signal will in general excite currents in each of the first two modes. The network input currents induced by the incident signal will be in-phase at $\phi = 0^\circ$, as is required by the method of defining the radiation fields of the first and second modes. That is, the first mode is excited in an annular region of the spiral of $\lambda/2\pi$ mean radius and the second mode in an annular region of λ/π mean radius. Both modes will be in-phase along one radial from the spiral center, and this direction is defined as $\phi = 0^\circ$.

If the phase changes due to path lengths are designated by α_1 for the first mode currents and α_2 for the second mode currents, the total current output at the antenna terminal will be

$$I = I'_1 + I'_2 \approx Be^{j(\phi_1 + \alpha_1)} \cos \theta + Ae^{j(\phi_2 + \alpha_2)} \sin \theta, \quad (1)$$

where I'_1 & I'_2 are the complex currents of the first and second modes appearing at the spiral input terminal and $\phi_2 = 2\phi_1$.

Neglecting the radial component of filament length, which in most practical spirals is quite negligible, the phase changes are given by (ref. 1)

$$\alpha_1 = \frac{a\pi}{\lambda} \psi_1^2 \quad (2)$$

$$\alpha_2 = \frac{a\pi}{\lambda} \psi_2^2, \quad (3)$$

where a = constant which determines pitch or spiral

ψ_1 = angular measure from spiral center to 1st mode ring

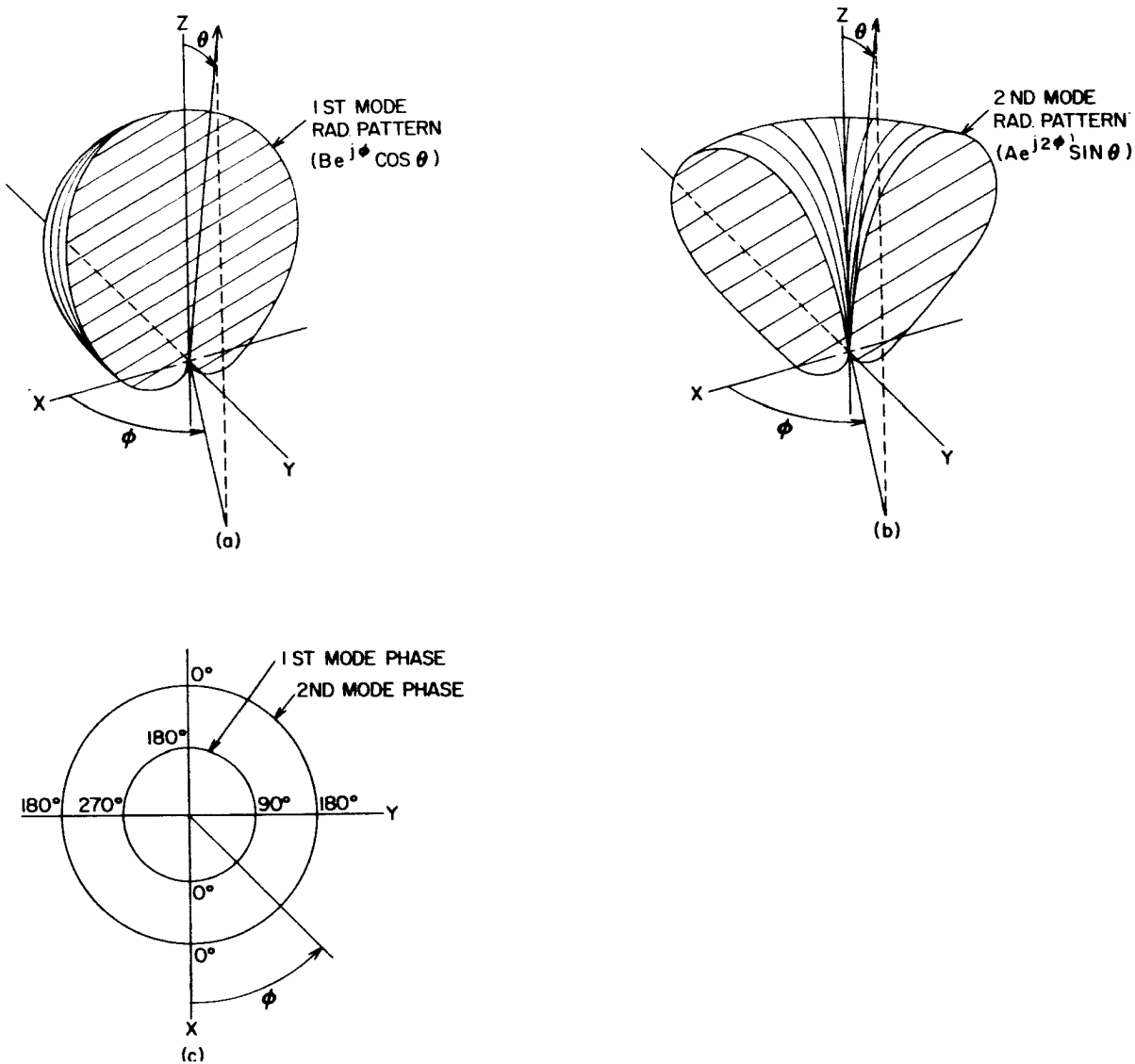


FIGURE 1.—Coordinate system of dual mode spiral.

ψ_2 = angular measure to 2nd mode ring.

With the current of (1) incident, the detected power at the antenna terminal is proportional to

$$P(\text{det}) \approx A^2 \sin^2 \theta + B^2 \cos^2 \theta + AB \sin 2\theta \cos [(\phi_2 - \phi_1) + (\alpha_2 - \alpha_1)] \quad (4)$$

and from (2) and (3)

$$\alpha_2 - \alpha_1 = \frac{\alpha\pi}{\lambda} (\psi_2^2 - \psi_1^2). \quad (5)$$

The factor $(\alpha_2 - \alpha_1)$ is a constant angle for a given

wavelength. On the other hand, $\phi_2 - \phi_1$ is constantly changing as ϕ ranges through 360 degrees, thereby producing changes in the detected output. Variation in detected power as a function of ϕ for $A=B$ and $\theta=45^\circ$ is shown in Figure 2a and $\theta=15^\circ$ in Figure 2b.

Equation (4) is not a single-valued function of ϕ throughout 360° . Such uniqueness is available from the detected outputs of a four filament spiral or from a two filament spiral with addition signal processing before detection (ref. 2, 3). The filaments of a four arm Archimedean spiral (when

starting from the spiral center) may be defined by

$$\begin{aligned} r_1 &= a\psi e^{j0} \\ r_2 &= a\psi e^{j\pi/2} \\ r_3 &= a\psi e^{j\pi} \\ r_4 &= a\psi e^{j3\pi/2}. \end{aligned} \quad (6)$$

By making the appropriate changes to the phases (exponents) of the induced signals in (1) in accordance with (6) the currents at the spiral input terminals become

$$\begin{aligned} I_1 &\approx Ae^{j(\phi_1+\alpha_1)} \cos \theta + Be^{j(\phi_2+\alpha_2)} \sin \theta \\ I_2 &\approx Ae^{j(\phi_1+\pi/2+\alpha_1)} \cos \theta + Be^{j(\phi_2+\pi+\alpha_2)} \sin \theta \\ I_3 &\approx Ae^{j(\phi_1+\pi+\alpha_1)} \cos \theta + Be^{j(\phi_2+2\pi+\alpha_2)} \sin \theta \\ I_4 &\approx Ae^{j(\phi_1+3\pi/2+\alpha_1)} \cos \theta + Be^{j(\phi_2+3\pi+\alpha_2)} \sin \theta. \end{aligned} \quad (7)$$

With square-law detectors at each of the spiral terminals, the detected outputs are

$$\begin{aligned} P_1 &\approx A^2 \sin^2 \theta + B^2 \cos^2 \theta + AB \sin 2\theta \cos [(\phi_2 - \phi_1) + (\alpha_2 - \alpha_1)] \\ P_2 &\approx A^2 \sin^2 \theta + B^2 \cos^2 \theta - AB \sin 2\theta \sin [(\phi_2 - \phi_1) + (\alpha_2 - \alpha_1)] \\ P_3 &\approx A^2 \sin^2 \theta + B^2 \cos^2 \theta - AB \sin 2\theta \cos [(\phi_2 - \phi_1) + (\alpha_2 - \alpha_1)] \\ P_4 &\approx A^2 \sin^2 \theta + B^2 \cos^2 \theta + AB \sin 2\theta \sin [(\phi_2 - \phi_1) + (\alpha_2 - \alpha_1)]. \end{aligned} \quad (8)$$

Figure 3 shows the variations in the four detected outputs as a function of ϕ when $A=B$ and $\theta=45^\circ$. Differencing detected outputs of opposite pairs produces

$$P_1 - P_3 \approx 2 AB \sin 2\theta \cos [(\phi_2 - \phi_1) + (\alpha_2 - \alpha_1)] \quad (9)$$

$$P_4 - P_2 \approx 2 AB \sin 2\theta \sin [(\phi_2 - \phi_1) + (\alpha_2 - \alpha_1)]. \quad (10)$$

Displaying (9) and (10) orthogonally to one another, e.g., on an X - Y indicator (ref. 4), in which (9) is applied to the X -axis and (10) to the Y -axis, yields an indicated power

$$P_i \approx 2 AB \sin 2\theta e^{j[(\phi_2 - \phi_1) + (\alpha_2 - \alpha_1)]}, \quad (11)$$

which is single valued in $(\phi_2 - \phi_1)$ throughout 360 degrees. This determines the azimuthal direction of an incoming signal with no ambiguities. Figure 4 illustrates the display obtained on an X - Y indicator; the amplitude function in (11) is a radius to the circle described by ϕ and α .

Since (5) shows a linear change in $(\alpha_2 - \alpha_1)$ with change in wavelength, there is an effect of rotating the radial along which the currents of the two modes are in-phase with change in frequency. For a spiral mounted in a rectangular coordinate system, where the positive X -axis lies in the

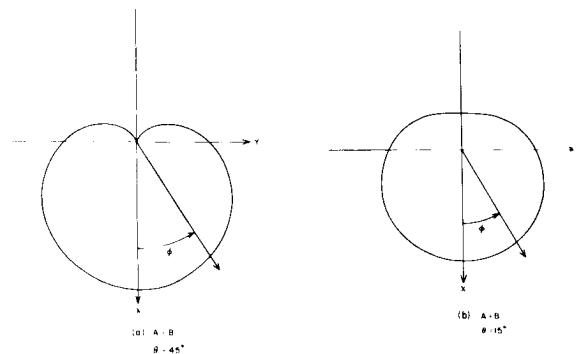


FIGURE 2.—Detected output from one filament of dual mode spiral vs. ϕ .

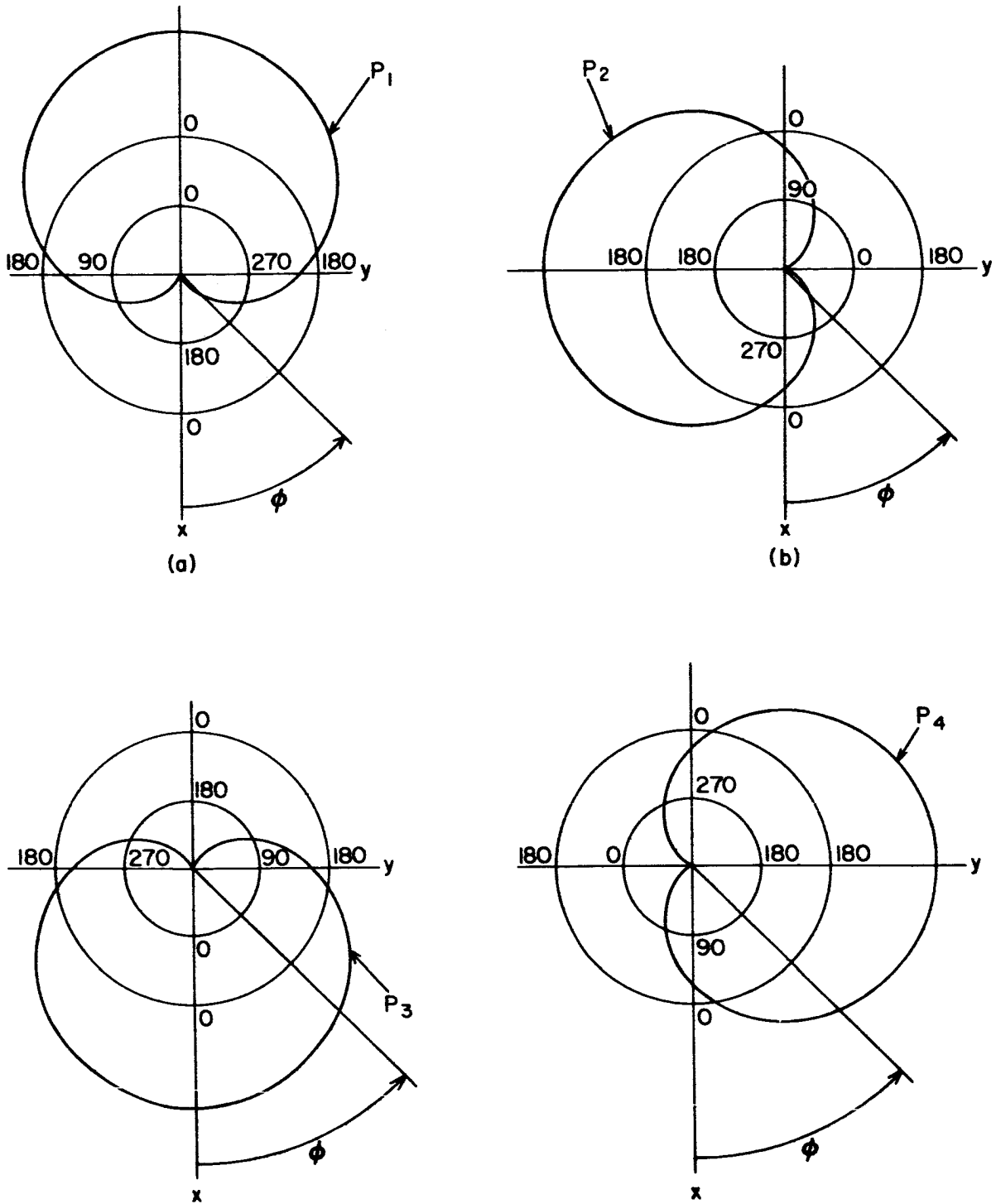


FIGURE 3.—Detected outputs of four wire spiral vs. ϕ (When $A = B$ and $\theta = 45^\circ$).

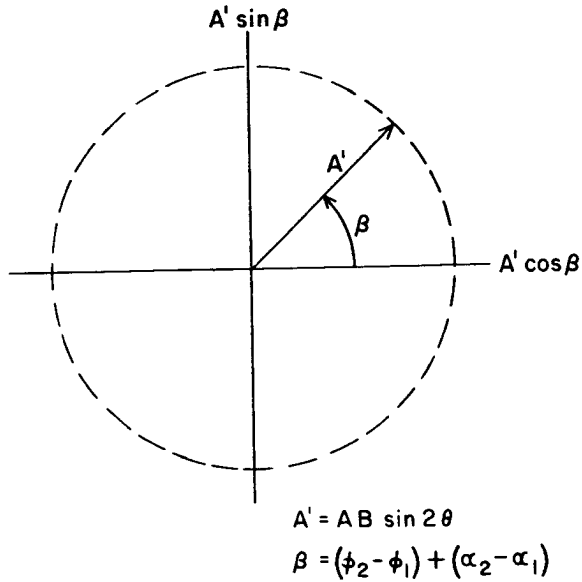


FIGURE 4.—Detected outputs of dual mode spiral displayed on X-Y indicator.

direction of in-phase currents on the spiral, the coordinate system will rotate with changing frequency because the X-axis rotates. The direction of rotation will be in the same direction as the winding sense of the spiral for increasing wavelength.

REFERENCES

1. J. A. KAISER, "The Archimedean Two-Wire Spiral Antenna," IRE Transactions on Antennas and Propagation, AP-8(3), May, 1960.
2. J. A. KAISER, H. B. SMITH, JR., W. H. PEPPER, J. H. LITTLE, "A Passive Automatic Elevation-Azimuth Direction Finder," NASA TN D-1763.
3. J. A. KAISER, H. B. SMITH, JR., W. H. PEPPER, J. H. LITTLE, "A Passive Automatic Direction Finder," Proc. IRE 9th ECCANE, Baltimore, Maryland, October, 1962.
4. J. A. KAISER, H. B. SMITH, JR., W. H. PEPPER, J. H. LITTLE, "An Automatic Microwave Phase Comparator," IRE Transactions on MIT, vol. MTT-10, No. 6, November, 1962.

DEVELOPMENT OF OPTIMUM FRAME SYNCHRONIZATION CODES FOR GODDARD SPACE FLIGHT CENTER PCM TELEMETRY STANDARDS*

JESSE L. MAURY, JR. AND FREDERICK J. STYLES

In PCM telemetry the acquisition of erroneous data or the loss due to false synchronization can be minimized through use of frame sync codes having the lowest probability of false sync (i.e., 'Optimum Frame Sync Codes'). In this paper are described the analysis of the criteria of optimality for frame sync codes, the application of these criteria to the derivation of formulas for evaluating the probability of false sync for any pattern, and the development of a high speed computer program for isolating the optimum frame sync code of any length n from the 2^n possible patterns. Also included are the Optimum Frame Sync Codes of length 7 to 30 which were developed as a result of this study and which have been presented for inclusion in the GSFC PCM Telemetry Standards.

INTRODUCTION

The use of a PCM telemeter, as with most time-division multiplexed methods, requires time synchronization of the ground receiving equipment and the spacecraft transmitting equipment in order to process the data. This synchronization is accomplished in two distinct phases (References 1, 2, 3, 4): 1) Bit sync, which distinguishes the individual bits of the data stream, 2) Frame sync, which defines the repetitive data blocks (frames) and therefore identifies (e.g. using a counter) each data word within the frame. The investigation assumes that bit sync has been acquired and can be maintained beyond signal-to-noise ratios for which the data is no longer useful. This study is concerned with the selection of the 'optimum' frame sync code of any given length, n , from $n=7$ to $n=32$ taken from the 2^n different binary patterns of that length.

In the past the development of so called optimum frame synchronization codes has, in general, followed a fairly uniform approach: Once the criteria for optimality have been defined, binary patterns which best fit these criteria are generated. There has been, for the most part, agreement as to the general criteria for optimality.

With the notable exception of Goode and Philips (Reference 5), the approach taken in the literature is that developed in the following sections. However, with various investigators, different methods of development as well as differences in the extent and application of the criteria have resulted in the publication of suggested frame sync codes with descriptions ranging from 'good' through 'preferred' and 'better' to 'best.'

Faced with this plurality of frame sync codes and primarily interested in arriving at a set of standard codes for general use in PCM telemetry, the authors felt that only through the application of an exhaustion technique (i.e., the examination of all binary patterns of a given length against specified criteria) could the truly optimum frame sync codes be established. Since such a technique is feasible only with the aid of a high speed computer, the criteria and analysis which follow were developed and programmed for the IBM 7094. As a result of this effort an optimum set of frame sync codes of length 7 through 30 have been determined and presented for inclusion in the Goddard Space Flight Center PCM Telemetry Standards.

CRITERIA

In order to understand the use of the word 'optimum' here, the criteria used in the frame sync code selection must be presented. These

*Published in the *Proceedings of the 1964 National Telemetering Conference*, June 2, 3, 4, 1964, Los Angeles, California, Session III, Multiplexing and Synchronizing, Pages 1-10.

criteria may be defined in terms of the PCM frame synchronization process as follows:

Once the ground station has begun receiving PCM telemetry from a particular spacecraft and is distinguishing each individual bit of the data (bit sync), it must locate in the continuous data stream the beginning of the major repetitive data block as indicated by a multi-bit code (frame sync). The recognition process is accomplished by a cross correlation technique (called auto-correlation by some authors, References 3 and 5) in which a code recognizer in the ground station examines the incoming bit stream for the frame sync code. (The term 'recognition pattern' will be used to designate the pre-set configuration of bits within the recognizer against which the recognizer makes its comparisons.) The recognizer has the capability of allowing a selected number of errors in the received frame sync code to compensate for transmission noise, receiver noise, etc. The incoming data, with the exception of the frame sync code, is assumed to be random. As the frame sync code is received, it is preceded and followed by random bits and the recognizer must examine three categories of bit patterns.

The first is completely random data and occurs when no portion of the frame sync code is seen by the recognizer. As the bit stream steps past the recognizer in a bit-by-bit fashion, the recognizer examines each code-length group of random bits. This will be called the 'random' region.

The second category of bit patterns seen by the recognizer is made up of both random data bits and one or more bits of the frame sync code, and can be called the 'overlap' region; that is, the actual frame sync code 'overlaps' into the group of bits being examined by the recognizer. As the bit stream continues to pass the recognizer, the number of bits of the frame sync code seen by the recognizer (degree of overlap) increases until the recognizer is examining one data bit and all but one bit of the frame sync code. This position is called the maximum degree of overlap. One more step and the complete frame sync code is in the recognizer. Another step and maximum degree of overlap exists again as the frame sync code begins to leave the recognizer.

The third category of bit patterns is the once-per-frame occurrence of the true frame sync code completely occupying the recognizer.

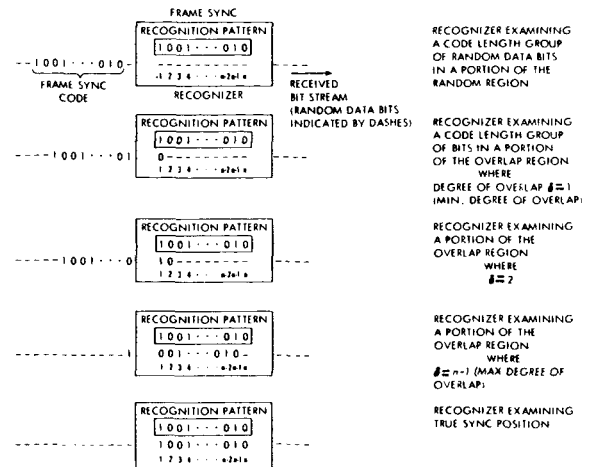


FIGURE 1.—Frame sync recognition process.

Figure 1 shows the movement of the data stream through the frame sync recognizer illustrating the three categories of bit patterns examined during the recognition process.

At any time there is a certain probability of the recognizer giving a sync indication. For both the random region and the true sync position, this probability is independent of the actual frame sync code used and therefore can form no basis for the selection of the optimum frame sync code. On the other hand the probability of a frame sync indication in the overlap region is a function of the particular code used since it is dependent upon the number of bits in agreement between the incoming bit stream and the recognition pattern within the recognizer.

In any overlap position, while each random data bit has a 50 percent probability of agreement independent of the recognition pattern, each overlapping bit of the incoming frame sync code precisely agrees or disagrees with that bit of the recognition pattern against which it is being compared. The number of overlapping code bits in agreement with the corresponding bits of the recognition pattern in a given overlap position will be designated as α . Neglecting for the moment the random data bits, the search for sync in any degree of overlap, δ , is a comparison of the first δ bits of the frame sync code (in the bit stream) with the last δ bits of the frame sync code (in the recognition pattern). If this comparison is made for each degree of overlap from one to maximum ($\delta =$ one less than the code length) an 'agreement'

or auto-correlation statement is developed which, when combined with the respective probabilities for the random data bit portion of the overlap region, allows the computation of the probability of false sync recognition over the entire overlap region. This probability, since it is, as indicated above, a function of the particular binary pattern arrangement, determines the optimum frame sync code from the 2^n different patterns of a given length n . The optimum code for a given length is that pattern with the smallest total probability of false sync recognition over the entire overlap region.

The above discussion of the recognition process has neglected the effects of noise expressed as the probability, η , of bit change and of the number, ϵ , of errors allowed by the recognizer to compensate for the effects of noise. These parameters will be elaborated upon in the analysis of the frame sync problem to follow.

ANALYSIS

For each degree of overlap, δ , there exists a probability, $F\delta$, of false sync occurring. The probability of not having false sync for a given δ is, then, $1 - F\delta$. By developing the product of all such terms over the full range from $\delta=1$ (the minimum degree of overlap) to $\delta=n-1$ (the maximum degree of overlap), the probability of not having false sync during the overlap period can be determined. One minus this product is the probability, F_Φ , that false sync will occur anywhere in the overlap region. Thus,

$$F_\Phi = 1 - \prod_{\delta=1}^{n-1} (1 - F\delta)$$

For false sync to occur, recognition must occur in both the δ code bits and the $n-\delta$ random bits under the recognizer. If the probability of recognition occurring in the δ bits is designated as J and the probability of recognition occurring in the corresponding random bits under the recognizer is designated by K , $F\delta$ becomes a product of J and K . However, the recognizer may have been set to allow a certain number, ϵ , of errors in its sync recognition. The errors allowed by the recognizer may be delegated in part or in whole to either the δ bits or the random bits. If i is used

to designate the number of errors to be allotted to the δ bits, $\epsilon-i$ errors may be allotted to the random bits. Since there exists a probability of recognition occurring in the δ bits without any errors allowed, with one error allowed, with two errors allowed, etc., to the limit of ϵ , it is necessary to sum the product of the probabilities J and K over the entire range $i=0$ to $i=\epsilon$. Obviously it would be meaningless to allow i to exceed δ (i.e., to allot more errors allowed to the δ portion than there are bits in that portion). Thus, it is necessary to modify the upper limit when this happens. Hence,

$$F_\Phi = \sum_{i=0}^{\epsilon} J_i K_{\epsilon-i}$$

Consider first, $K_{\epsilon-i}$, the probability of recognition occurring in just the random bits. Here, the probability of recognition with no errors allowed would be $\frac{1}{2}^{n-\delta}$. The probability with one error allowed would be $\frac{1}{2}^{n-\delta}$ multiplied by $n-\delta$, the number of different configurations of all but one of the random bits agreeing with the corresponding recognition pattern bits. Or, in other words,

$$\frac{1}{2^{n-\delta}} \binom{n-\delta}{1}$$

If two errors are allowed in the random bits, the probability of recognition in this portion is

$$\frac{1}{2^{n-\delta}} \binom{n-\delta}{2}. \text{ Etc.}$$

Since each of the above is a possible condition, the probability when $\epsilon-i$ errors are allowed in the random bits is the sum of these terms from 0 to $\epsilon-i$. Again, it would be meaningless to allow more errors in the random area under the recognizer than there are bits (i.e., for $\epsilon-i$ to exceed $n-\delta$). Thus, the limit of the sum must be the number of bits in the region when this occurs. Therefore,

$$K_{\epsilon-i} = \sum_{k=0}^{\epsilon-i} \frac{1}{2^{n-\delta}} \binom{n-\delta}{k}$$

The analysis of J_i , the probability of recognition occurring in just the δ code bits is somewhat more complex. Here there exist two known values, the degree of agreement, α , and the counterpart degree of disagreement, $\delta - \alpha$. Since these values are based on the relationship between the configurations of the recognition pattern and the incoming frame sync code, the possibility of bit changes in the code due to noise occurring between transmission and reception of the code must be considered. Designating the probability of any one bit being changed by noise as η , the probability that some quantity, j , of bits will be changed is η^j . These j bit changes may occur in either the agreement bits or the disagreement bits. If j agreement bits are changed to disagreement and, remembering that i is the number of errors allowed in this term, $i - j$ disagreement bits may remain as they are. Since the probability of a bit being changed is η , the probability of a bit not being changed is $1 - \eta$, and the probability that $i - j$ bits will not be changed is $(1 - \eta)^{i - j}$. When j changes are allotted to the agreement bits, the remaining $\alpha - j$ agreement bits must remain as they are. The probability of this is $(1 - \eta)^{\alpha - j}$. For sync to occur in the δ region the $\delta - \alpha$ disagreement bits, except for $i - j$ of them, must be changed (by noise) to agreement bits. That is, $(\delta - \alpha) - (i - j)$ bits. The probability of this occurring is

$$\eta^{[(\delta - \alpha) - (i - j)]}$$

In recapitulation,

$$\eta^j$$

is the probability that j agreement bits will be changed by noise,

$$(1 - \eta)^{\alpha - j}$$

is the probability that $\alpha - j$ agreement bits will not be changed by noise,

$$(1 - \eta)^{i - j}$$

is the probability that $i - j$ disagreement bits will not be changed by noise,

$$\eta^{[(\delta - \alpha) - (i - j)]}$$

is the probability that $(\delta - \alpha) - (i - j)$ disagreement bits will be changed by noise.

It is now necessary to consider the number of different configurations wherein the above probabilities can be realized for a given δ . Consider first the number of ways that j of the agreement bits can be changed leaving $\alpha - j$ unchanged. This is

$$\binom{\alpha}{j}$$

Thus, the probability of j changes in the agreement bits while $\alpha - j$ agreement bits remain unchanged is:

$$\binom{\alpha}{j} \eta^j (1 - \eta)^{\alpha - j}$$

The second consideration is the number of ways in which $i - j$ disagreement bits can remain unchanged while the remaining $(\delta - \alpha) - (i - j)$ disagreement bits are changed to agreement. This is

$$\binom{\delta - \alpha}{i - j}$$

Thus, the probability of the event of $(\delta - \alpha) - (i - j)$ changes in the $\delta - \alpha$ disagreement bits while $i - j$ disagreement bits remain unchanged is:

$$\binom{\delta - \alpha}{i - j} (1 - \eta)^{i - j} \eta^{[(\delta - \alpha) - (i - j)]}$$

Since there exists a probability of recognition occurring in the δ bits with j changes occurring in the α agreement bits and the remaining needed $(\delta - \alpha) - (i - j)$ changes occurring in the $\delta - \alpha$ disagreement bits, recognition occurring with $j + 1$ changes in the α bits and $(\delta - \alpha) - (i - j + 1)$ changes occurring in the $\delta - \alpha$ bits, etc., it is necessary to sum these terms over every possible value of j . The range of these values can easily be recognized by noting it would be meaningless to allow more bit changes in the disagreement bits than there are disagreement bits. Thus, one limit can be set by insuring that the number of disagreement bits, $\delta - \alpha$, is greater than or equal to the number of changes, $i - j$, being considered for the disagreement bits; that is, $\delta - \alpha \geq i - j$ or $j \geq i - (\delta - \alpha)$. Also it is meaningless to allow a negative number of changes in the agreement bits. Therefore, $j \geq 0$. The second limit can be set to preclude allowing more changes to occur in the agreement bits than there are such bits; that is, $j \leq \alpha$. In addition it would be useless to allow the

number of changes in the agreement bits to exceed j , the number of errors allowed in this term, since this would guarantee a zero probability of recognition occurring in the δ portion. Thus, $j \leq i$. Therefore,

$$i \text{ if } i \leq \alpha$$

$$\alpha \text{ if } i > \alpha$$

$$J_i = \sum \binom{\alpha}{j} \eta^j (1-\eta)^{\alpha-j} \binom{\delta-\alpha}{i-j} (1-\eta)^{i-j} \eta^{(\delta-\alpha)-(i-j)}$$

$$j = i - (\delta - \alpha) \text{ if } i - (\delta - \alpha) \geq 0$$

$$j = 0 \text{ if } i - (\delta - \alpha) < 0$$

Although this form of J_i is more amenable to analysis, it can be reduced to a simpler form.

$$i \text{ if } i \leq \alpha$$

$$\alpha \text{ if } i > \alpha$$

$$J_i = \sum \binom{\alpha}{j} \binom{\delta-\alpha}{i-j} (1-\eta)^{(\alpha+i-2j)} \eta^{(\delta-\alpha-i+2j)}$$

$$j = i - \delta + \alpha$$

It is now possible to develop a complete formula for evaluating F_Φ , the probability of false sync occurring anywhere in the overlap region. First, examine the formula

$$F_\Phi = 1 - \prod_{\delta=1}^{n-1} (1 - F_\delta) \text{ where } F_\delta = \sum_{t=0}^{\epsilon} J_i K_{t=i}$$

$$\epsilon \text{ if } \epsilon \leq \delta$$

$$\delta \text{ if } \epsilon > \delta$$

Since the values of F_δ are quite small, a valid approximation of F_Φ is

$$F_\Phi \approx \sum_{\delta=1}^{n-1} F_\delta$$

Therefore,

$$F_\Phi \approx \sum_{\delta=1}^{n-1} \sum_{t=0}^{\delta} \left(\sum_{j=i-\delta+\alpha}^{\alpha} \binom{\alpha}{j} \binom{\delta-\alpha}{i-j} (1-\eta)^{(\alpha+i-2j)} \right)$$

$$\left[\eta^{(\delta-\alpha-i+2j)} \right] \left\{ \frac{1}{2^{n-\delta}} \sum_{k=0}^{n-\delta} \binom{n-\delta}{k} \right\}$$

$$\epsilon - i \text{ if } \epsilon - i \leq n - \delta$$

$$n - \delta \text{ if } \epsilon - i > n - \delta$$

Initially this formula was developed for determining F_Φ only for patterns of length seven. It was then extended to the general case of any length n . In the development and use of this formulation, consideration of the probability of false sync was made only for that period of overlap when the code was entering the recognizer. This is of little consequence since the probability of false sync being realized as the code leaves the recognizer is exactly the same as the probability when entering the recognizer. Thus, where both entrance and exit are considered as belonging to the overlap region, the probability of false sync is merely twice the above value.

SEARCH

Early in this study several characteristics of binary patterns became apparent. First, although the value of η , the probability of a bit being changed by noise, does change the probability of false sync, it does not affect the relative merit of the patterns. That is, for a given length n and a given number of errors allowed, ϵ , the ordering of all patterns by their related F_Φ values is not changed by changing the value of η . It also became apparent that, though ϵ does affect the selection of the optimum frame sync code of a given length, the primary factor in the relative merit of the patterns is the configuration of the agreement or α values over the range of $\delta=1$ to $\delta=n-1$. This configuration of α values for a given pattern is designated as the agreement vector of that pattern. An example of its development is given in Figure 2 for a seven bit pattern. The random region has been designated by dashes.

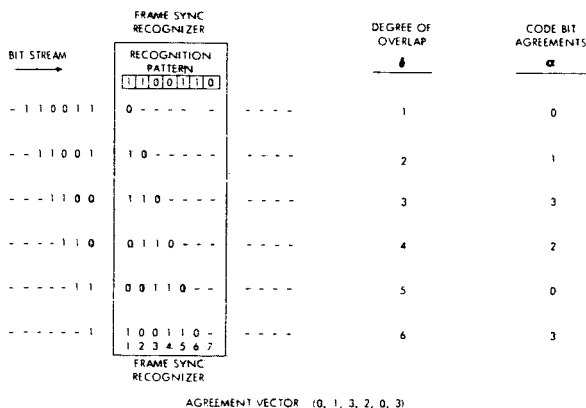


FIGURE 2.—Agreement vector development.

The last characteristic noted was that for any given pattern there exist three related configurations. These three related patterns are the complement wherein each 1 bit of the pattern is replaced by a 0 while each 0 is replaced by a 1, the mirror pattern which is developed by reversing the pattern, and the mirror complement developed by complementing and reversing the pattern. An example of these related patterns is given below.

1 1 0 0 1 1 0	pattern
0 0 1 1 0 0 1	complement
0 1 1 0 0 1 1	mirror
1 0 0 1 1 0 0	mirror complement

The designation of any one specific member of this group as the basic pattern is arbitrary and there exist patterns in which either the mirror or mirror complement is the same as the basic pattern. Thus, there may be only two distinct members in a group. The important distinguishing characteristic of any group is that all its members have the same agreement vector. Because of this all members have the same F_{Φ} and therefore it is necessary to examine only one member of each group in the search for the optimum frame sync code of a given length. The agreement vector of each group is unique.

The first program written was essentially an elementary pilot study composed of a pattern generator, a routine for developing the agreement vector, and a routine for evaluating F_{Φ} . This program was capable of searching for the optimum frame sync code for any designated length and any designated number of errors allowed. Pattern generation was realized through a binary count. The pattern complements were excluded by considering only those patterns having a 1 bit in the first position. This also excluded either the mirror or mirror complement depending upon whether the last bit of the pattern was a 1 or a 0. The exclusion of the fourth member of each group was achieved by treating the pattern as a binary value and comparing it to the value of its mirror or mirror complement whichever remained. The lower valued of the two was used. The agreement vector was developed through a simple comparison technique simulating the auto-correlation of the recognizer. The evaluation of F_{Φ} was made by plugging in the values and solving the

formulas for F_{Φ} . To this basic program were added several auxiliary routines to enable more study of the patterns. The first of these was a routine to print, for each generated pattern, the agreement vector, F_{Φ} , and the pattern itself. This was an optional routine to be used in place of listing only the optimum frame sync code. The second routine was an option to sort the patterns according to F_{Φ} before listing them. Another routine added to the basic pilot study was a function to exclude patterns purely on the basis of various agreement vector criteria. A great many variations of this agreement vector criteria routine were tried in an attempt to isolate good patterns. The criteria employed were, by and large, functions of δ , the degree of overlap, and n , the code length. One of the limitations tried was to exclude each pattern whose agreement vector contained a value greater than the integral value of $(\delta+2)/2$. Another was the exclusion of all patterns whose agreement vector contained a value greater than the integral value of $(n+2)/2$, etc. Although many variations similar to these were tested, a satisfactory agreement vector limitation that did not exclude some good patterns was never found and, consequently, this idea was abandoned.

This pilot study yielded the optimum frame sync codes up to 14 bits in length. In addition, examination of the listings of patterns ordered according to F_{Φ} disclosed that the best patterns of any length have different first and last bits, approximately the same number of 1 and 0 bits, and a number of trailing 0 or 1 bits. The number of trailing 0 bits increases irregularly as the length of the code increases. The number of 1 bits in the best patterns of each length (including the optimum codes) lies inclusively between

$$\left[\frac{n-1}{2} \right]$$

and

$$\left[\frac{n+2}{2} \right]$$

where the term

$$\left[\frac{a}{b} \right]$$

represents the integral part of the fraction a/b . This study also disclosed that ϵ could affect the

These two methods for reducing F_{ϕ} evaluation time were incorporated in a special program which was tailored to the requirements of the search for the 27-bit code. The program excluded all patterns which had fewer than 13, i.e.

$$\left[\frac{n-1}{2} \right],$$

or more than 14, i.e.,

$$\left[\frac{n+2}{2} \right],$$

1 bits. It considered only those patterns having independent agreement functions. That is, where a pattern was considered, the complement, mirror, and mirror complement were not. In addition, the search was modified to insist on five trailing 0's on the end and a 1 bit in the first position. The value of ϵ was fixed at 2 while the value of η was set at 0.10. This program developed the 27-bit optimum frame sync code in approximately three and one-half hours.

Up to this point all programming had been done in the CAMEO system (reference 6). The use of CAMEO was advantageous to the study because of the ease it provided in the writing and debugging of programs and for the facility it afforded programming modifications. However, it became apparent that a more efficient program could be written if, in certain areas of the program, advantage was taken of some special machine language logic instructions.

In the final version of the search program, the input-output and initializing routines were written in CAMEO. The central portion was written in machine language in order to place particular emphasis of efficient use of the computer. Each routine was positioned to take maximum advantage of the instruction overlap time of the IBM 7094. All unnecessary loops were removed and replaced with straight line coding. This program developed the optimum 27-bit frame sync code in forty-three minutes. A flow chart of the program is given in Figure 5.

Two tolerances based on optimum frame sync code characteristics determined from the pilot study were included as options in this program: Limits could be placed on the number of 1 bits to be allowed and provision could be made to insist on trailing 0's on the end of the pattern. Both

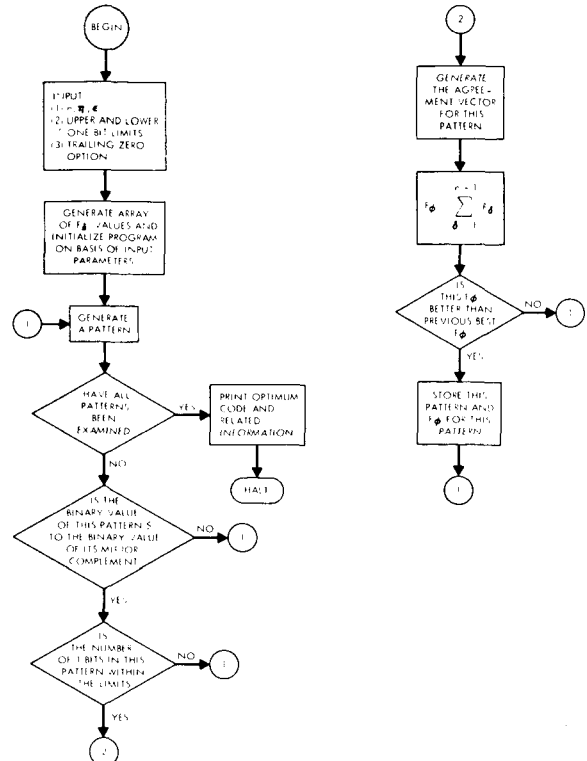


FIGURE 5.—Search program flow chart.

of these tolerances, which could be preset at the discretion of the program user, were included to reduce, when desirable, the length of time necessary to develop an optimum code. The program insisted on a 1 bit in the first position of the pattern and a 0 in the last. It was used to develop the optimum frame sync codes for lengths 7 through 20, using ϵ values 0 through 4 and the optimum codes of lengths 21 through 25 using ϵ values 0 through 3. A slight modification to the program was then made. This modified version was able to examine, without any limitation, all patterns having distinct agreement vectors. It was used to again develop the optimum codes of length 7 through 20 using ϵ values 0 through 4 in order to check the validity of assuming that the optimum codes have trailing 0's and that the number of 1 bits lies between

$$\left[\frac{n-1}{2} \right]$$

and

$$\left[\frac{n+2}{2} \right].$$

The results were the same as those developed by the unmodified version.

As has been noted, the use of different values of ϵ , the number of errors allowed, may result in the development of different optimum frame sync codes of the same length. Since the values of F_{ϕ} for such codes when evaluated under the same ϵ are so nearly equal, the choice of one such code over another is essentially arbitrary. Therefore, because of the amount of computer time involved in the search for the longer optimum codes, (e.g., 10.5 hours for the 30-bit code), the decision was made to develop the codes of length 26 through 30 for only one value of ϵ . The value used was $\epsilon=2$. The tolerances used in the development of these longer codes restricted consideration to those patterns having a 1 bit in the first position, a 1 bit total between the limits

$$\left[\frac{n-1}{2} \right]$$

and

$$\left[\frac{n+2}{2} \right],$$

and five trailing 0's. The value of η , the probability of bit changes incurred as a result of noise, was set at 0.10.

RESULTS

Table 1 is a listing of optimum frame sync codes derived as a result of this study, together with their agreement vectors and probabilities of false sync. Note that the term 'optimum' is qualified by the fact that these codes were determined under the condition of two errors allowed by the recognizer. Though this choice of ϵ is arbitrary, as pointed out previously, it was made in the interest of obtaining standard frame sync codes for PCM telemetry usage.

While the GSFC Standards for PCM Telemetry allow the use of codes up to 32 bits in length, codes of length 31 and 32 have not yet been determined because of the amount of computer time their determination would require—approximately 60 hours.

Figure 6 is a set of curves showing the total probability, F_{ϕ} , of false sync over the overlap region from $\delta=1$ to $\delta=n-1$, for the optimum frame sync codes given in Table 1. The values

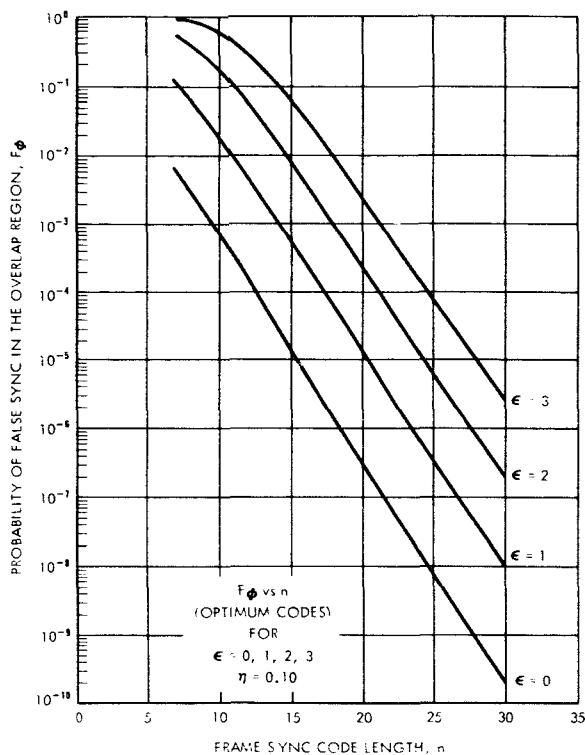


FIGURE 6.—Graphic representation of F_{ϕ} .

plotted were calculated using a bit error rate, $\eta=0.10$, and number of errors allowed by the recognizer, $\epsilon=0, 1, 2$ and 3.

FURTHER STUDY

As this investigation continued it became apparent, principally from the literature, that though a number of people were working in the area of optimization of frame synchronization techniques and codes for PCM telemetry, very few were directing their efforts toward the optimization of data channel usage by the selection of the proper frame sync code lengths. In the past the code length has been selected primarily on the basis of simplicity of implementation in the spacecraft and adaptation to the telemetry format. Adequacy of frame sync codes to synchronize specific length frames under anticipated signal-to-noise conditions has been based more on the intuitive feelings of the designer than on any quantitative evidence. For this reason, the work on optimum frame sync codes has been extended to the determination of code lengths to be used in the spacecraft, as well as number of errors to be

TABLE 1.—*Optimum Frame Sync Codes*

CODE 1 0 1 1 0 0 0 AGREEMENT VECTOR (03 02 02 01 01 00)	PROBABILITY OF FALSE SYNC 5.723×10^{-1}	LENGTH 07
CODE 1 0 1 1 1 0 0 0 AGREEMENT VECTOR (04 03 01 02 01 01 00)	PROBABILITY OF FALSE SYNC 4.235×10^{-1}	LENGTH 08
CODE 1 0 1 1 1 0 0 0 0 AGREEMENT VECTOR (05 04 02 02 01 01 01 00)	PROBABILITY OF FALSE SYNC 2.950×10^{-1}	LENGTH 09
CODE 1 1 0 1 1 1 0 0 0 0 AGREEMENT VECTOR (06 04 03 03 02 01 01 00 00)	PROBABILITY OF FALSE SYNC 1.783×10^{-1}	LENGTH 10
CODE 1 0 1 1 0 1 1 1 0 0 0 AGREEMENT VECTOR (05 04 04 03 03 02 02 01 01 00)	PROBABILITY OF FALSE SYNC 9.065×10^{-2}	LENGTH 11
CODE 1 1 0 1 0 1 1 0 0 0 0 0 AGREEMENT VECTOR (06 06 05 03 04 03 02 01 01 00 00)	PROBABILITY OF FALSE SYNC 5.142×10^{-2}	LENGTH 12
CODE 1 1 1 0 1 0 1 1 0 0 0 0 0 AGREEMENT VECTOR (07 07 05 04 04 04 03 01 01 00 00 00)	PROBABILITY OF FALSE SYNC 2.821×10^{-2}	LENGTH 13
CODE 1 1 1 0 0 1 1 0 1 0 0 0 0 0 AGREEMENT VECTOR (08 06 06 05 04 05 03 03 02 01 00 00 00)	PROBABILITY OF FALSE SYNC 1.514×10^{-2}	LENGTH 14
CODE 1 1 1 0 1 1 0 0 1 0 1 0 0 0 0 AGREEMENT VECTOR (07 07 07 06 05 05 03 04 02 02 01 00 00 00)	PROBABILITY OF FALSE SYNC 6.611×10^{-3}	LENGTH 15
CODE 1 1 1 0 1 0 1 1 1 0 0 1 0 0 0 0 AGREEMENT VECTOR (08 08 06 07 06 05 05 02 03 03 02 01 00 00 00)	PROBABILITY OF FALSE SYNC 3.460×10^{-3}	LENGTH 16
CODE 1 1 1 1 0 0 1 1 0 1 0 1 0 0 0 0 0 AGREEMENT VECTOR (09 09 07 07 06 06 04 05 04 03 03 01 00 00 00 00)	PROBABILITY OF FALSE SYNC 1.657×10^{-3}	LENGTH 17
CODE 1 1 1 1 0 0 1 1 0 1 0 1 0 0 0 0 0 0 AGREEMENT VECTOR (10 10 08 08 07 06 05 05 05 03 03 02 01 00 00 00 00)	PROBABILITY OF FALSE SYNC 8.228×10^{-4}	LENGTH 18
CODE 1 1 1 1 1 0 0 1 1 0 0 1 0 1 0 0 0 0 0 0 AGREEMENT VECTOR (11 09 09 09 07 07 06 05 05 04 04 03 02 00 00 00 00 00)	PROBABILITY OF FALSE SYNC 3.837×10^{-4}	LENGTH 19
CODE 1 1 1 0 1 1 0 1 1 1 1 0 0 0 1 0 0 0 0 0 0 AGREEMENT VECTOR (12 10 10 09 09 08 07 05 05 04 01 03 03 02 01 01 00 00 00)	PROBABILITY OF FALSE SYNC 2.175×10^{-4}	LENGTH 20
CODE 1 1 1 0 1 1 0 1 0 1 0 0 1 0 1 1 0 0 0 0 0 0 AGREEMENT VECTOR (11 11 08 08 08 07 07 07 05 04 04 04 02 01 01 01 00 00 00)	PROBABILITY OF FALSE SYNC 1.051×10^{-4}	LENGTH 21
CODE 1 1 1 1 0 0 1 1 0 1 1 0 1 0 1 0 0 0 0 0 0 0 AGREEMENT VECTOR (12 12 10 09 08 08 07 06 05 05 05 04 03 02 02 01 00 00 00 00)	PROBABILITY OF FALSE SYNC 4.906×10^{-5}	LENGTH 22
CODE 1 1 1 1 0 1 0 1 1 1 0 0 1 1 0 1 0 0 0 0 0 0 0 AGREEMENT VECTOR (13 13 11 11 10 10 09 08 04 07 05 06 04 03 03 02 01 01 00 00 00)	PROBABILITY OF FALSE SYNC 2.533×10^{-5}	LENGTH 23
CODE 1 1 1 1 1 0 1 0 1 0 1 1 1 1 0 0 1 1 0 0 0 0 0 0 AGREEMENT VECTOR (14 12 12 12 10 10 10 09 08 07 05 05 03 05 04 03 02 02 00 00 00 00)	PROBABILITY OF FALSE SYNC 1.255×10^{-5}	LENGTH 24
CODE 1 1 1 1 1 0 0 1 0 1 1 0 1 1 1 0 0 0 1 0 0 0 0 0 0 AGREEMENT VECTOR (15 13 13 11 11 10 08 08 09 07 08 06 05 05 04 04 03 03 01 00 00 00 00)	PROBABILITY OF FALSE SYNC 6.449×10^{-6}	LENGTH 25
CODE 1 1 1 1 1 0 1 0 0 1 1 0 1 0 1 1 0 0 0 1 0 0 0 0 0 0 AGREEMENT VECTOR (14 14 12 12 11 08 09 10 08 07 07 07 05 05 04 04 03 02 01 00 00 00 00)	PROBABILITY OF FALSE SYNC 3.144×10^{-6}	LENGTH 26
CODE 1 1 1 1 1 0 1 0 1 1 0 1 0 0 1 1 0 0 0 0 0 0 0 0 AGREEMENT VECTOR (15 13 15 13 13 10 10 10 09 09 08 07 07 06 06 05 04 04 03 01 01 00 00 00)	PROBABILITY OF FALSE SYNC 1.583×10^{-6}	LENGTH 27
CODE 1 1 1 1 0 1 0 1 1 1 1 0 0 1 0 1 1 0 0 1 1 0 0 0 0 0 0 AGREEMENT VECTOR (16 14 14 12 13 13 12 11 08 10 08 08 08 06 06 04 04 04 03 02 01 01 00 00 00)	PROBABILITY OF FALSE SYNC 8.036×10^{-7}	LENGTH 28
CODE 1 1 1 1 0 1 0 1 1 1 1 0 0 1 1 0 0 1 1 0 1 0 0 0 0 0 0 0 AGREEMENT VECTOR (17 15 15 15 14 12 13 11 10 10 09 09 08 04 07 05 06 04 03 03 02 02 01 01 00 00 00)	PROBABILITY OF FALSE SYNC 4.093×10^{-7}	LENGTH 29
CODE 1 1 1 1 1 0 1 0 1 1 1 1 0 0 1 1 0 0 1 1 0 1 0 0 0 0 0 0 0 AGREEMENT VECTOR (18 16 16 16 14 13 13 12 11 11 10 09 08 05 08 05 06 05 04 03 03 02 01 01 00 00 00)	PROBABILITY OF FALSE SYNC 2.070×10^{-7}	LENGTH 30

allowed by the code recognizer in the ground station, as a function of telemetry frame length and bit error rate (or signal-to-noise ratio) of the received PCM signal.

A computer program to determine these parameters for the GSFC PCM Telemetry Standards frame sync codes has been completed and run on the IBM 7094. At present, a second report is being prepared which will present the results of this effort in graphical form with recommendations as to code length and errors to be allowed under the above varying conditions. The computer program has the additional capability of taking any bit pattern up to a 40-bit length and performing the same evaluation on it as is performed on the 'optimum' code. In this way, frame sync codes proposed by other authors may be directly compared to the codes of this study in terms of average time to acquire true frame sync and percent data lost due to loss of frame sync.

ACKNOWLEDGMENT

The authors of this paper are grateful to Miss Barbara S. Goss for much of the programming in

the early phases of this study and especially for her assistance in the analysis of the problem.

REFERENCES

1. It is generally conceded that the third type of synchronization, word sync, is wasteful of information bandwidth and, in the presence of low signal-to-noise conditions, cannot be reliably detected. Word sync has not been considered in this effort. (References 2, 3, and 4).
2. "PCM Telemetry Synchronization," M. W. WILLIARD, Proceedings of the NTC, May 1961, pp. 11-51 to 11-74.
3. "Synchronization Methods for PCM Telemetry," E. R. HILL and J. L. WEBLEMOE, Proceedings of the NTC, May, 1961, pp. 11-87 to 11-106.
4. "Digital Synchronization of PCM Telemeters," J. P. MAGNIN, Proceedings of the NTC, May, 1962, pp. 5-3-1 to 5-3-10.
5. "Optimum PCM Frame Synchronization Codes and Correlation Detection," G. E. GOODE and J. L. PHILLIPS, Proceedings of the NTC, May 1961, pp. 11-15 to 11-50.
6. "CAMEO: Computer-independent Abstract Machine-language Encoder and Operating-system, System Description," T. P. GORMAN, in preparation.

TIMING ERRORS—THEIR DETECTION AND CORRECTION IN THE IMP INFORMATION PROCESSING SYSTEM*

WILLIAM H. MISH

One of the major problems encountered by many experimenters in the analysis of data retrieved by experiments flown on board satellites has been that of obtaining correct Universal Time in conjunction with this data.

This paper presents the scheme that was used by the IMP Information Processing System in the processing of the data from IMP-I which provided a means of eliminating most of the timing errors before they reached the experimenter. In addition this scheme proved valuable in the initial debugging of the IMP analog to digital line and most significant in evaluating the stability of and providing a check on the IMP-I spacecraft telemetry system.

INTRODUCTION

The generation of Universal Time at the ground station, the merging of this time with recorded spacecraft data, and the subsequent processing of this time through analog to digital conversion equipment has, historically, had a multiplicity of problems associated with it. It would appear that a certain amount of these problems are unavoidable when one considers the complexity of the over-all situation.

The "aggregate" timing of data from most spacecrafts is a product of a relatively large number of Prime Minitrack Time Standards located throughout a world wide network of Space Tracking and Data Acquisition Stations. Each of these time standards is designed to provide a time base for the spacecraft data that is in agreement with Universal Time, i.e., WWV, to within one millisecond when the time standard has been synchronized correctly with WWV and the propagation delay times between WWV and the station have been compensated (see reference 12, page 2-95). The prime function of these time standards is to provide a time code that can be written on analog tape simultaneously with the recording of the spacecraft data at the various telemetry receiving stations. This analog tape is then shipped to a central information processing facility where it provides the input for analog to

digital conversion equipment, one function of which, is to decode the time and convert it to a digital format in conjunction with the data, and finally production of a digital magnetic tape which can subsequently be processed with computers.

As the practice of flying scientific experiments which required high resolution of the ground station time became prevalent it became obvious to the experimenters that they were *not* always realizing the design capabilities of the timing standards, i.e., one millisecond agreement with WWV when propagation effects have been accounted for, due to the degradation of the timing as it traversed the above mentioned series of operations. The occurrences of these "timing errors" was so frequent in some cases as to require a considerable expenditure of effort on the part of the individual experimenters to verify the integrity of the timing associated with their data.

One solution to this problem was to incorporate as part of the information processing facilities a systematic method, necessarily computer orientated because of the large volumes of data, of detecting and correcting timing errors before they reached the experimenter.

Up to the advent of the IMP-I spacecraft and its crystal controlled telemetry system a scheme for the detection and correction of timing errors of the nature described in this paper was, in most cases, impractical because the telemetry systems,

*Published as *Goddard Space Flight Center Document X-612-64-328*, October 1964.

even those that telemetered "pseudo clocks," (see reference 11, page 9) were not stable enough with time to allow the telemetry patterns or pseudo clocks to be used as an independent "clock" against which ground station time could be accurately compared. However, with the successful launch of IMP-I on November 27, 1963, a satellite became available which had a telemetry format consisting of a repeating pattern of telemetry sequences with a relatively stable period that could be used to perform an independent check on ground station time. This aspect of the IMP-I telemetry system was taken advantage of in the design of the IMP Information Processing System (IMP-IPS). (See reference 13.)

This paper gives a brief summary of the IMP-IPS, discusses in detail those phases of the IMP-IPS which are concerned with the detection and correction of timing errors and how this aspect of IMP-IPS was employed to trouble-shoot the IMP analog to digital line and monitor the IMP-I spacecraft telemetry system throughout the useful life of the spacecraft.

APPLICABILITY OF THE IMP-I TELEMETRY FORMAT TO TIME ERROR DETECTION AND CORRECTION

The IMP-I telemetry system is a modification of the basic type of PFM telemetry system flown on some of the past satellites, e.g., Explorer XII and UK-1. For a discussion of this telemetry system see references 1, 15 and 17.

There are *two* significant features of the IMP-I telemetry format that permit the effective implementation of a time error detection and correction scheme of the type described in this paper:

- A. The telemetry format is a synthesis of a repeating pattern of four telemetry sequences each approximately 81.914 seconds in length (see Figure 1). Each of these sequences with the exception of the fourth, consists of an array of sixteen channels by sixteen frames. The first three sequences contain an octal satellite clock reading in channels 11 through 15, frames 4 and 12, which increases by one decimal unit each sequence including the fourth. The low order octal digit of the satellite clock reading is decoded by the analog to digital line into

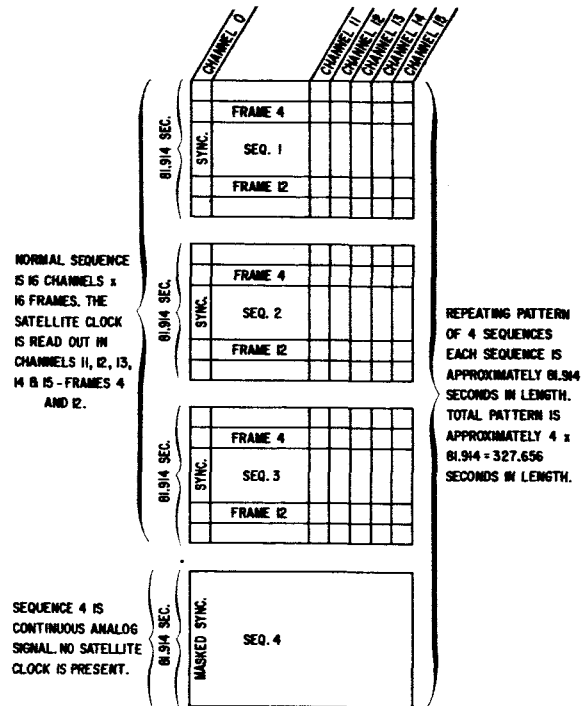


FIGURE 1.—The IMP-1 telemetry format.

a 1, 2 or 3 depending on whether it is the first, second, or third telemetry sequence respectively. This information is included as part of the digitized data and is utilized by the time error detection software. The fourth sequence (Rubidium vapor magnetometer sequence) contains continuous analog data and consequently is missing the satellite clock reading. (The satellite clock still up-dates itself by one decimal unit in this sequence even though it is not read out.)

- B. The periodicity with time of the channel rate is insured through the use of a 20kc crystal as a basic component of the spacecraft telemetry system. This basic crystal frequency is divided down to obtain the desired sampling rate and is backed up with a free running multi-vibrator synchronized to one of the subdivided frequencies from the crystal. When not synchronized with the crystal the multi-vibrator runs some 25 percent lower in frequency than when in synchronization with the crystal. The significance of this feature becomes apparent in later discussions in this paper.

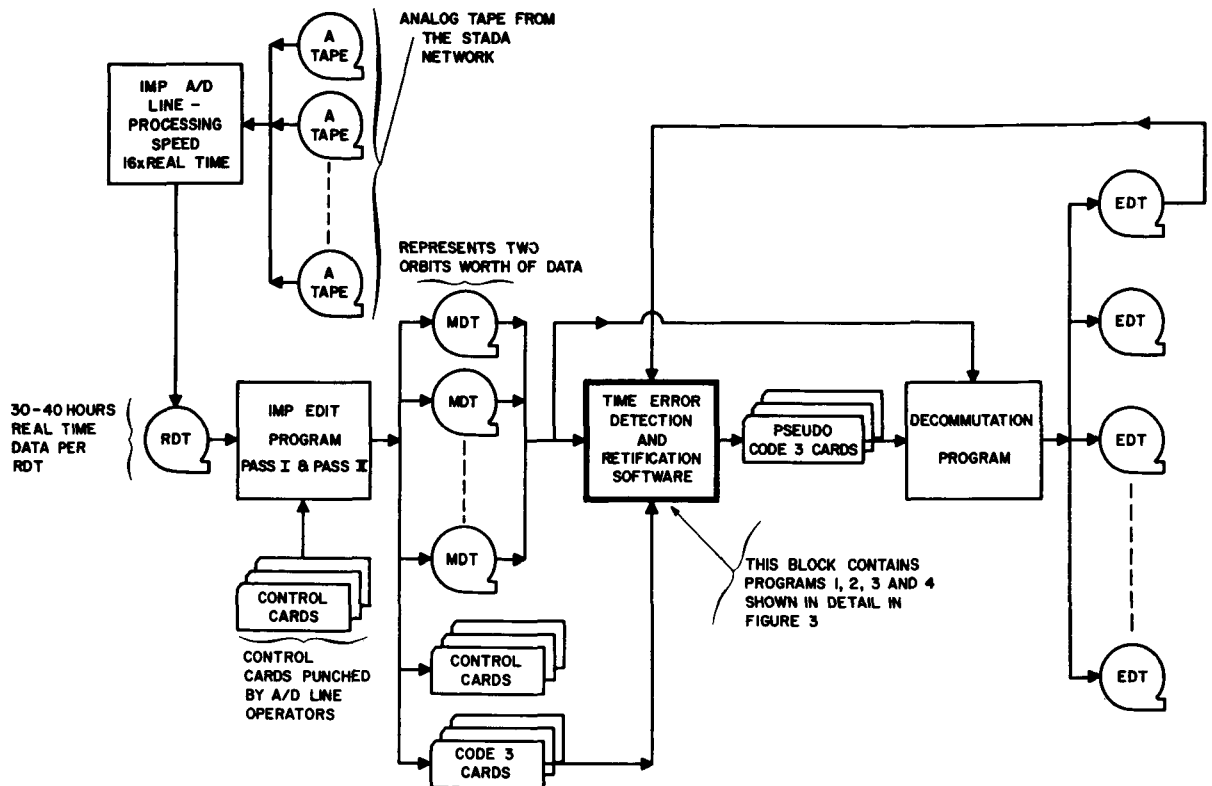


FIGURE 2.—The IMP information processing system (IMP-IPS).

IMP INFORMATION PROCESSING SYSTEM (IMP-IPS)

The IMP-Information Processing System is an integrated set of analog to digital conversion equipment and computer programs which receive as input information in an analog format and output information in a standard Binary Coded Decimal format. Documentation of this system is extensive and details can be found in References 6, 7, 9, 13, 14, 16, and 18.

Because information is processed serially through the system a brief explanation of each of the steps is appropriate to allow the presentation of the Time Error Detection and Rectification Software in its proper place with respect to the over-all series of operations (see Figure 2).

The flow of information through the IMP-IPS commences with the reception of the analog tapes from the network of stations covering the spacecraft. These analog tapes are processed through the IMP analog to digital (A/D) line, (see refer-

ences 7, 9, and 18) which digitizes the analog data and time and creates various flags and symbols that are a function of the time-data array at 16 times "real time." The digitized output of the A/D line is written on tape in a BCD format, this tape is called a Raw Data Tape (RDT) and contains from 30 to 40 hours of digitized "real time" data. This 30 to 40 hours of data constitutes many satellite passes* from a number of different stations. The RDT, along with control cards punched by the A/D line operators, serve as input to the IMP Edit Program (see reference 16). This edit program, which is operationally broken into two passes on the computer, performs in the most general sense the following functions:

- A. The Program checks for and corrects operational and equipment A/D line errors, i.e., "cleans up" the RDT.

*A pass is defined as one uninterrupted data recording made at a single station (average pass on IMP-1 is about 4 hours).

- B. Uses the IMP-I satellite clock* to tag each telemetry sequence with a monotonically increasing number for each 90 days of satellite operation.
- C. Uses the error detecting and correcting properties of the IMP-I satellite clock to create a flag for *each* telemetry sequence which gives a measure of the quality of the data in the sequence.

The final output of the IMP Edit Program is a Master Data Tape (MDT) which contains 120 to 160 hours of real time data, and control cards that are punched by the program. Of special interest is the Code 3 Control Card which, after being operated on by the Time Error Detection and Rectification Software, serves as the *timing interface* between the MDT, which contains "raw time," and the Experimenter Data Tapes (EDTs) which contain "corrected and smoothed" time.

At this point in the series of computer operations the MDTs are allowed to backlog until *two* consecutive orbits' worth of data have accumulated. As each complete orbit takes about 93 hours this constitutes about eight days of real time data on each MDT.

The two orbits' worth of data are then processed through the Time Error Detection and Rectification Software system. In this series of computer operations, which will be covered in detail in the following sections, errors in time are detected and provision is made for the subsequent correction of these errors and "smoothing" of the time through the creation of the Pseudo Code 3 Cards. These Pseudo Code 3 Cards and the MDTs then provide the input to the decommutation program (see reference 6) which "smooths" and corrects, if necessary, the time for each pass contained in the two orbits' worth of data being processed, and creates the experimenter tapes containing each individual experimenter's data. Finally, one of the experimenter tapes is fed back into the Time Error Detection and Rectification Software where the alteration made to the timing through the use of the Pseudo Code 3 Cards is reviewed.

*The spacecraft "clock," a 15-bit accumulator, can accumulate a maximum of 32,767 counts, i.e., 2^{15} counts, before it returns to zero. To fill this accumulator takes approximately 30 days. This clock, through the use of programming techniques on the ground, has been extended to a 90-day clock in the data that the experimenter receives (see Appendix G.)

THE TIME ERROR DETECTION AND RECTIFICATION SOFTWARE

The Time Error Detection and Rectification Software consists of four separate IBM 1410 computer programs which have been represented by a single block in Figure 2. The contents of this block are shown in detail in Figure 3, where the four programs are serially numbered from 1 to 4.

This series of programs is designed to accomplish the following objectives:

- A. The MDT Time Verification Program (program 1 in Figure 3) produces a plot tape, which, when sorted chronologically with respect to time and plotted, provides a means of visually detecting timing errors as well as providing information on the performance of the spacecraft telemetry system and the A/D line, and when input to the Pseudo Code 3 Card Program (program 3 in Figure 3) allows the timing to be analyzed, corrected and smoothed.
- B. The Time Verification Sort Program (program 2 in Figure 3) insures that the data sent to the experimenter is in chronological order and provides a sorted input tape for the Pseudo Code 3 Card Program.
- C. The Pseudo Code 3 Card Program performs an analysis on the timing and corrects timing errors that may exist by altering the Code 3 Cards to Pseudo 3 Cards which are used to control the timing that is generated by the decommutation program for the experimenters' tapes.
- D. The California EDT Time Verification Program (program 4 in Figure 3) again produces a plot tape which when plotted illustrates the alterations that have been made to the timing by the use of the Pseudo Code 3 Cards.

The MDT Time Verification Program—A Tool for Detecting Time Errors and Monitoring the A/D Line and Spacecraft Telemetry System Performance.

Referring to Figure 1, the MDT Time Verification Program (program 1 in Figure 3) takes advantage of the fact that channel 0 of frame 0 of each telemetry *sequence 1* occurs approximately

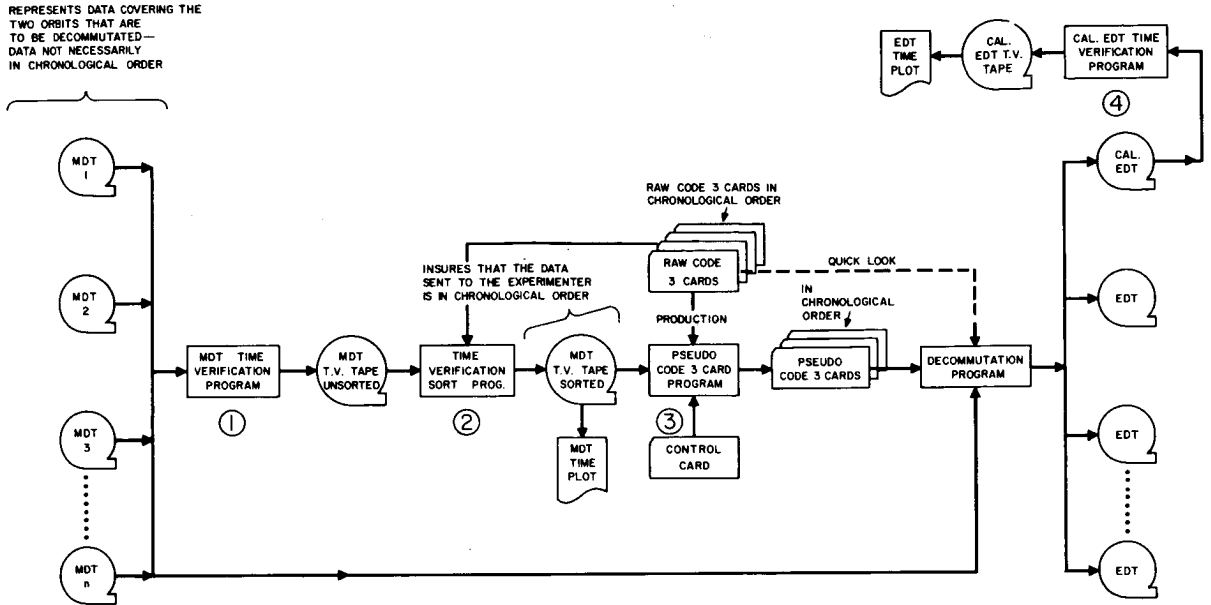


FIGURE 3.—The time error detection and rectification software.

every 327.6 seconds. This enables the program to assign a unique number defined as N to each sequence 1, i.e., the program counts the number of sequence 1s that have occurred since some chosen reference sequence 1 (at the reference sequence 1, N equals zero). N is determined by the program as follows:

$$N = \left[\frac{t_n - t_0}{\Delta t} \right] \quad (1)$$

(Note that the brackets indicate that N has been rounded to an integer.)

Where:

- t_0 is defined as the time for channel 0 of frame 0 of the reference sequence 1,*
- t_n is the time for the current channel 0 of frame 0 of the sequence 1 being counted, and
- Δt is a chosen supercommutated sequence rate† and is equal to 327.565 seconds. (It should be noted that the above supercommutated sequence rate was chosen at the same time that t_0 was chosen and does not represent the

*It has been necessary to change this reference point once during the useful life of the spacecraft. (See Appendix A orbit 25.)

†The sequence rate is assumed to be 81.914 seconds (see Figure 1).

average supercommutated sequence rate from t_0 to t_n .)

The above quantities are used by the program to compute the parameter Delta for each sequence 1 that is counted. Delta is defined by the following formula:

$$\text{Delta} = (t_n - t_0) - N(\Delta t) \quad (2)$$

Where:

- $(t_n - t_0)$ represents the actual elapsed time for N sequence 1s,
- $N(\Delta t)$ represents the elapsed time that would be predicted for N sequence 1s if the actual supercommutated sequence rate was a constant, and Delta represents the algebraic difference of the two foregoing quantities.

A sample computation of Delta is presented in Appendix F. Note that the computation of N is accurate as long as t_n is in error by less than 50 percent of 327.6 sec. minus the accumulated error due to the slightly improper choice of the supercommutated sequence rate.

As previously mentioned the parameter Delta is computed for each sequence 1 that is counted

by the program. This parameter Delta along with its corresponding N is written on magnetic tape by the program in a format that allows, after being sorted chronologically with respect to time (program 2 in Figure 3), Delta versus N to be directly plotted with an EAI* data plotter. Appendix A contains the complete library of these plots for the useful lifetime of the spacecraft.

These plots are an extremely useful tool in verifying the integrity of the timing associated with each pass of data that has been recorded and processed through the A/D line during any particular orbit. In addition, a great deal of information about the performance of the spacecraft telemetry system, i.e., stability of the bit rate, can be derived from a review and analysis of these plots.

The following paragraphs treat the applicability of these plots to the detection of timing errors originating at the station and during analog to digital conversion. In addition the plots proved valuable in the initial debugging of the A/D line. A discussion encompassing the usefulness of the plots in the evaluation of the performance of the spacecraft telemetry system is also included.

Detection of Timing Errors Originating at the Ground Station

- A. The Ground Station clock set incorrectly with WWV during the recording of a pass.

This relatively common error (see Appendix E) is reflected in the plot of Delta versus N as a positive or negative displacement in the ordinate values, i.e., Delta, from the nominal curve during the pass. Appendix A, orbit number 1, contains a number of passes recorded at Rosman, North Carolina which contain this particular error. Two of these passes have had the BCD time code and the WWV signal simultaneously strip charted to illustrate the discrepancy in the two time codes. These strip charts are displayed on the same page with the plot of orbit Number 1 in Appendix A.

- B. Errors introduced by the initial disagreement of the station clocks.

During the life of IMP-I the tracking stations were *not* compensating for the propagation time from WWV to the station site. This results in all the station clocks initially being in disagreement with each other and with WWV. Appendix D contains a tabulation of these delay times.

This initial disagreement of the station clocks combined with the fact that the distance from the spacecraft to the tracking stations changes abruptly as coverage passes from station to station is readily apparent in all the plots appearing in Appendix A as small positive and negative steps in the curve of Delta versus N .

- C. Analog to Digital line could not synchronize properly with the data because of a weak or noisy telemetry signal. This adversely effected the timing.

This anomaly is reflected in the plot of Delta versus N as a predominant scattering of the points during the pass. Appendix A, orbit 5 around sequence 1 counts of 4189 and 4789 are good examples.

- D. Both BCD and Serial Decimal time were found to be unacceptable and the pass was processed with relative time.†

This type of error is generally reflected in the plots as a gap in the data where the pass would, under normal circumstances, be located. This is due to the fact that the values of Delta and N will generally be completely erroneous during the pass and consequently are off scale when the plots are made.

- E. Incorrect ID

Because in the computation of t_n the program utilizes the ID of the file (see sample calculation in Appendix F) an incorrect ID will also result in completely erroneous values of Delta and N . Again when the data is plotted a gap will be present where

†Relative time is defined as time that is generated by the A/D line accumulator after the accumulator has been set to zero. (See reference 18.)

*Electronic Associates, Inc. Model 3440.

the pass would, under normal circumstances, be located.

This ID check is an important feature of this scheme, as incorrect IDs are very common (particularly when the recording of a pass at the station commences very close to midnight) and when not detected and corrected can cause the experimenter considerable difficulty in rectifying the proper date.

Trouble-Shooting the IMP Analog to Digital Line

As the A/D line had completed the final stages of construction just prior to the launch of the spacecraft there were a certain amount of undiscovered problems that remained to be resolved and improvements made during the processing of the first orbits of data.

- A. At the time that data from the sixth orbit was being processed a refinement was made to the line in the form of a circuit that served as a window to look for phase errors at the end of each channel between the system flywheel frequency, which determines when time is sampled, and the 50 cycle channel rate derived from the recorded data on the analog tape. During installation this window was mistakenly inverted. The plot of Delta versus N reflected this malfunction as small "ramps" in the plots of each pass. (See Appendix A orbits 6 through 13; orbit 10 around sequence 1 counts of 9709 and 9909 is a particularly good example.)

The chances of seeing an error of this magnitude would have been remote if a scheme, such as described in this paper, for detecting time errors was not being employed. The reason for the ramps in the plots were investigated and subsequently discovered and corrected (see reference 7) as can be seen in Appendix A in the plots after the 13th orbit.

- B. Another interesting and yet incompletely explained structure in the plots of Delta versus N are the "tails" observed on many of the passes. (See Appendix A orbits 2 and 3.) It is conjectured by the designers of the A/D Line that these tails are due to a disagreement between the frequency of the

flywheel, which determines the rate at which time is sampled, and the 50 cycle channel rate derived from the recorded data on the analog tape at the start of processing of the pass. As can be seen in the plots as time progresses the flywheel slowly "syncs" in on the 50 cycle channel rate. This malfunction seems to have remedied itself during the processing of data in the later orbits but again appears in plots after orbit 38.

Monitoring the Spacecraft Telemetry System

Because the program is looking at the difference between the actual time for each sequence 1 and the predicted time for each sequence 1 (assumed supercommutated sequence rate of 327.656 sec.) a considerable amount of information is made available about the periodicity with time of the sequence rate. One of the distinct advantages of this analysis is the fact that a change in the sequence period will be reflected in the curve.

- A. In Appendix A, orbit 1 a sudden change in the period of a single supercommutated sequence introduced an approximate 71 second discontinuity in the curve (see Appendix A, orbit 1, sequence 1 count of 800). As explained in the section titled *Applicability of the IMP-I Telemetry System to Time Error Detection and Correction* the stability of the sampling rate of the spacecraft is assured through the use of a 20kc crystal which is backed up with a free running multi-vibrator synchronized to one of the subdivided frequencies from the crystal. Apparently this discontinuity was caused by an overload on the satellite power system which prevented the multi-vibrator from synchronizing to the crystal. Consequently it "free ran" about 25 percent low for about 5 minutes. (See reference 16, Appendix F, page 8.)
- B. Orbit number 42 in Appendix A covers that period when the spacecraft was eclipsed by the Earth for 7^h 57^m, as is evident from the plot the spacecraft transmitted for 1^h 17^m after entering the Earth's shadow. It was completely shut off for a period of 15^h 17^m and became operational again on May 7 at 07^h 38^m U.T. (see reference 2 and 4). It is

estimated that during the period of time that the spacecraft was not operational the temperature of the satellite electronics dropped to temperatures between -60° and -80°C (see reference 8, page 3). This "cold soak" evidently changed the operating frequency of the crystal and consequently the sequence rate.

The telemetered temperature of the spacecraft battery, which is physically located close to the encoder and thus the crystal, is shown plotted against time following spacecraft turn on after the shadow in Figure 4. The "raw time" plot of Delta

in Appendix C also reveal this rather "radical" change in the sequence rate.

- C. Appendix C provides a summary of the stability characteristics of the IMP-I telemetry system during the useful life of the spacecraft.* Graph 1 was generated by plotting the apogee value of Delta against the appropriate orbit number. Referring to equation (2), i.e.,

$$\Delta t = (t_n - t_0) - N(\Delta t)$$

it would be expected that if Δt was chosen incorrectly it would introduce a constant slope in the plot of the Apogee Values of Delta versus the Orbit Number.

As it turned out Δt equal to 327.656 was not precisely the average supercommutated sequence rate. This fact is reflected in the plot of the Apogee Values of Delta versus the Orbit Number in Appendix C.

The interesting aspect of this plot is not the slope itself but the rate of change of the slope which gives a measure of the long range stability of the telemetry system. Construction of a tangent to a least squares quadratic fit of data from the 5th to the 24th orbit at the 5th orbit indicates the sequence rate had drifted on the order of 4 seconds in 6.4 million seconds through the 24th orbit. From the 25th orbit until the satellite was eclipsed for an extended length of time in the 42nd orbit the stability undergoes some deterioration.

- D. In addition to the Apogee Values of Delta plotted against the Orbit Numbers, Appendix C contains a plot of the Average Sequence Rate Per Pass plotted against the Satellite Clock Reading at the beginning of the pass. The instabilities that were mentioned in the previous paragraph are particularly evident in this plot.

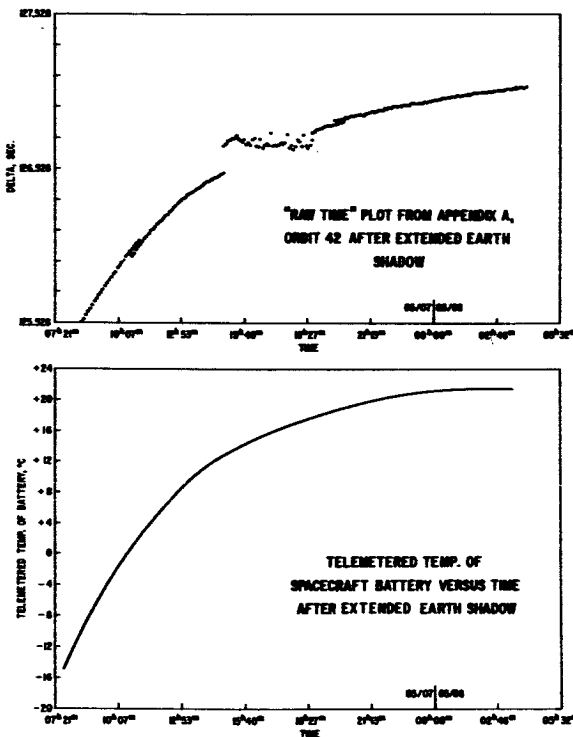


FIGURE 4.—Telemetered temperature of battery after spacecraft turn on following the extended earth shadow compared with the "Raw Time" plot during this time.

versus N for this same period of time has also been included to illustrate the changing sequence rate. As can be seen from Figure 4 it took about 19.5 hours before the crystal frequency stabilized close to its original value. The graphs of Average Sequence Rate per Pass versus Satellite Clock Num-

The Pseudo Code 3 Card Program—A Means of Correcting and Smoothing Time

The Pseudo Code 3 Card Program (program 3 in Figure 3) uses the sorted MDT Time Verifica-

*With two exceptions the spacecraft operated continuously from launch on November 27, 1963 until May 30, 1964 at $09^{\text{h}} 18^{\text{m}}$ U.T. (see reference 3). After May 30 the satellite became extremely intermittent.

tion Tape, generated by the MDT Time Verification Program (program 1 in Figure 3) and sorted by the Time Verification Sort Program (program 2 in Figure 3), as input. This tape contains the tabulation of Delta versus N (the sequence 1 count) for the two consecutive orbits, henceforth referred to as the Delta- N data array, of data being processed.

At the outset the program operates on the Delta- N data array algebraically subtracting the delay time of WWV propagation to the station site (see Appendix D) from each of the Deltas computed for each sequence 1 that comprise a pass. This compensation for the delay time of WWV synchronization to the station site is carried out on each of the passes that comprise the orbit being processed.

As an example in Appendix F there is a sample calculation for the Delta corresponding to the first sequence 1 in a pass that was recorded at Woomera, Australia on January 8, 1964. In this calculation Delta is computed to be +86.697 sec. Now referring to Appendix D the delay time from WWV to the station site at Woomera is 59.1 milliseconds. Thus the adjusted Delta would be:

$$+86.697 - .059 = 86.638$$

Next the program operates on the adjusted Delta- N data array for *each* individual orbit fitting it, using least squares, to a quadratic equation. It is readily apparent from the plots in Appendix A that a second degree equation of the form:

$$\text{DELTA} = A_1 + A_2N + A_3N^2 \quad (3)$$

will approximately fit the data if we proceed from perigee to perigee.* Appendix B shows a comparison of the raw time and the adjusted and least squares smoothed and corrected time for orbits 36 and 37. Note that the adjusted and smoothed time curve is below the raw time curve during most of the orbit due to the propagation delay time adjustments that have been made. After making the least squares fit the program operates on the Code 3 Cards which were generated by the IMP Edit program (see Figures 2 and 3). There is one of these cards per pass, containing the

*Other mathematical expressions have been suggested by Cyrus J. Creveling of the Data Systems Division, GSFC. In particular, he feels that fitting the time curve to an analytically-generated curve of a cycloid shows promise of increasing the accuracy of the operation with an accompanying saving in computer time (see reference 5).

following information:

1. The start time of each pass
2. The corresponding satellite clock reading for the start time
3. The stop time of each pass
4. The corresponding satellite clock reading for the stop time
5. "Logging" information, e.g., ID of the file, MDT tape number, etc.

(A complete write up on the Code 3 Cards can be found in reference 16, Appendix B.)

In this operation the program obtains from the sorted MDT Time Verification Tape the ground station time for the first sequence 1 in each pass and the corresponding N and Delta and uses this N in the second degree equation that it has just generated for the orbit, i.e., equation 3, to compute a theoretical Delta. It then subtracts the actual Delta taken from the sorted MDT Time Verification tape from the theoretical Delta and algebraically adds this difference to the ground station time of the first sequence 1 in the pass.

In a similar manner the time for the last sequence 1 of the pass is adjusted to agree with the second degree equation. The adjusted start and stop times and corresponding satellite clock readings† are then punched into the Pseudo Code 3 Cards along with the logging information.

Use of The Pseudo Code 3 Cards in The Decommutation Program

The Pseudo Code 3 Cards, along with the MDTs, now provide the inputs to the Decommutation Program (see Figure 3) where information from the Pseudo Code 3 Cards is used to compute the average sequence rate per pass which in turn is used to increment the time that goes on the experimenter tape.

The average sequence rate per pass is computed by dividing the difference between the start time on the Pseudo Code 3 Card and the stop time on the Pseudo Code 3 Card by the difference between the two 30-day clock readings associated with these start and stop times. This computation results in an average sequence rate for each pass. This average sequence rate is then used by the program to update the time for each sequence

†The satellite clock (30-day clock) and N (the sequence 1 count) are related such that one can easily be computed from the other (see Appendix G).

throughout the pass using the start time as the reference time. Thus the time that is written on the experimenter tape for each pass is a linear approximation to that particular segment of the least squares fit where the pass occurs.

Actually then, the smoothed time plots which appear in Appendix B are made up of straight line segments each of which represents one pass. This feature is not readily apparent in the plots in Appendix B because the curvature is very small.

The California EDT Time Verification Program

This program is identical to the MDT Time Verification Program except that it is written to accept the California Experimenter Data Tape instead of the Master Data Tape. Its purpose is to provide a means of checking the alterations that have been made to the time through the use of the Pseudo Code 3 Cards in the Decommutation Program (the *same* time goes on all experimenter tapes, thus it is necessary to check only one of them). The smoothed time plots that appear in Appendix B are a result of plotting the plot tape that is created by this program.

This final check on the timing is an important feature of the entire scheme as the time that actually is present on the experimenters' tapes is being reviewed in this operation.

CONCLUSIONS

As is evident from Appendix A the scheme just described for detecting and correcting time errors was successfully applied throughout the useful life of IMP-I*, i.e., for almost 48 orbits, with the following results:

- A. A large percentage of the timing discrepancies were filtered out before the data was received by the experimenters. This has saved the experimenters considerable time and effort in the analysis of their data and permitted rapid evaluation of the scientific significance of the data.

- B. At least two malfunctions of the A/D line were detected that otherwise would have probably gone unnoticed.
- C. The scheme provided a means of continuously keeping track of the spacecraft telemetry system.
- D. Finally, it provided a relatively easy way in which statistics could be compiled on the timing failures when they did occur. Appendix E contains a tabulation of these timing failures through the 23rd orbit.

This scheme is currently being employed to analyze the timing associated with the data from IMP-II which was launched on October 4 of 1964.

REFERENCES

1. BEAUCHAMP, K. C., *The Encoding, Telemetry and Processing of Data From the U. K. Satellite "Ariel"*, Space Research Management Unit—RAE, Farnborough, England, September 1963.
2. BUTLER, PAUL, *IMP Weekly Progress Report for the Period Ending 13 May 1964*, Greenbelt, Maryland, Goddard Space Flight Center, May 14, 1964.
3. BUTLER, PAUL, *IMP Weekly Progress Report for the Period Ending 3 June 1964*, Greenbelt, Maryland, Goddard Space Flight Center.
4. CARR, FRANK A., *IMP-I Shadow Fact Sheet*, Greenbelt, Maryland, Goddard Space Flight Center, May 7, 1964.
5. CREVELING, C. J., *Time Correction Programs for the IMP Information Processing System*, Memorandum, Greenbelt, Maryland, Goddard Space Flight Center, May 13, 1964.
6. HANEY, HERBERT E., and EDWARD SCHOONOVER, *IMP Master to Experimenter Data Tape FP 010A and FP 010B*, Greenbelt, Maryland, Goddard Space Flight Center, November 1, 1963.
7. HOUSE, CLARENCE B., "Magnetic Core Flywheel and Synchronizer", *Transactions of the 1964 International Space Electronics Symposium*.
8. International Geophysics Bulletin, *Initial Results from the First Interplanetary Monitoring Platform (IMP-I)*, National Academy of Sciences, Number 84, June 1964.
9. LOKERSON, DONALD C., "Synchronization and Control of the IMP Information Processing System", *Transactions of the 1964 International Space Electronics Symposium*.
10. LONGANECKER, GERALD W., *IMP, S-74 Weekly Progress Report for Period Ending 11 December, 1963*, Greenbelt, Maryland, Goddard Space Flight Center, December 12, 1963.

*The *Pseudo Code 3 Card Program* was written after the launch of IMP-I when the initial "raw time" plots from the first few orbits of data indicated that the procedure for correcting and smoothing the time discussed in this paper would be a feasible approach to the problem of time errors, consequently, the first eleven orbits were not fit. In addition, orbits 42 and 43 were not fit due to lack of transmission from the spacecraft during portions of these orbits.

11. MEYERSON, HERBERT, *Fourth Fact Sheet, Explorer XIV*, Greenbelt, Maryland, Goddard Space Flight Center, October 17, 1963.
12. *Minitrack System Instruction Manual For 136 Mc Minitrack Interferometer System*, Volume I of II, NASA.
13. MISH, WILLIAM H., *A Description of the IMP-A (S-74) Ground Station Time Error Detection and Rectification Software FP-026, 030, 033 and 040, X-612-64-128*, Greenbelt, Maryland, Goddard Space Flight Center, March, 1964.
14. NESS, NORMAN F., *IMP Information Processing*, Greenbelt, Maryland, Goddard Space Flight Center, June 29, 1962.
15. ROCHELLE, R. W., *Pulse Frequency Modulation*, NASA Technical Note D-1421, Greenbelt, Maryland, Goddard Space Flight Center, October 1962.
16. WHITE, HOSEA, *Description of IMP Edit Program FP 08A and FP 08B (Preliminary) and Appendices*, Greenbelt, Maryland, Goddard Space Flight Center, January 30, 1964.
17. WHITE, HOSEA, *On the Design of PFM Telemetry Encoders*, NASA Technical Note D-1672, Greenbelt, Maryland, Goddard Space Flight Center, June 1964.
18. WOLFGANG, JOHN L., JR., "Design Philosophy of the IMP Information Processing System," *Transactions of the 1964 International Space Electronics Symposium*.

APPENDIX A

“Raw” time plots of DELTA versus N (The Sequence 1 Count) for the useful life of IMP-I

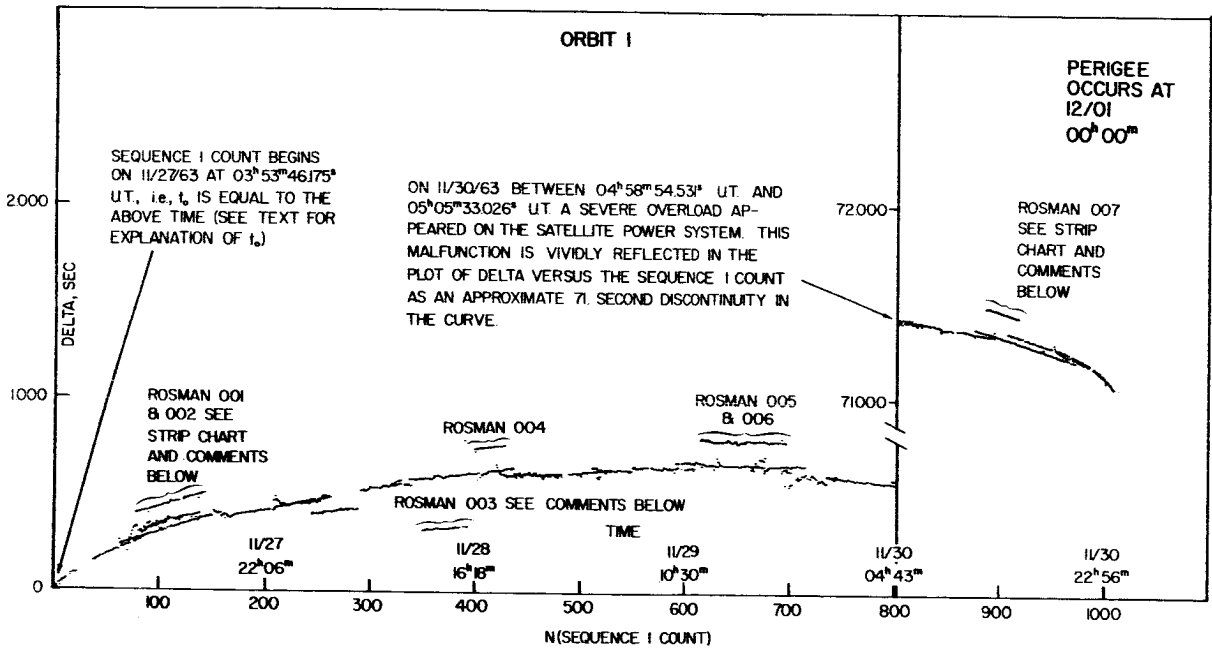


FIGURE A-1.

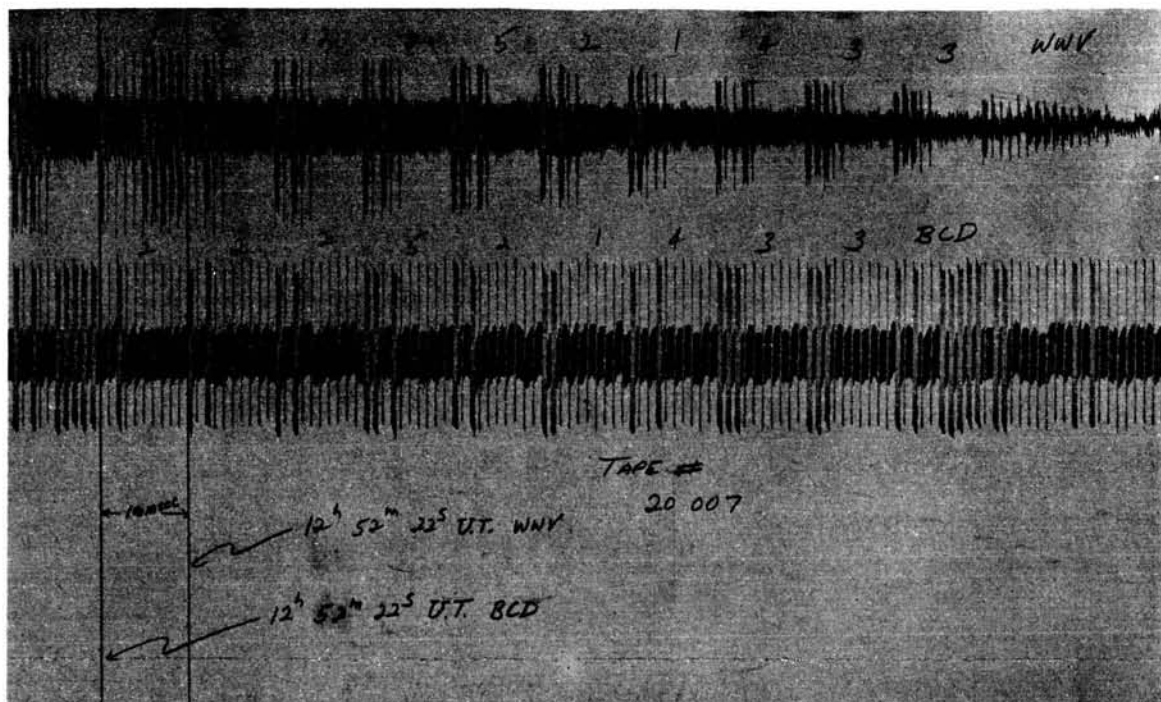


FIGURE A-2.

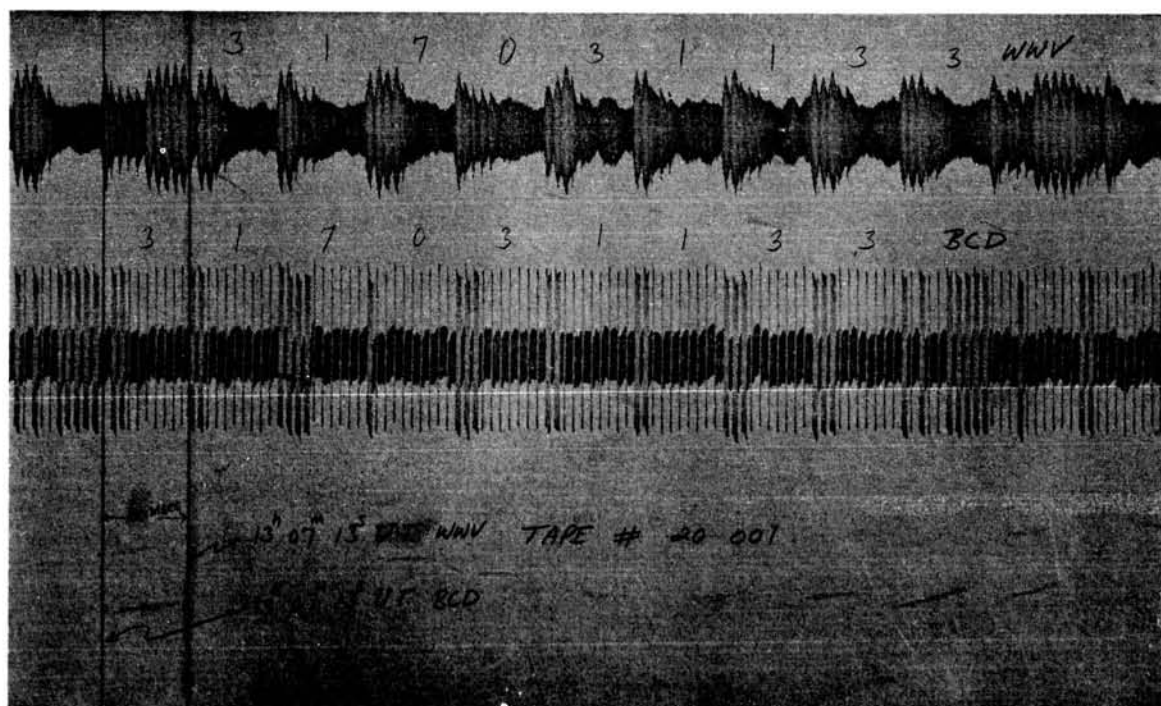


FIGURE A-3.

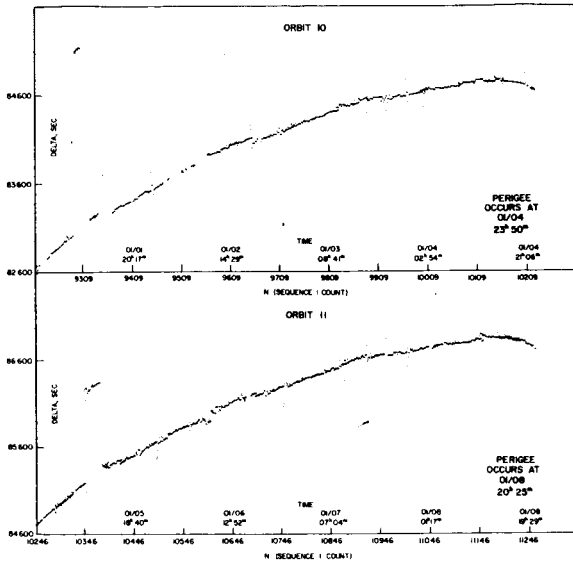


FIGURE A-8.

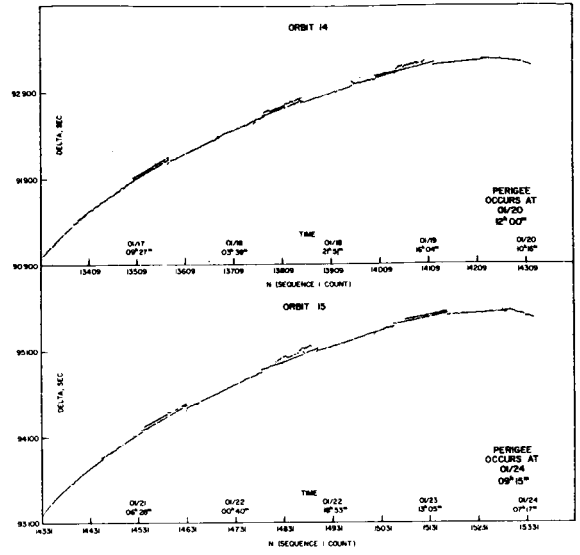


FIGURE A-10.

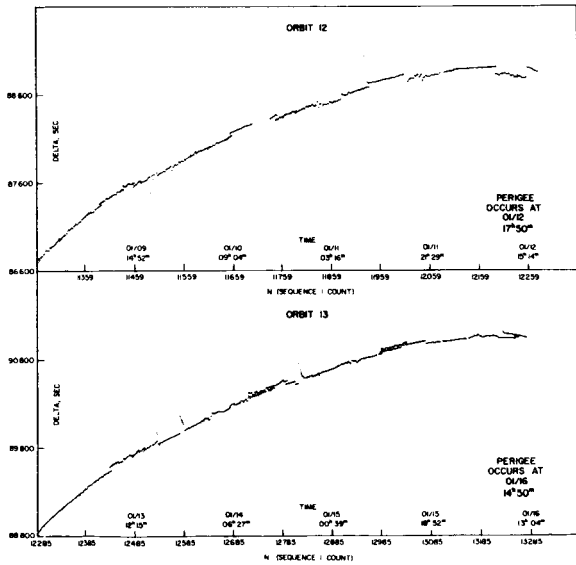


FIGURE A-9.

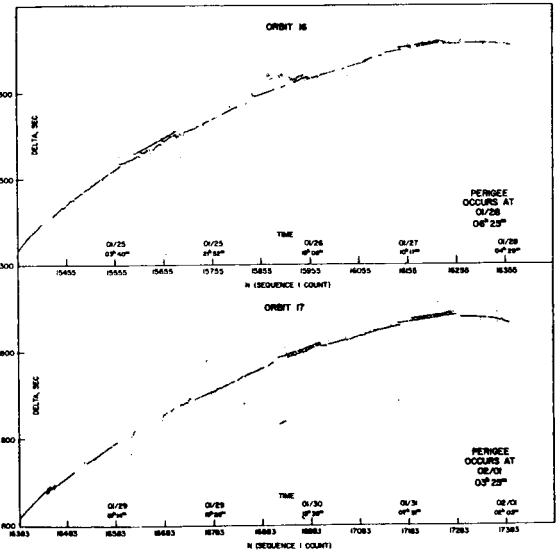


FIGURE A-11.

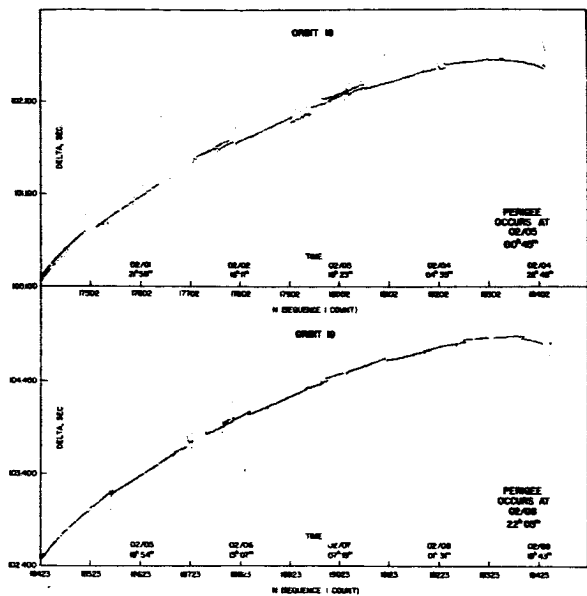


FIGURE A-12.

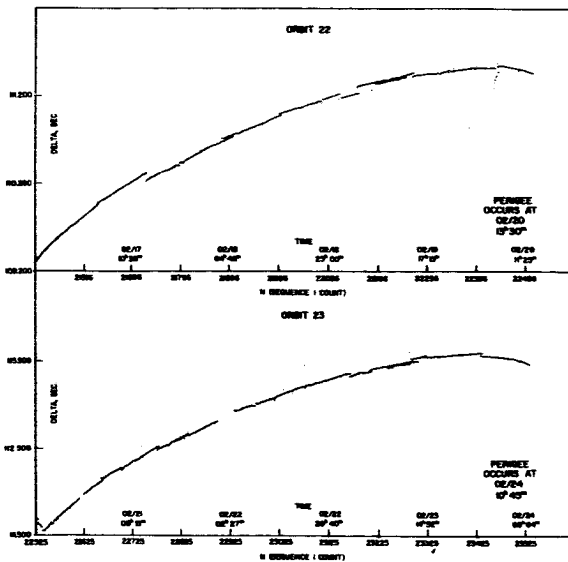


FIGURE A-14.

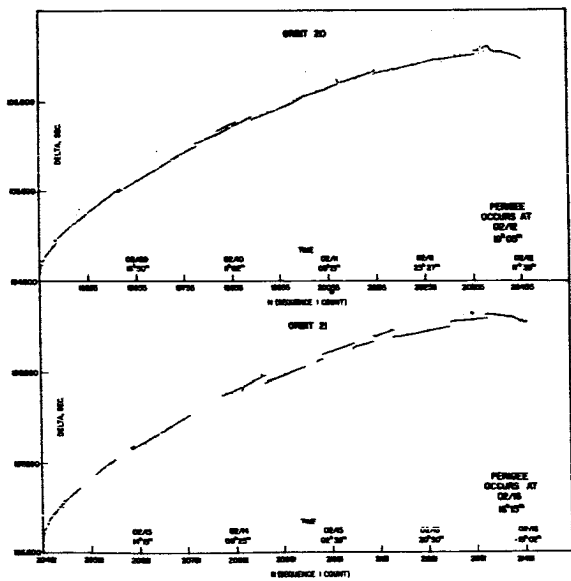


FIGURE A-13.

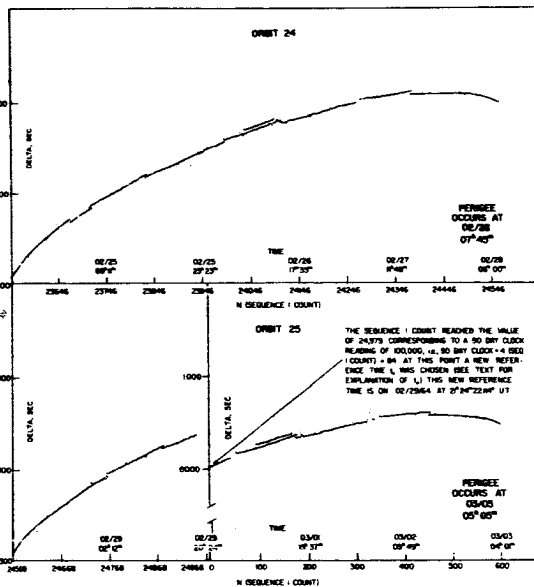


FIGURE A-15.

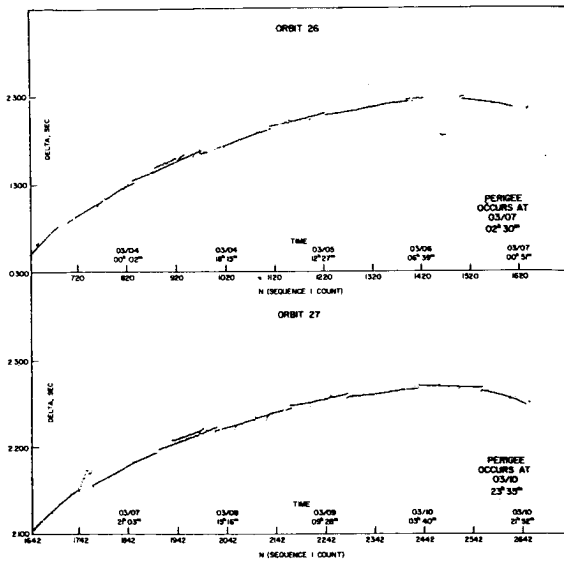


FIGURE A-16.

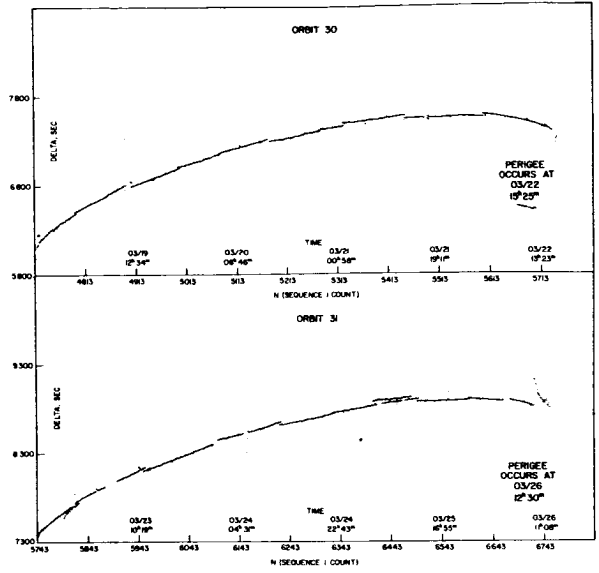


FIGURE A-18.

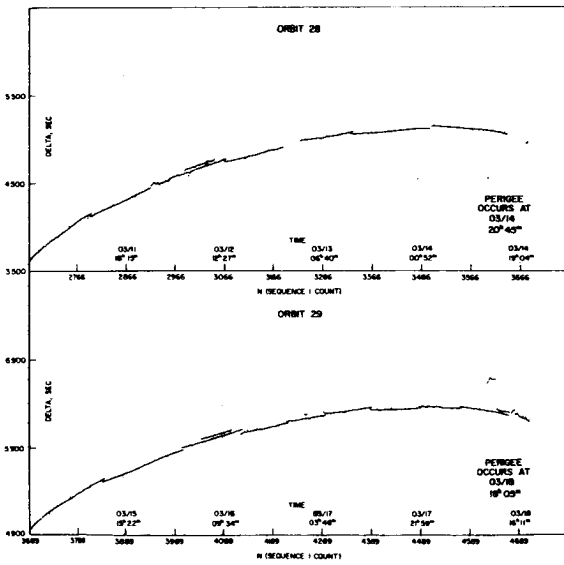


FIGURE A-17.

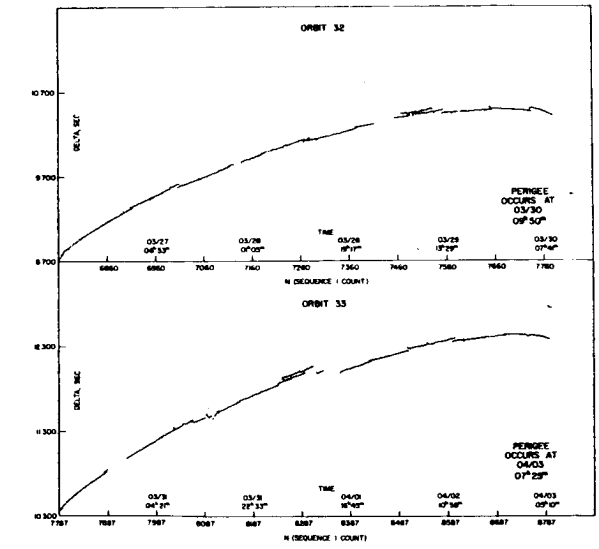


FIGURE A-19.

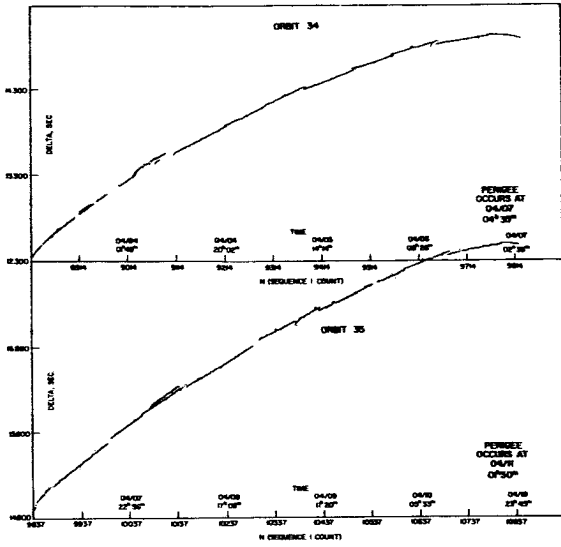


FIGURE A-20.

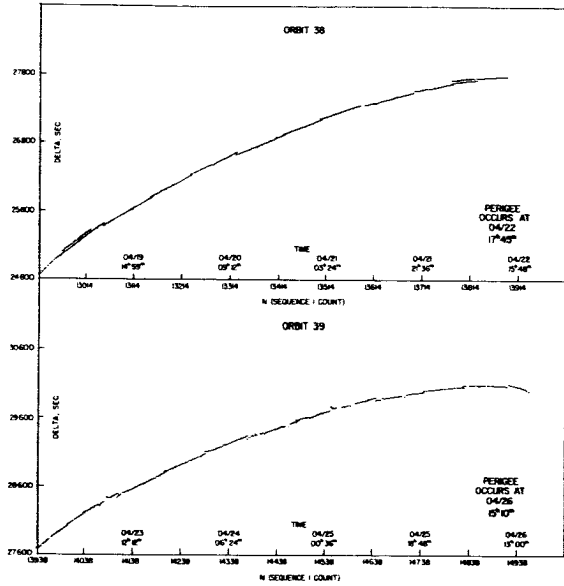


FIGURE A-22.

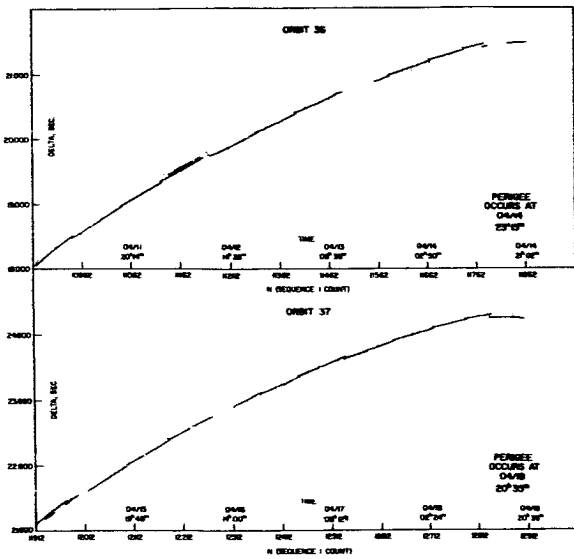


FIGURE A-21.

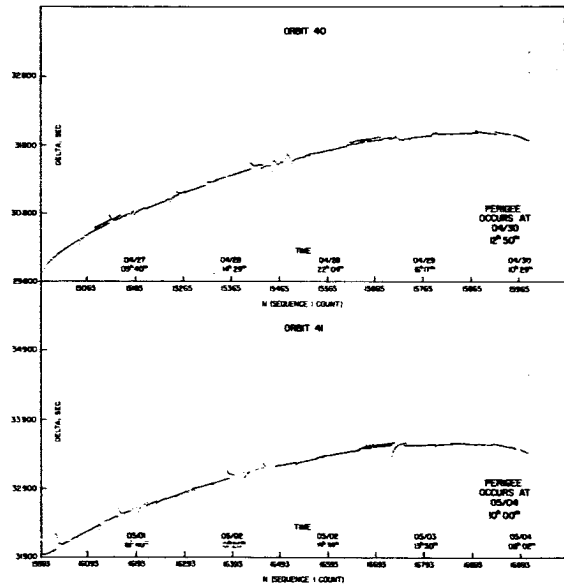


FIGURE A-23.

APPENDIX B

“Raw” time plots of DELTA versus N for orbits 36 and 37 compared with the corrected and smoothed plots

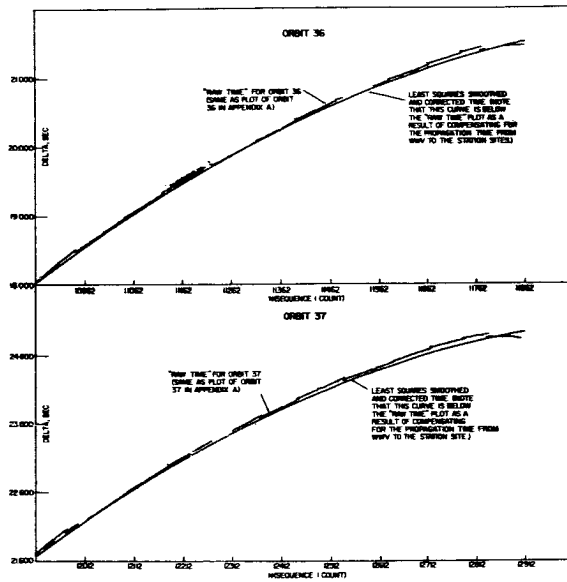


FIGURE B-1.

APPENDIX C

Telemetry system stability plots

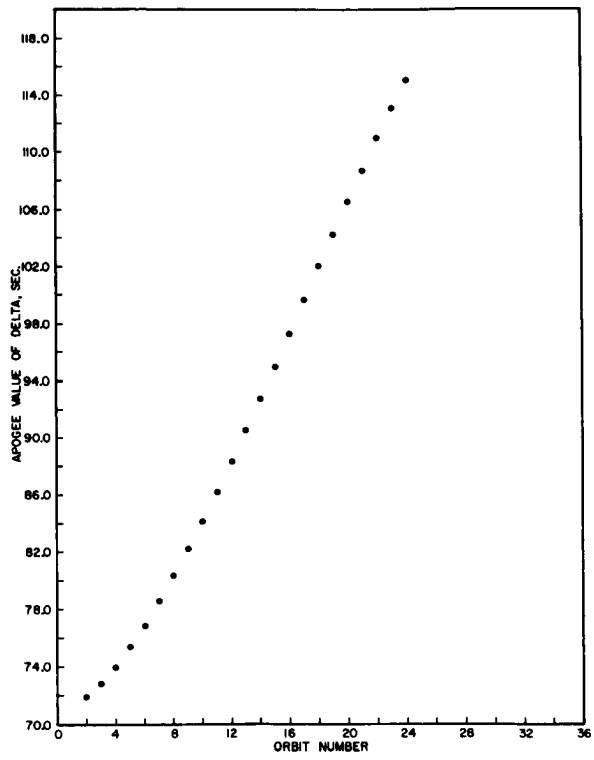


FIGURE C-1.

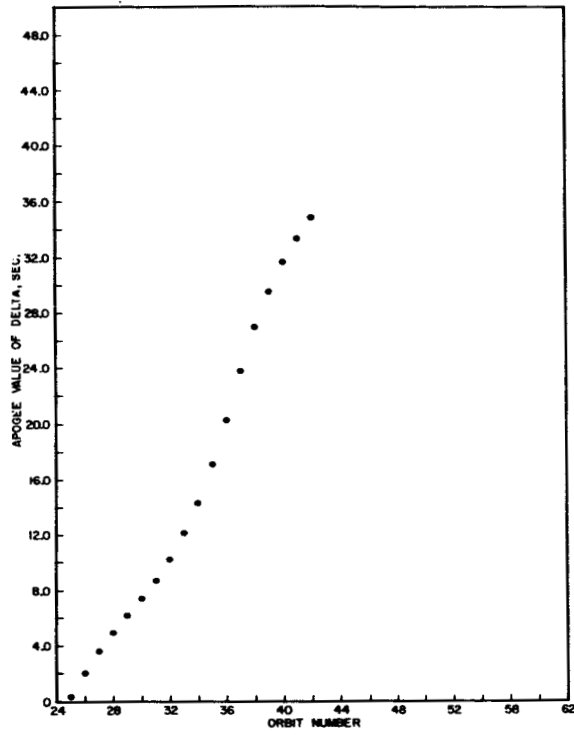


FIGURE C-2.

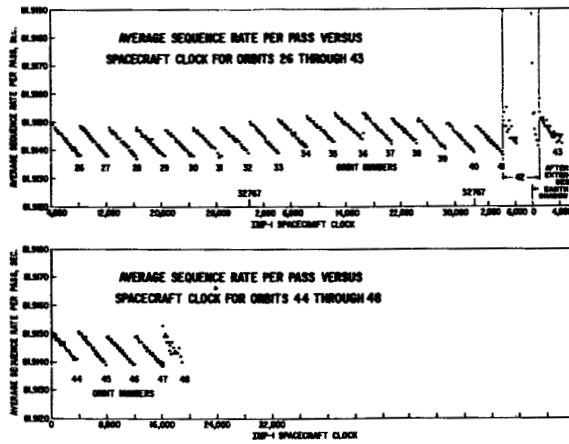


FIGURE C-3.

APPENDIX D

Propagation delay times from WWV to the Space Tracking and Data Acquisition Station Site

*Propagation Delay Times From WWV To The Space Tracking and Data Acquisition Station Site**

Station No.	Station	Propagation Time† Milliseconds
01.....	Blossom Point, Maryland.....	.5
03.....	Fort Myers, Florida.....	5.8
05.....	Quito, Ecuador.....	16.1
06.....	Lima, Peru.....	20.2
07.....	Antofagasta, Chile.....	24.6
08.....	Santiago, Chile.....	28.7
12.....	St. Johns, Newfoundland.....	8.1
13.....	Fairbanks, Alaska.....	27.4
14.....	East Grand Forks, Minnesota.....	7.2
15.....	Winkfield, England.....	20.8
16.....	Johannesburg, South Africa.....	45.7
17.....	Goldstone Lake, California.....	12.6
18.....	Woomera, Australia.....	59.1
19.....	Gilmore Creek, Alaska.....	27.4
20.....	Rosman, North Carolina.....	3.6
55.....	South Point, Hawaii.....	28.2
56.....	Ascension Island.....	30.1
61.....	College Park, Maryland.....	0.0

*Propagation times calculated by the Network Timing Section, Network Engineering and Operations Division, Goddard Space Flight Center.

†Includes .3ms for WWV Receiver; Accurate to within ± 5 ms.

APPENDIX E

Timing failure rates on IMP-I for the first 23 orbits*

- I. Station clock set *incorrectly* with WWV.
14/855 = 1.6 percent.
 - II. BCD time present and acceptable at start of pass but found *deficient* at some later time in the pass.
10/855 = 1.2 percent.
 - III. BCD time code found *unacceptable*—pass processed with serial decimal time.
1/855 = 0.2 percent.
 - IV. Pass contained *no usable* time code—pass processed with relative time.
8/855 = 0.9 percent.
 - V. A/D line could not synchronize properly with data because of a weak or noisy telemetry signal—this adversely affected the timing.
13/855 = 1.5 percent.
 - VI. Cause of unacceptable timing unknown.
4/855 = 0.5 percent.
- Total number of passes which contained a timing failure.
50/855 = 5.8 percent.

*Statistics compiled on a pass by pass basis, i.e., a pass is defined as one uninterrupted data recording made at a single station, not by analog tape. This policy was adopted because one analog tape can and often does contain more than a single pass.

Statistics on timing failures were compiled during the entire time the Fields and Plasmas Branch had operational responsibility for IMP-IPS, this constituted approximately the first 23 orbits.

APPENDIX F

Sample calculation for DELTA and N .

Sample sequence 1 recorded at Woomera, Australia on January 8, 1964 (this information obtained from the ID of the file) at $20^h 39^m 51.776^s$ U.T. during orbit number 12.

From equation (1) in text:

$$N = \left[\frac{t_n - t_0}{\Delta t} \right]$$

(Note that the brackets indicate that N has been rounded to an integer.)

Where:

$$\Delta t = 327.656 \text{ sec.}$$

$$t_0 = \text{November 27, 1963 at } 03^h 53^m 46.175^s$$

$$t_n = \text{January 8, 1964 at } 20^h 39^m 51.766^s$$

All time is referenced from January 1, 1963 thus:

$$t_0 = 86,400 \frac{\text{Sec.}}{\text{Day}} \times (330 \text{ Days})$$

$$+ 3,600 \frac{\text{Sec.}}{\text{Hour}} \times (3 \text{ Hours})$$

$$+ 60 \frac{\text{Sec.}}{\text{Min.}} \times (53 \text{ Min.})$$

$$+ 46.175 \text{ sec.}$$

$$t_0 = 28,526,026.175 \text{ sec.}$$

$$t_n = 86,400 \frac{\text{Sec.}}{\text{Day}} \times (372 \text{ Days})$$

$$\times 3,600 \frac{\text{Sec.}}{\text{Hour}} \times (20 \text{ Hours})$$

$$+ 60 \frac{\text{Sec.}}{\text{Min.}} \times (39 \text{ Min.})$$

$$+ 51.776 \text{ Sec.}$$

$$t_n = 32,215,191.776$$

$$N = \left[\frac{32,215,191.776 - 28,526,026.175}{327.656} \right]$$

$$N = 11,259$$

From equation (2) in text:

$$\text{DELTA} = (t_n - t_0) - N(\Delta t)$$

$$\text{DELTA} = 3,689,165.601 - 11,259 \times (327.656)$$

$$\text{DELTA} = +86.697 \text{ Sec.}$$

APPENDIX G

A comparison of the 30-Day Clock, 90-Day Clock and the Sequence 1 Count for approximately 6 months of satellite operation.

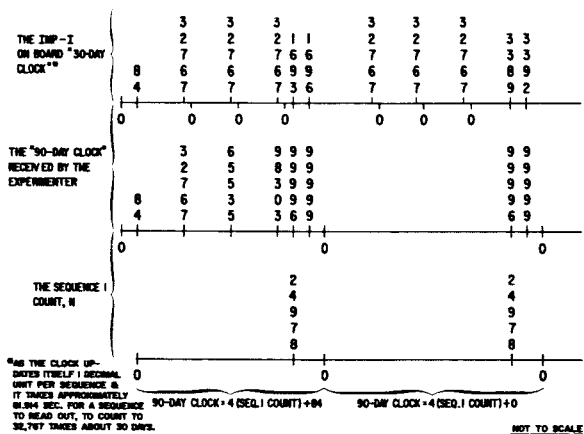


FIGURE G-1.

GEOS-1 LASER RETROREFLECTOR DESIGN AND PRELIMINARY SIGNAL CALCULATIONS*

H. H. PLOTKIN

INTRODUCTION

GEOS will be attitude stabilized along the gradient of the gravitational field, so that the normal to the bottom plane surface will point toward the center of the earth. On this bottom plane will be mounted an array of cube-corner prisms, each with an entrance aperture approximately 1 inch in diameter. These will form a retroreflecting surface and will enhance the reflected energy received in the vicinity of a ground laser source directed at the satellite. The aims of this paper are (a) to discuss the advisability of tilting the symmetry axes of the cube-corners at various angles to the plane normal in an attempt to equalize the effective reflecting area presented to rays incident over a large range of angles, and (b) to perform preliminary signal calculations which would be useful for the design of ground stations.

VARIATION OF EFFECTIVE REFLECTING AREA WITH ANGLE OF INCIDENCE

As a satellite moves over the laser reflection station during a pass, the angle between the line of sight and the satellite symmetry axis will change, as shown in Fig. 1. The individual cube-corner prisms will be identical to those used on the S-66 Beacon Explorer Satellite in which the corners of the triangular front face have been cut off to form an aperture in the form of a regular hexagon. For those prisms, the effective reflecting area has previously been calculated by P. O. Minott (Monthly Research and Advanced Technological Development Activity Report for April 1963). In addition to the angle of incidence, α , the variation of area depends upon the width across flats, w , the depth of the prism, d , (which can be

*Published as *Goddard Space Flight Center Document X-524-64-205*, July 1964.

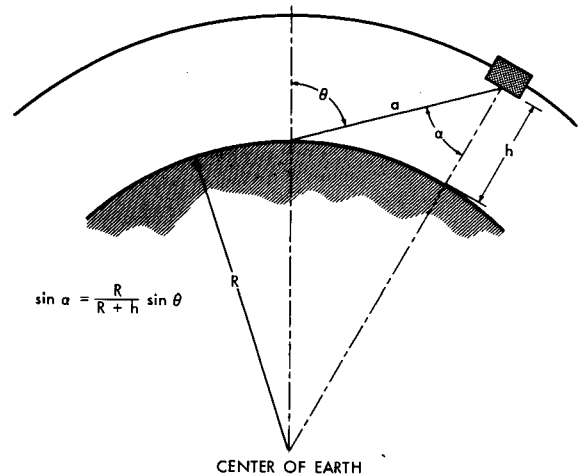


FIGURE 1.—Geometry of a satellite pass.

expressed in terms of w) and its index of refraction, N . When rotated about an axis parallel to one of the sides of the hexagon, the expression for the effective area can be written

$$A = (w \tan 30^\circ) \cos \alpha \left[(w - 2d \tan \alpha') + (w - 2d \tan \alpha')^2 \frac{\sin 30^\circ}{w} \right]$$

where

$$\alpha' = \arcsin \left(\frac{\sin \alpha}{N} \right) \quad (1)$$

A plot of the relative projected area versus angle of incidence is presented in Fig. 2, in which the area at normal incidence has been normalized to unity. The curve resulting from rotation about a corner of the hexagon instead of a side does not vary significantly from this.

In order to maximize the angle of the cone about zenith in which a ground station would find effective laser tracking possible, it has been suggested that the GEOS reflector consist of an

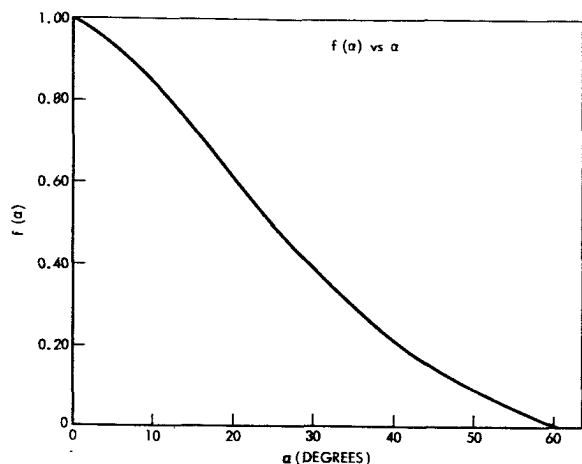


FIGURE 2.— $f(\alpha)$ vs. α . Relative effective area of a single fuzed quartz cube-corner as a function of the angle between incident ray and symmetry axis.

array of cube corners with their axes tilted at a variety of angles with respect to the satellite symmetry axis. Thus, if the distribution of tilt angles were chosen properly, it was hoped that the satellite might present a reasonably large effective area even when the zenith distance, θ , was relatively large. Further consideration of the proposal shows that there is little or nothing to be gained by pursuing this approach.

Suppose the zenith distance were such that the line of sight made an angle α_1 with the normal to the bottom surface. Then, if all the cube corners were parallel and pointing down, the relative projected area would be f_1 (see Fig. 3). Now, if a group of prisms were tilted so as to point closer to the line of sight, making an angle α_2 with the laser beam, these reflectors would have a larger effective relative area f_2 . At the same time, it is clear that another equal group of prisms would have to be tilted away from the line of sight by the same angle in order to maintain rotational symmetry about the satellite axis. These reflectors would present the smaller relative area f_3 . Where the function $f(\alpha)$ is linear, the loss of area in the group tilted away from the laser just offsets the gain experienced by those tilted in the favorable direction. Where the second derivative of $f(\alpha)$ is negative (near zenith), the result is a net loss of reflecting area, and where the second derivative of $f(\alpha)$ is positive, (nearer to the horizon) there is a net gain.

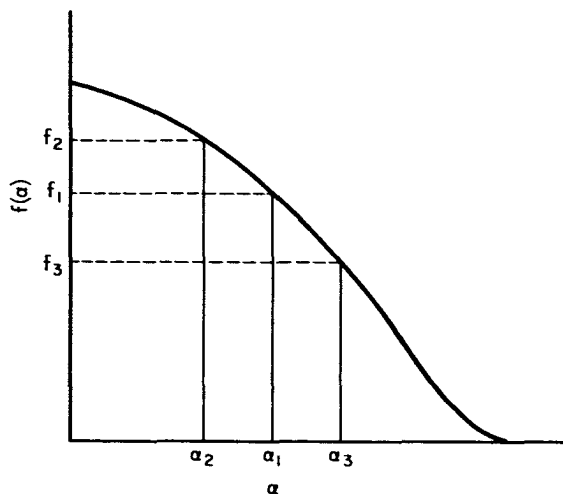


FIGURE 3.—The effect of tilting groups of cube-corner prisms.

The small improvement in effective area that might be experienced at large zenith distances by distributing the cube corner directions in this manner does not contribute appreciably to the usefulness of the satellite because other parameters which also determine the signal strength will be found to vary in a much more significant way. In particular, the received energy falls off very rapidly as the satellite moves away from zenith because of increasing range and increasing atmospheric attenuation. In addition, any attempt to tilt the cube corners without lifting them out of the plane of the bottom surface would result in shadowing effects and/or degraded geometrical packing efficiency. Our conclusions therefore are to arrange as many reflectors as possible into a mosaic on the available area, with all of them pointing parallel to the gravitational gradient.

REFLECTED SIGNAL VS. ZENITH ANGLE

We use the following parameters:

E_t = the energy in the transmitted laser pulse
 δ_t = half-angle of the transmitted beam divergence cone. We assume the energy is uniformly distributed over this cone.

$A_0 f(\alpha)$ = effective retro-reflecting area of the array, for light incident at an angle α to the symmetry axis. A_0 is the maximum projected area and $f(0) = 1$. Fig. 2 is a plot of $f(\alpha)$.

ρ = effective reflectivity

δ_c = half-angle of the reflected beam divergence. Again, we assume that the energy is uniformly distributed in this cone.
 $\tau(\theta)$ = one-way transmission through the atmosphere for light at zenith distance θ .
 α = slant range from observer to satellite
 d = receiving telescope aperture
 η = receiving telescope transmission.

The ratio of received signal to transmitted pulse energy is then given by

$$\frac{Es}{E} = \frac{A_0 f(\alpha) [\tau(\theta)]^2 \eta d^2}{4\pi \delta_c^2 \delta_c^2 a^4} = KG(\theta) \quad (2)$$

where

$$K = \frac{\rho A_0 \eta d^2}{4\pi \delta_c^2 \delta_c^2} \quad (3)$$

is a constant, and

$$G(\theta) = \frac{f(\alpha) [\tau(\theta)]^2}{a^4} \quad (4)$$

is a variable over the duration of a pass since it is a function of zenith distance θ . From Fig. 1 and the law of cosines,

$$a = -R \cos \theta + [R^2 \cos^2 \theta + h(2R+h)]^{1/2} \quad (5)$$

From the law of sines,

$$\sin \alpha = \frac{R}{R+h} \sin \theta \quad (6)$$

These equations enable us to calculate expected signal levels for a satellite at different positions in the sky.

VARIATION OF ATMOSPHERIC TRANSMISSION

To estimate the effect of zenith angle on atmospheric transmission, we assume the very rough model shown in Fig. 4. The atmosphere here is a uniformly dense layer of thickness T . A light beam passing once through the entire atmosphere in a vertical direction will be attenuated so that its intensity is reduced to a fraction τ_0 of its initial value. At a zenith angle θ , the atmospheric path length is $t(\theta)$ and the one-way transmission is

$$\tau(\theta) = \tau_0 \frac{t(\theta)}{T} \quad (7)$$

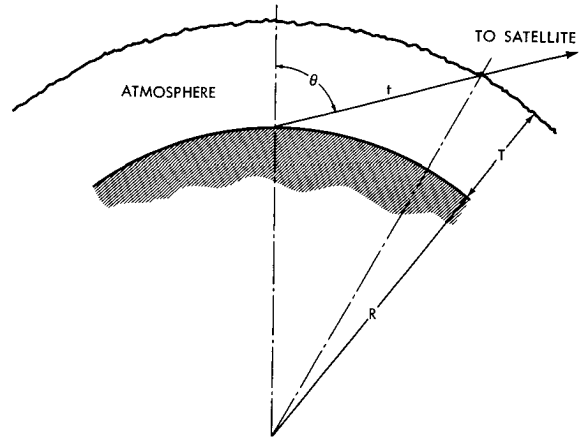


FIGURE 4.—Atmospheric attenuation geometry.

$$\text{where } t(\theta) = \left\{ -R \cos \theta + [R^2 \cos^2 \theta + T(2R+T)]^{1/2} \right\} \quad (8)$$

Note that in the limit of a flat earth ($R \rightarrow \infty$), equation (6) may be written

$$\begin{aligned} t(\theta) &= R \cos \theta \left\{ \left[1 + \frac{2T}{R \cos^2 \theta} + \frac{T^2}{R^2 \cos^2 \theta} \right]^{1/2} - 1 \right\} \\ &= R \cos \theta \left\{ 1 + \frac{T}{R \cos^2 \theta} + \frac{T^2}{R^2} \left(\frac{1}{2 \cos^2 \theta} - \frac{1}{2 \cos^4 \theta} \right) \right. \\ &\quad \left. + \dots - 1 \right\} t(\theta) \doteq T \sec \theta \quad (9) \end{aligned}$$

and

$$\tau(\theta) \doteq \tau_0 \sec \theta \quad (10)$$

This is the usual expression used by astronomers for atmospheric transmission. It is only accurate for zenith angles such that

$$\left| \frac{T^2}{R^2} \left(\frac{1}{2 \cos^2 \theta} - \frac{1}{2 \cos^4 \theta} \right) \right| < 0.1 \frac{T}{R \cos^2 \theta}$$

or

$$\sec^2 \theta < 1 + 0.2 \frac{R}{T} \quad (11)$$

If we take for the thickness of our model atmosphere $T = 25$ km, and for the radius of the earth $R = 6,310$ km, we find that the condition in eq. (11) is satisfied by

$$\theta < 80^\circ \quad (12)$$

Since we will never operate at zenith angles even as large as this, we may as well use eq. (10) instead

of the more complicated expressions (7) and (8).

The value of τ_0 is chosen to be 0.7, from several conservative estimates in the literature. Fig. 5 is the resulting variation of atmospheric transmission with zenith angle.

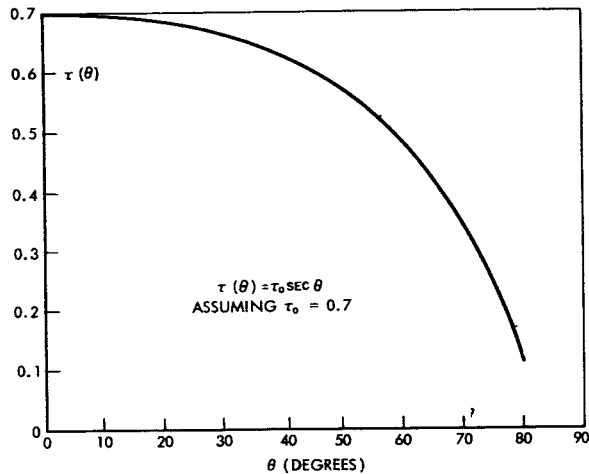


FIGURE 5.—One-way atmospheric transmission vs. zenith angle. $\tau(\theta) = \tau_0 \sec \theta$
Assuming $\tau_0 = 0.7$.

EXAMPLE OF NUMERICAL SIGNAL CALCULATIONS

The expected orbit for GEOS-I is one with a perigee of 600 nautical miles and an apogee of about 1200 nautical miles. We choose an altitude of 1500 km (about 800 nautical miles) as a typical example because it results in fairly large off-axis incidence angles α so we can study the drop-off in effective area. The corresponding relation between α and zenith angle θ is shown in Fig. 6. When the effective reflecting area is then plotted as a function of θ instead of α , as required in eq. (4), it appears as in Fig. 7. The last of the variable factors, slant-range, a , is plotted in Fig. 8.

We next turn to a calculation of the constant K in eq. (3). The specification on the cube corner reflectors is that 50 percent of the light incident on the effective reflecting area shall be retro-reflected within a cone 20 seconds of arc in diameter. We thus have

$$\rho = 0.5$$

$$\delta_c = 10 \text{ sec} = 5 \times 10^{-5} \text{ radians.}$$

The transmitted laser beam will be taken with a divergence equal to that planned for S-66 since it seems to be a good compromise between aiming ability and energy concentration.

$$\delta_t = 5 \times 10^{-4} \text{ radians}$$

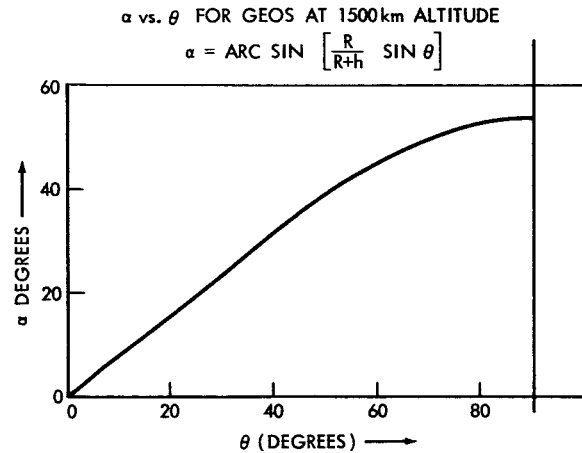


FIGURE 6.— α vs. θ for GEOS at 1500 km altitude
 $\alpha = \arcsin \left[\frac{R}{R+h} \sin \theta \right]$

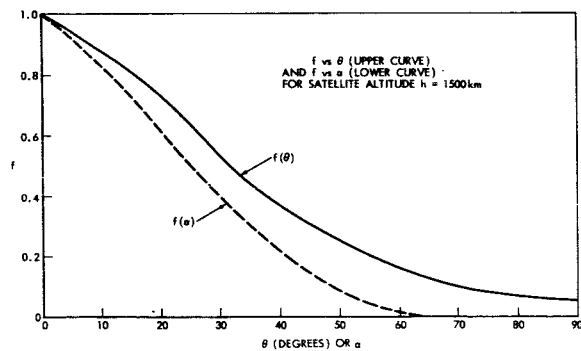


FIGURE 7.— f vs. θ (upper curve) and f vs. α (lower curve) for satellite altitude $h = 1500$ km.

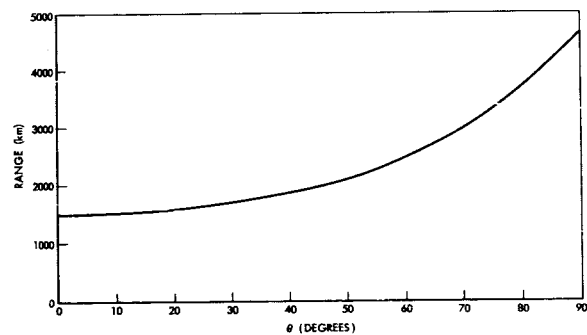


FIGURE 8.—Range vs. zenith angle for 1500 km altitude.

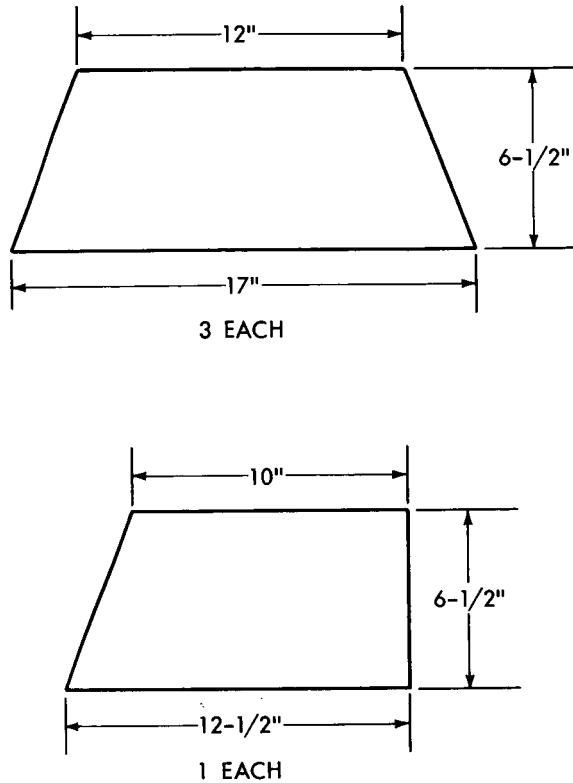


FIGURE 9.—Approximate geometry of reflector panels.

The receiving telescope is assumed to have an aperture with a diameter

$$d = 16 \text{ inches} = 0.4064 \text{ meters}$$

and the transmission of the receiving optics is

$$\eta = 0.5$$

The reflector arrays will be fitted into four flat panels which take on the trapezoidal shapes shown in Fig. 9. Around the perimeter of each panel will probably be a waster strip about $\frac{1}{2}$ inch wide because of the hexagonal shape of the reflectors. The usable area of these panels adds up to about

$$A_0 = 0.178 \text{ sq. meters.}$$

When the appropriate values are inserted into eq. (3) we obtain

$$K = 0.9358 \text{ km}^4$$

The ratio of received to transmitted energy from

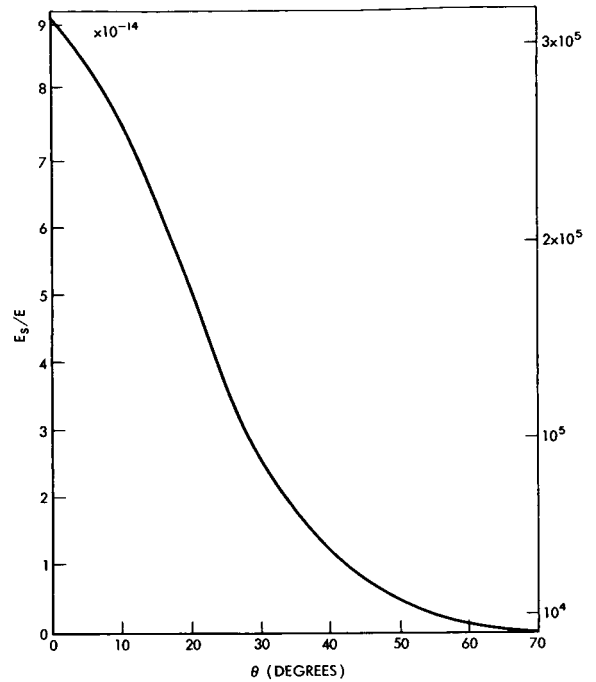


FIGURE 10.—Received signal per joule of transmitted ruby laser energy. Photons/joule.

eq. (2) is plotted in Fig. 10. The scale on the right has been converted into received photons at the sensor, per joule of energy in the transmitted ruby laser pulse.

Using a photomultiplier detector we have found that 10^4 photons results in a good signal, even against a bright sky, with proper filtering. This would allow us, from Fig. 10, to get range information from GEOS down to a zenith angle of at least 53° , with much better signals near zenith. For other geometrical arrangements, other detectors and lasers, similar calculations may be performed. However, it is already clear from the foregoing that the laser reflection panels on GEOS will be very valuable for range and photographic tracking with geodetic precision.

ACKNOWLEDGMENT

Much of this paper developed from valuable discussions with William M. Meyers of General Electric Missile and Space Division.

THERMAL RADIATION TO A FLAT SURFACE ROTATING ABOUT AN ARBITRARY AXIS IN AN ELLIPTICAL EARTH ORBIT: APPLICATION TO SPIN-STABILIZED SATELLITES*

EDWARD I. POWERS

The derivation of total thermal radiation incident upon a flat plate rotating about an arbitrary axis is presented. The functional relationships between position in an elliptical earth orbit and direct solar radiation, earth-reflected solar radiation (albedo), and earth-emitted radiation (earthshine) are included. The equations have been programmed for the IBM 7090 digital computer, resulting in solutions which relate the incident radiation to spin axis orientation and orbital position. Several representative orbits for a typical geometrical configuration were analyzed and are presented as examples.

INTRODUCTION

The satisfactory operation of an artificial satellite depends upon maintaining the payload temperature within prescribed limits. For example, the standard batteries employed in present-day spacecraft generally restrict the temperature limits to 0° and 40°C. Often experiments located within the satellite structure further restrict this variation.

The temperature level of a satellite may be determined by solving the instantaneous energy balance

$$P + S \alpha_s A_p + q_{alb} + q_{es} = \sigma \epsilon A_s \bar{T}^4 + WC_p \frac{dT}{dt}$$

where

- P = internal power dissipation,
- S = solar constant,
- α_s = solar absorptance,
- A_p = instantaneous projected area for sunlight,
- q_{alb} = reflected solar radiation (albedo),
- q_{es} = earth-emitted radiation (earthshine),
- σ = Stefan-Boltzmann constant,
- ϵ = infrared emittance,

- A_s = total surface area,
- \bar{T}^4 = mean fourth power, surface temperature,
- WC_p = heat capacity of the satellite,
- $\frac{dT}{dt}$ = time rate of change of satellite temperature.

If the average orbital temperature is being computed, the last term is dropped. In this case all heat input terms represent integrated orbital values.

It should be noted that the above equation, as a representation of the entire satellite, is greatly oversimplified. The values for ϵ and α_s generally vary over the surface and great fluctuations in skin temperature may exist. In practice the thermal analysis consists of the development of a thermal model which represents a fine mesh of interconnected isothermal nodes. The appropriate relationship between nodes in regard to thermal conduction and radiation interchange must be established.

The energy balance for each node of a thermal model may be written

$$P_n + S \alpha_{s_n} A_{p_n} + q_{alb_n} + q_{es_n} + \sum_m C_{nm} (T_m - T_n) + \sigma \sum_m E_{nm} F_{nm} (T_m^4 - T_n^4) = \sigma \epsilon A_{s_n} T_n^4 + (WC_p)_n \frac{dT_n}{dt}$$

*Published as NASA Technical Note D-2147, April 1964.

where

- C_{nm} = conductance between nodes n and m ,
 E_{nm} = effective emissivity between nodes n and m ,
 F_{nm} = shape factor-area product between nodes n and m .

Since the major interest is to determine the satellite temperature level and not the gradients within, a discussion of the terms in the first equation is in order.

For most passive controlled satellites P is relatively small compared with the total radiation input and does not have a significant effect on the satellite mean temperature. The magnitude of this effect, however, depends upon the ϵ of the surface (e.g. a surface with a low absolute ϵ may raise the internal temperature significantly because the skin has a limited capacity for re-radiation).

The remaining three heat sources, direct solar heating, earth-reflected solar heating (albedo), and earth-emitted radiation (earthshine), represent the significant inputs to the satellite. It is apparent that an adequate thermal design is predicated upon a reasonably accurate knowledge of these thermal radiation inputs. The major source of energy—direct solar radiation—is fortunately the most accurately obtained. Since the sun's rays impinging upon the satellite are virtually parallel, the problem is simply one of determining the instantaneous orientation of each external face with respect to the solar vector.

The calculation of earthshine is considerably less precise, requiring fundamental assumptions to reduce the complexity of computation. These include the assumptions that the earth is a diffusely emitting blackbody and that the surface temperature is uniform at 450°R . This permits direct calculation of energy input from all visible positions on the earth. Since all locations supply varying inputs, an integral equation must be solved to obtain the total incident energy.

The albedo determination involves a similar integration. The earth in this case is assumed to be a diffusely reflecting sphere. In addition, the source intensity is a function of the satellite location relative to the sunlit portion of the earth.

For the nonrotating satellite whose orientation is "fixed in space" the calculation of the thermal

radiation fluxes is performed at any interval during the orbit. At each orbital position a particular flat surface may or may not "see" the entire part of the earth's cap visible from the satellite. In the latter case integration for albedo and earthshine must exclude the shaded portion of the cap. Spin-stabilized satellites which rotate uniformly about the spin axis greatly complicate the analysis. The rotation around an arbitrary axis means in general that the heat fluxes vary, since the visible portion of the earth is a function of this rotation.

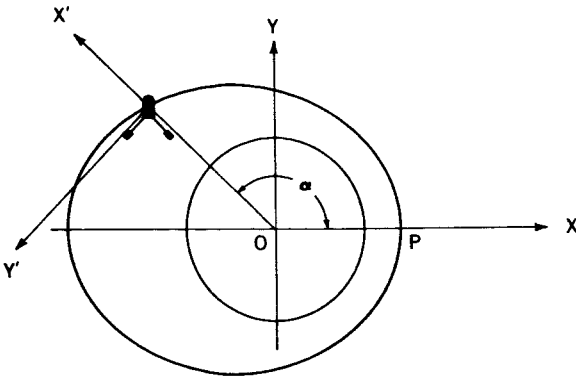
The analysis presented herein is based upon the energy impinging upon a flat surface whose orientation is defined in vector notation (by the normal vector), and which is rotating about an arbitrary axis. This has a much broader application than may be realized at first. Although the obvious application is for spin-stabilized satellites, the results apply to any body of revolution. Here the interest lies in the variation of flux about the axis rather than an average value per spin. Thus, with the orientation of a single flat plate, the impinging fluxes on cylinders and cones are obtained. A sphere or a shape with a variable surface curvature along its axis requires several or many such plates, depending upon the precision required. It can be seen that the thermal radiation to the entire satellite surface may be found by simply considering a handful of appropriately oriented flat plates.

The purpose of this paper is to present, in general form, the derivation of the governing equations for the radiation energy sources as stated. The equations refer specifically to a flat surface rotating about an axis whose orientation is arbitrary. The dependence upon orbital position is included; the results of the numerical integration of these equations, encompassing a suitable range of applicable orbits, is presented.

COORDINATE SYSTEMS AND VECTOR REPRESENTATION

The primary coordinate system is fixed with respect to the orbit (Figure 1). The XY plane lies in the plane of the orbit with the X axis coincident with the line connecting the center of the earth (the focus of the ellipse) and perigee.

As the satellite traverses the orbit, the instantaneous location is defined by the angle α .

**LEGEND**

O—EARTH CENTER
P—PERIGEE

FIGURE 1.—Coordinate systems.

The altitude, therefore, may be determined at any instant by

$$A(\alpha) = \frac{(P + R_e)(1 + e)}{1 + e \cos \alpha} - R_e \quad (1)$$

where

$A(\alpha)$ = altitude,
 P = altitude at perigee,
 R_e = radius of the earth,
 e = eccentricity of the orbit.

The orientation of the unit vectors (Figure 2) representing the normals to the flat plates, \mathbf{N} , and the solar vector, \mathbf{S} , are specified by the following angles:

β = angle between the projection of \mathbf{N} on the XY plane and the X axis,

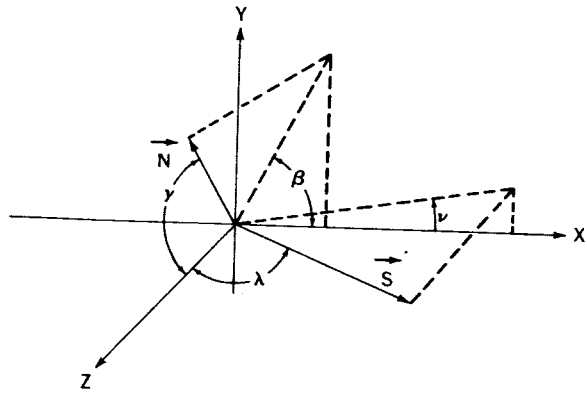


FIGURE 2.—Orientation of normal and solar vectors.

γ = angle between \mathbf{N} and the Z axis,
 ν = angle between the projection of \mathbf{S} on the XY plane and the X axis,
 λ = angle between \mathbf{S} and the Z axis.

The vectorial representations in the fixed coordinate system are thus

$$\mathbf{N} = \cos \beta \sin \gamma \mathbf{i} + \sin \beta \sin \gamma \mathbf{j} + \cos \gamma \mathbf{k} \quad (2)$$

and

$$\mathbf{S} = \cos \nu \sin \lambda \mathbf{i} + \sin \nu \sin \lambda \mathbf{j} + \cos \lambda \mathbf{k} \quad (3)$$

It is convenient to introduce an instantaneous coordinate system ($X' Y' Z'$) whose origin is fixed in the satellite (Figure 1). The $X' Y'$ plane lies in the orbital plane with the X' axis coincident with the line from the earth's center to the satellite. Z' and Z have the same orientation.

The vectors \mathbf{N} and \mathbf{S} in the primed system are

$$\mathbf{N} = \cos (\beta - \alpha) \sin \gamma \mathbf{i}' + \sin (\beta - \alpha) \sin \gamma \mathbf{j}' + \cos \gamma \mathbf{k}' \quad (4)$$

$$\mathbf{S} = \cos (\nu - \alpha) \sin \lambda \mathbf{i}' + \sin (\nu - \alpha) \sin \lambda \mathbf{j}' + \cos \lambda \mathbf{k}' \quad (5)$$

In a similar manner the expression for the spin axis vector \mathbf{A} may be shown to be

$$\mathbf{A} = \cos (\delta - \alpha) \sin \mu \mathbf{i}' + \sin (\delta - \alpha) \sin \mu \mathbf{j}' + \cos \mu \mathbf{k}' \quad (6)$$

where

δ = angle between the projection of \mathbf{A} on the XY plane and the X axis,
 μ = angle between \mathbf{A} and the Z axis.

PLATE ROTATING ABOUT THE SPIN AXIS

For spin-stabilized satellites the normal vectors \mathbf{N} , which represent the orientation of the exterior surfaces, rotate about the spin axis. In deter-

mining both instantaneous and average heat fluxes, the orientation of \mathbf{N} as a function of this rotation with respect to the primed coordinate system ($X' Y' Z'$) must be known. This is accomplished by defining a third coordinate system $X'' Y'' Z''$ (Figure 3). The X'' axis is

coincident with the spin axis \mathbf{A} . Y'' is defined so that it lies in the plane formed by X'' and an arbitrary fixed position of the normal vector \mathbf{N}_0 .*

In Figure 3, the following terms may be evaluated:

$$\cos \tau = \cos (\beta - a) \sin \gamma \cos (\delta - a) \sin \mu + \sin (\beta - a) \sin \gamma \sin (\delta - a) \sin \mu + \cos \gamma \cos \mu, \quad (7)$$

$$\sin \tau = \sqrt{1 - \cos^2 \tau},$$

$$\mathbf{N} = \cos \tau \mathbf{i}'' + \sin \tau \cos \kappa \mathbf{j}'' + \sin \tau \sin \kappa \mathbf{k}'' \quad (8)$$

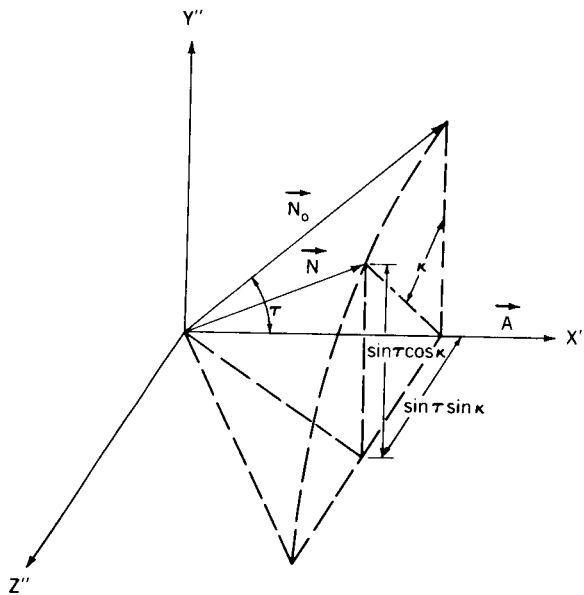


FIGURE 3.—Double-primed coordinate system.

It can be seen that the angle between the spin axis and \mathbf{N}_0 can be computed only in the first and second quadrants of the $X'' Y''$ plane. This has little effect since κ varies from 0 to 2π (a rotation). [If the vector \mathbf{N}_0 lies in the third quadrant (of the $X'' Y''$ plane) the initial position for rotation remains in the second quadrant.] The objective now is to relate the double-primed unit vectors to the primed vectors in order to determine \mathbf{N} in the primed system.

The projections of \mathbf{j}'' and \mathbf{k}'' in the primed coordinate system are found in the following manner: The value for the unit vector \mathbf{k}'' is

$$\mathbf{k}'' = \frac{\mathbf{A} \times \mathbf{N}_0}{|\mathbf{A} \times \mathbf{N}_0|} = \frac{g\mathbf{i}' + h\mathbf{j}' + l\mathbf{k}'}{\sqrt{g^2 + h^2 + l^2}}, \quad (9)$$

where

$$g = \sin (\delta - a) \sin \mu \cos \gamma - \cos \mu \sin (\beta - a) \sin \gamma,$$

$$h = \cos \mu \cos (\beta - a) \sin \gamma - \cos (\delta - a) \sin \mu \cos \gamma,$$

$$l = \cos (\delta - a) \sin \mu \sin (\beta - a) \sin \gamma - \sin (\delta - a) \sin \mu \cos (\beta - a) \sin \gamma$$

Similarly

$$\mathbf{j}'' = \frac{(\mathbf{A} \times \mathbf{N}_0) \times \mathbf{A}}{|(\mathbf{A} \times \mathbf{N}_0) \times \mathbf{A}|} = \frac{m\mathbf{i}' + n\mathbf{j}' + o\mathbf{k}'}{\sqrt{m^2 + n^2 + o^2}} \quad (10)$$

*Since a complete rotation occurs, the orientation of \mathbf{N}_0 merely indicates the starting point.

where

$$m = h \cos \mu - l \sin (\delta - \alpha) \sin \mu ,$$

$$n = l \cos (\delta - \alpha) \sin \mu - g \cos \mu ,$$

$$o = g \sin (\delta - \alpha) \sin \mu - h \cos (\delta - \alpha) \sin \mu .$$

\mathbf{N} may now be determined in the primed coordinate system ($X' Y' Z'$). By rewriting the double-primed unit vectors,

$$\left. \begin{aligned} i'' &= pi' + qj' + rk' \\ j'' &= si' + tj' + uk' \\ k'' &= vi' + wj' + xk' \end{aligned} \right\} , \quad (11)$$

where

$$p = \cos (\delta - \alpha) \sin \mu ,$$

$$q = \sin (\delta - \alpha) \sin \mu ,$$

$$r = \cos \mu ,$$

$$s = \frac{m}{\sqrt{m^2 + n^2 + o^2}} ,$$

$$t = \frac{n}{\sqrt{m^2 + n^2 + o^2}} ,$$

$$u = \frac{o}{\sqrt{m^2 + n^2 + o^2}} ,$$

$$v = \frac{g}{\sqrt{g^2 + h^2 + l^2}} ,$$

$$w = \frac{h}{\sqrt{g^2 + h^2 + l^2}} ,$$

$$x = \frac{l}{\sqrt{g^2 + h^2 + l^2}} .$$

By substituting for i'' , j'' , and k'' in Equation 8

$$\mathbf{N} = ai' + bj' + ck' , \quad (12)$$

where

$$a = p \cos \tau + s \sin \tau \cos \kappa + v \sin \tau \sin \kappa ,$$

$$b = q \cos \tau + t \sin \tau \cos \kappa + w \sin \tau \sin \kappa ,$$

$$c = r \cos \tau + u \sin \tau \cos \kappa + x \sin \tau \sin \kappa ,$$

where κ is the angle of rotation. \mathbf{N} may now be evaluated at any position during rotation throughout the orbit in the primed system ($X' Y' Z'$).

DETERMINATION OF EARTHSHINE

The general expression for the radiation intensity dq , from a diffusely emitting source dA_e , impinging upon a unit area, is as follows:

$$dq = \frac{I \cos \omega \cos \eta dA_e}{D^2} , \quad (13)$$

where

I = source intensity,

ω = angle between the line connecting the unit area with dA_e and the normal to dA_e (\mathbf{N}_e),

η = angle between the line connecting dA_e with the unit area and the normal to the unit area,

D = distance between the unit area and dA_e .

With the application of this equation to a flat plate of unit area at an altitude $A(\alpha)$, the radiation intensity impinging upon the plate is

$$q_E = \frac{\sigma T_e^4}{\pi} \int_0^{\theta_m} \int_0^{2\pi} \frac{\cos \omega \cos \eta dA_e}{D^2} \quad (14)$$

where

σ = Stefan-Boltzmann constant,

T_e = mean black body temperature of the earth's surface,

σT_e^4 = the heat flux leaving the source dA_e .

From the geometry of Figure 4 (which is similar to a sketch used by Katz*) and the vector notation presented, the terms in the equation may be

*Katz, A. J., "Determination of Thermal Radiation Incident upon the Surfaces of an Earth Satellite in an Elliptical Orbit," Grumman Aircraft Engineering Corp., Rept. XP 12.20, May 1960.

defined more specifically:

$$\begin{aligned}
 \mathbf{D} &= -H\mathbf{i}' + R_e \sin\theta \cos\phi \mathbf{j}' + R_e \sin\theta \sin\phi \mathbf{k}' , \\
 &= -[A(\alpha) + R_e(1 - \cos\theta)] \mathbf{i}' + R_e \sin\theta \cos\phi \mathbf{j}' + R_e \sin\theta \sin\phi \mathbf{k}' , \\
 D^2 &= [A(\alpha) + R_e(1 - \cos\theta)]^2 + R_e^2 \sin^2\theta , \\
 \omega &= \theta + \tan^{-1} \left[\frac{R_e \sin\theta}{A(\alpha) + R_e(1 - \cos\theta)} \right] , \\
 \cos\eta &= \frac{\mathbf{D} \cdot \mathbf{N}}{|\mathbf{D}|} , \\
 &= \left\{ -a[A(\alpha) + R_e(1 - \cos\theta)] + b R_e \sin\theta \cos\phi + c R_e \sin\theta \sin\phi \right\} \cdot \\
 &\quad \cdot \frac{1}{\sqrt{[A(\alpha) + R_e(1 - \cos\theta)]^2 + (R_e^2 \sin^2\theta)}} , \\
 \theta_m &= \cos^{-1} \left[\frac{R_e}{A(\alpha) + R_e} \right] .
 \end{aligned}$$

Equation 14 may now be written in the form

$$\begin{aligned}
 q_E &= \frac{\sigma T_e^4}{\pi} \int_0^{\cos^{-1} \left[\frac{R_e}{A(\alpha) + R_e} \right]} \int_0^{2\pi} \cos \left\{ \theta + \tan^{-1} \left[\frac{R_e \sin\theta}{A(\alpha) + R_e(1 - \cos\theta)} \right] \right\} \cdot \\
 &\quad \cdot \left\{ a[-A(\alpha) + R_e(1 - \cos\theta)] + b R_e \sin\theta \cos\phi + c R_e \sin\theta \sin\phi \right\} \cdot \\
 &\quad \cdot \left\{ \frac{1}{[A(\alpha) + R_e(1 - \cos\theta)]^2 + R_e^2 \sin^2\theta} \right\}^{3/2} R_e^2 \sin\theta \, d\phi \, d\theta . \quad (15)
 \end{aligned}$$

The limits of integration are such that the part of the earth cap visible from the instantaneous location of the satellite is included. For most practical cases, at least a portion of the cap is not visible from a plate because of its orientation. Attempts to define the appropriate integration limits for such cases are extremely laborious. Since the equations cannot be solved without the aid of a high speed computer, an alternative approach is employed.* The numerical integration includes the entire visible part of the earth cap as

*Katz, A. J., op. cit.

stated, but the contributions of the elemental area that the plate does not see are deleted if the local value of $\cos\eta$ is negative.

Equation 15 represents the flux for the instantaneous orientation of the plate. Interest also lies in the determination of the flux for a complete rotation of the plate about the satellite spin axis. The average value for q_E is therefore

$$q_{E_{av.}} = \frac{1}{2\pi} \int_0^{2\pi} q_E \, d\kappa \quad (16)$$

where κ is the rotational angle.

of the plate with respect to the solar vector:

$$q_{SR} = S (S \cdot N)$$

$$= S [a \cos (\nu - \alpha) \sin \lambda + b \sin (\nu - \alpha) \sin \lambda + c \cos \lambda] , \quad (20)$$

where S is the solar constant.

Because of its orientation the plate may or may not be facing the sun during a rotation. Negative values of $S \cdot N$ indicate that it is not facing the sun.

The average flux per spin is

$$q_{SR_{av.}} = \frac{1}{2\pi} \int_0^{2\pi} q_{SR} d\kappa \quad (21)$$

Because the orientation of the plate is fixed in space, the heat flux is constant throughout the sunlit portion of the orbit. To determine whether or not the satellite is within the earth's shadow at any orbital position requires two simple checks.* If both of the following expressions are satisfied, the satellite received no direct input from the sun:

$$\cos \Omega < 0 ,$$

$$[R_e + A(\alpha)] \cdot |\sin \Omega| < R_e ,$$

where Ω is the angle between the solar vector S and X' ,

$$\cos \Omega = S \cdot i' ,$$

$$\cos \Omega = \cos (\nu - \alpha) \sin \lambda .$$

APPLICATION

Hypothetical geometric configurations (Figure 5) have been chosen to illustrate the use of the

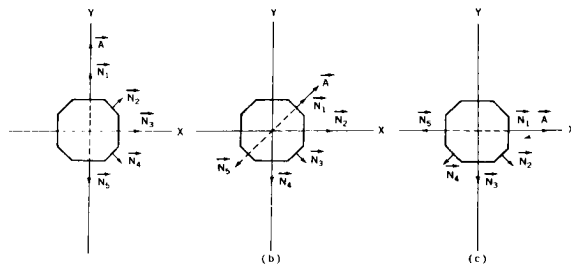


FIGURE 5.—Geometrical configurations.

*Katz, A. J., op. cit.

equations. The external surfaces are represented by the indicated normal vectors. The three sketches represent three spin axis orientations.

Circular orbits of 150, 550, 1050, 2000, and 5000 statute miles, in which the solar vector lies in the plane of the orbit (minimum sunlight), were considered. Figures 6-35 indicate the orbital varia-

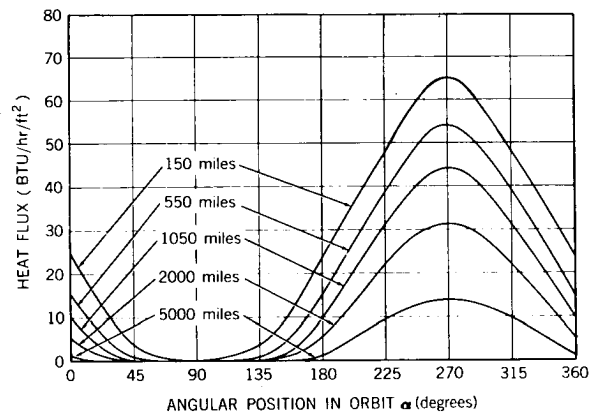


FIGURE 6.—Earth-emitted radiation for circular orbits, $\delta = 90$ degrees, $\beta = 90$ degrees.

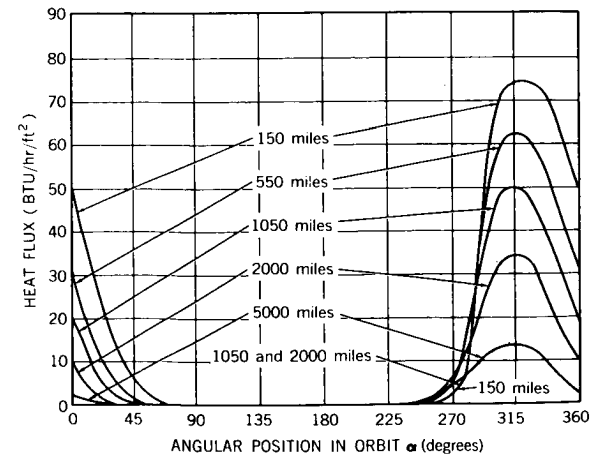


FIGURE 7.—Albedo for minimum sunlit circular orbits, $\delta = 90$ degrees, $\beta = 90$ degrees.

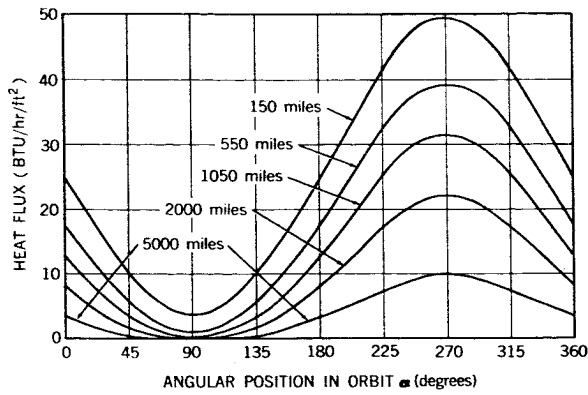


FIGURE 8.—Earth-emitted radiation for circular orbits, $\delta=90$ degrees, $\beta=45$ degrees.

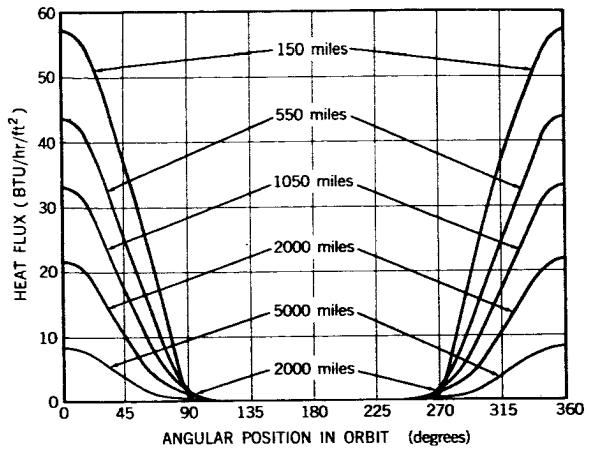


FIGURE 11.—Albedo for minimum sunlit circular orbits, $\delta=90$ degrees, $\beta=0$ degrees.

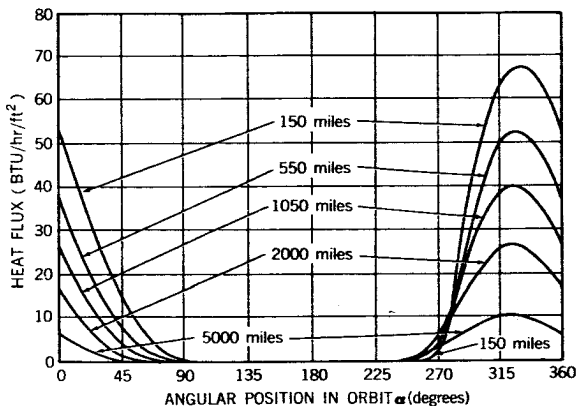


FIGURE 9.—Albedo for minimum sunlit circular orbits, $\delta=90$ degrees, $\beta=45$ degrees.

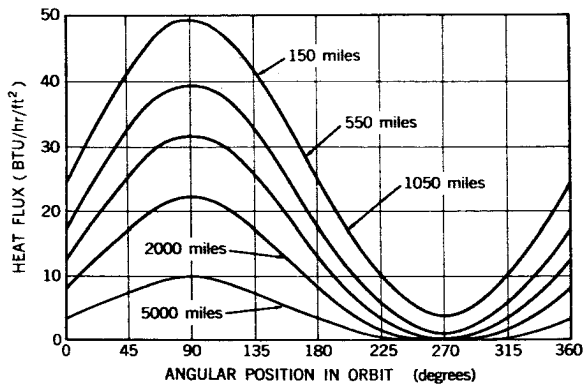


FIGURE 12.—Earth-emitted radiation for circular orbits, $\delta=90$ degrees, $\beta=315$ degrees.

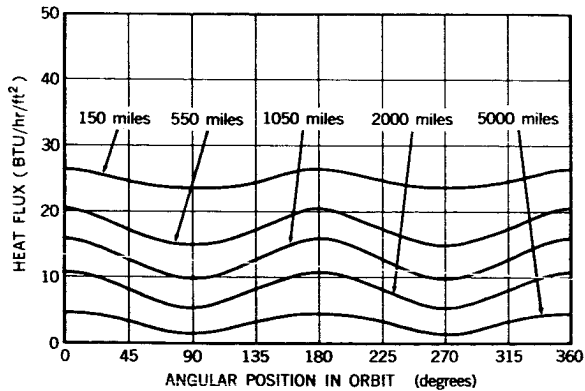


FIGURE 10.—Earth-emitted radiation for circular orbits, $\delta=90$ degrees, $\beta=0$ degrees.

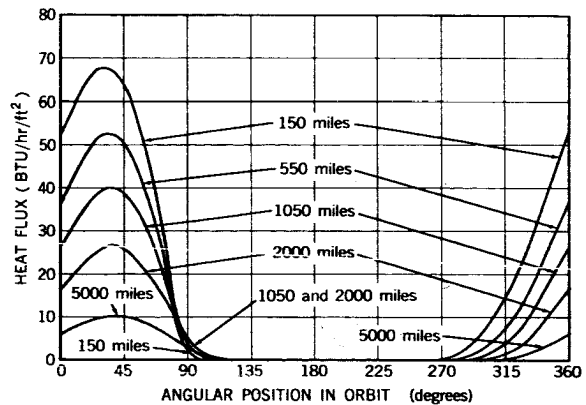


FIGURE 13.—Albedo for minimum sunlit circular orbits, $\delta=90$ degrees, $\beta=315$ degrees.

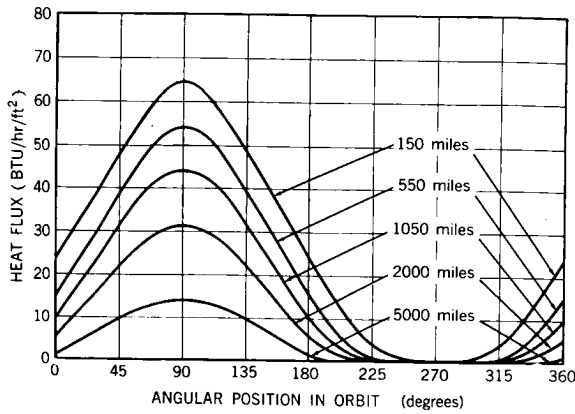


FIGURE 14.—Earth-emitted radiation for circular orbits, $\delta = 90$ degrees, $\beta = 270$ degrees.

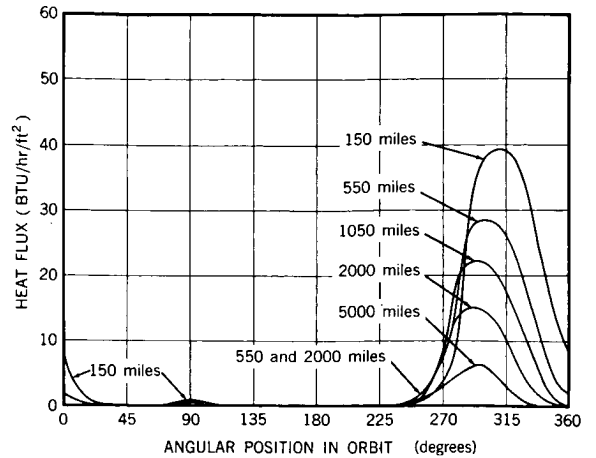


FIGURE 17.—Albedo for minimum sunlit circular orbits, $\delta = 45$ degrees, $\beta = 45$ degrees.

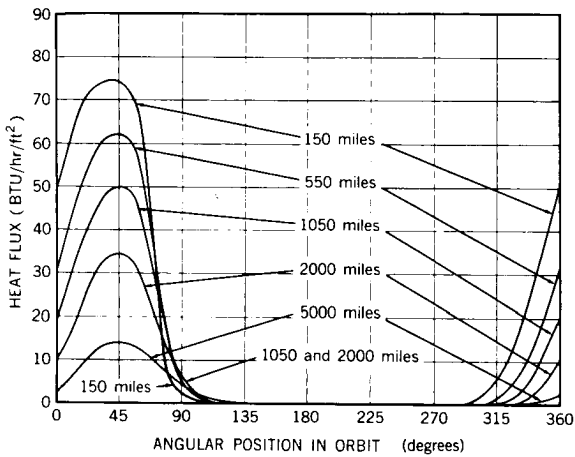


FIGURE 15.—Albedo for minimum sunlit circular orbits, $\delta = 90$ degrees, $\beta = 270$ degrees.

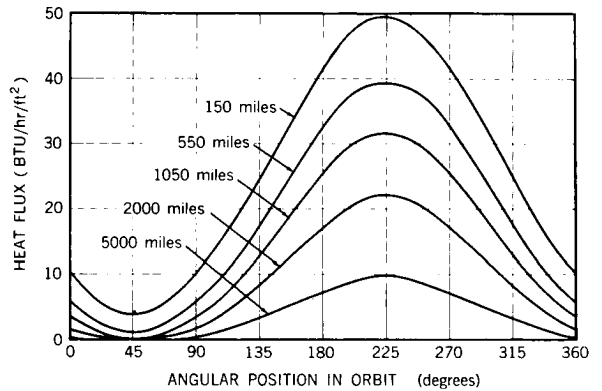


FIGURE 18.—Earth-emitted radiation for circular orbits, $\delta = 45$ degrees, $\beta = 0$ degrees.

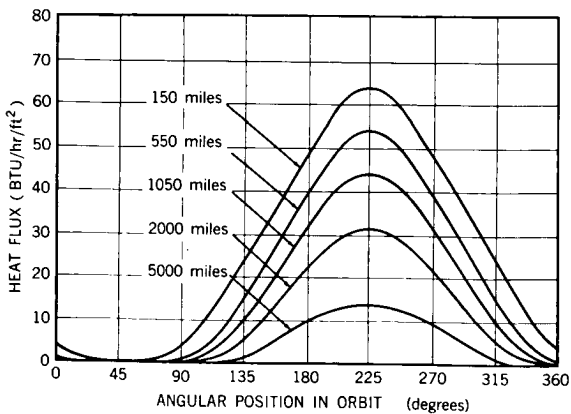


FIGURE 16.—Earth-emitted radiation for circular orbits, $\delta = 45$ degrees, $\beta = 45$ degrees.

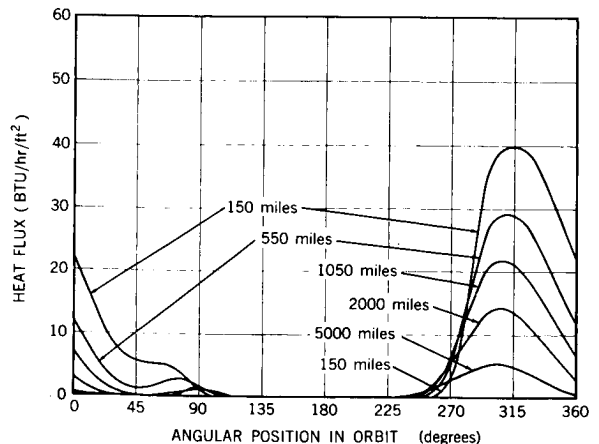


FIGURE 19.—Albedo for minimum sunlit circular orbits, $\delta = 45$ degrees, $\beta = 0$ degrees.

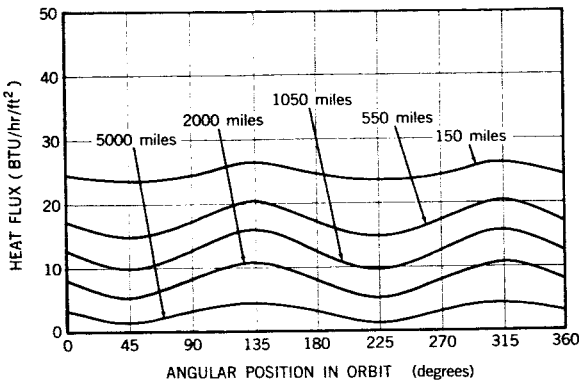


FIGURE 20.—Earth-emitted radiation for circular orbits, $\delta = 45$ degrees, $\beta = 315$ degrees.

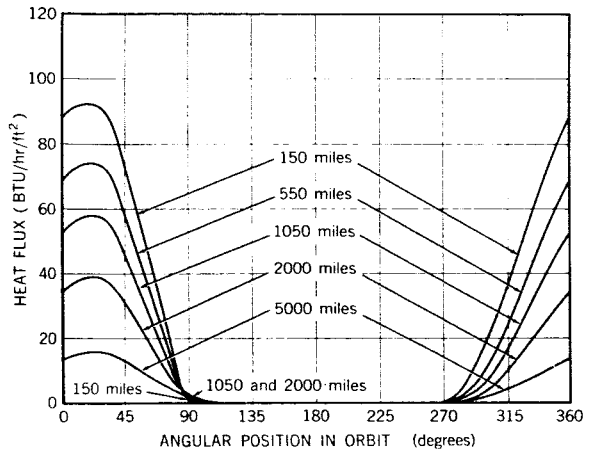


FIGURE 23.—Albedo for minimum sunlit circular orbits, $\delta = 45$ degrees, $\beta = 270$ degrees.

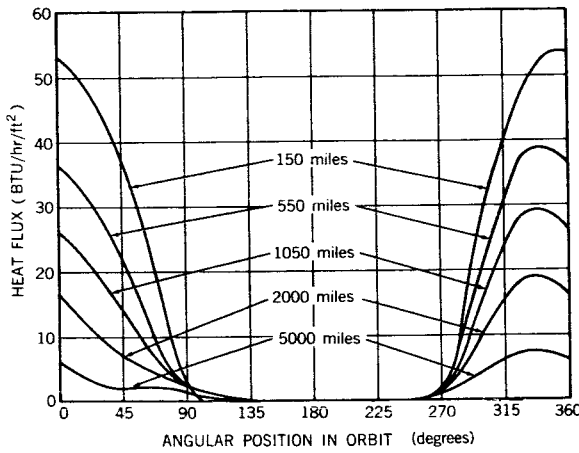


FIGURE 21.—Albedo for minimum sunlit circular orbits, $\delta = 45$ degrees, $\beta = 315$ degrees.

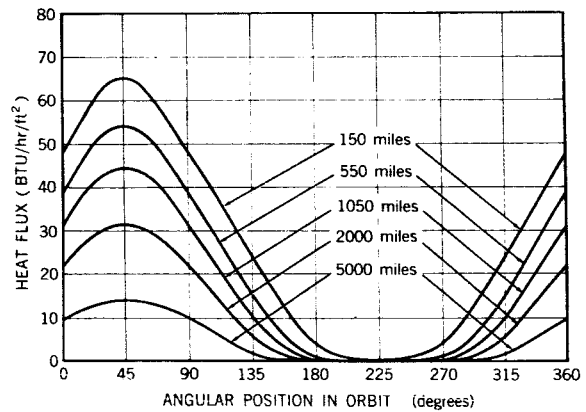


FIGURE 24.—Earth-emitted radiation for circular orbits, $\delta = 45$ degrees, $\beta = 225$ degrees.

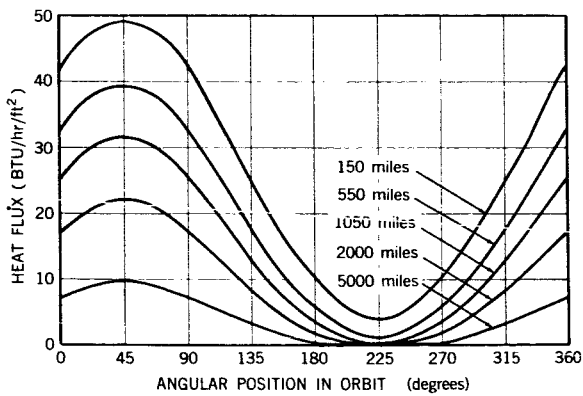


FIGURE 22.—Earth-emitted radiation for circular orbits, $\delta = 45$ degrees, $\beta = 270$ degrees.

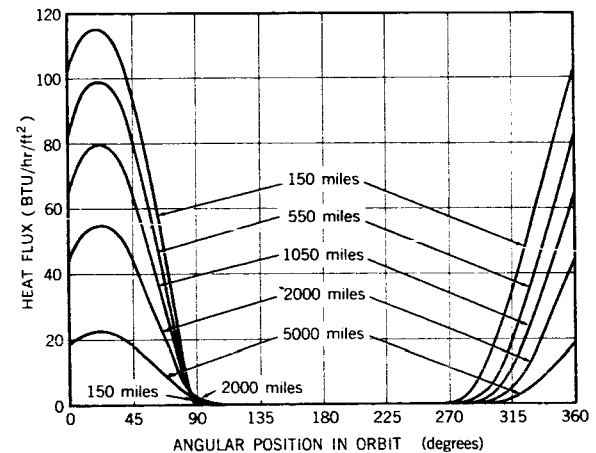


FIGURE 25.—Albedo for minimum sunlit circular orbits, $\delta = 45$ degrees, $\beta = 225$ degrees.

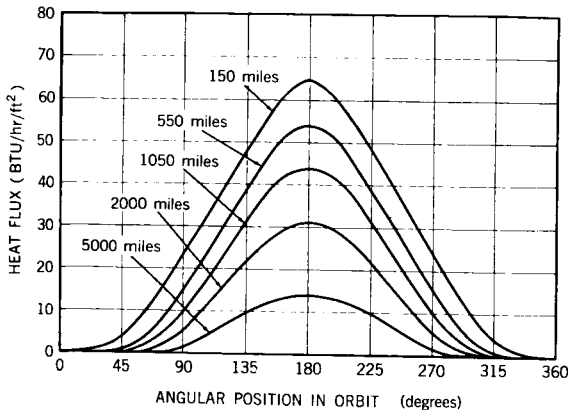


FIGURE 26.—Earth-emitted radiation for circular orbits, $\delta=0$ degrees, $\beta=0$ degrees.

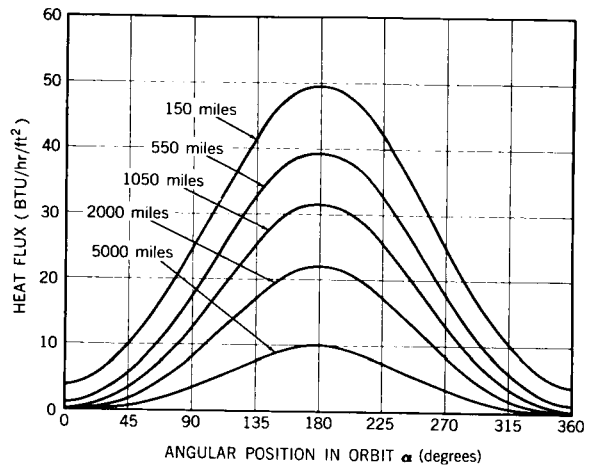


FIGURE 28.—Earth-emitted radiation for circular orbits, $\delta=0$ degrees, $\beta=315$ degrees.

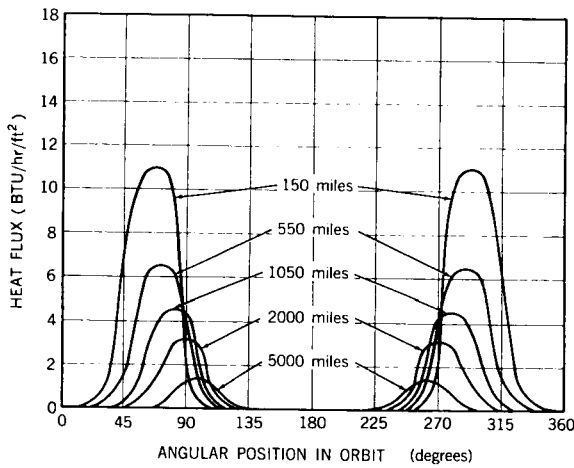


FIGURE 27.—Albedo for minimum sunlit circular orbits, $\delta=0$ degrees, $\beta=0$ degrees.

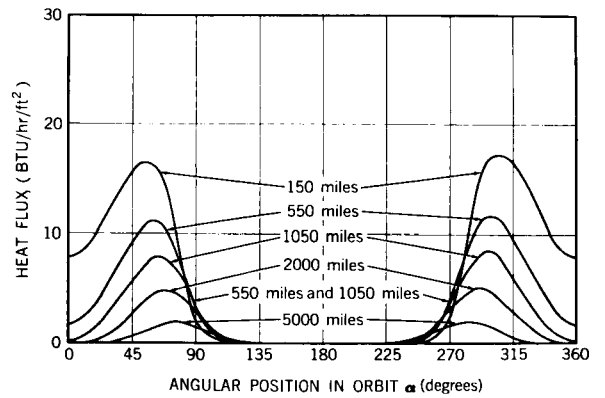


FIGURE 29.—Albedo for minimum sunlit circular orbits, $\delta=0$ degrees, $\beta=315$ degrees.

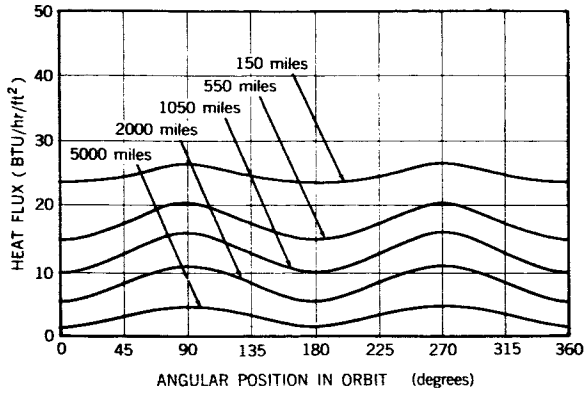


FIGURE 30.—Earth-emitted radiation for circular orbits, $\delta = 0$ degrees, $\beta = 270$ degrees.

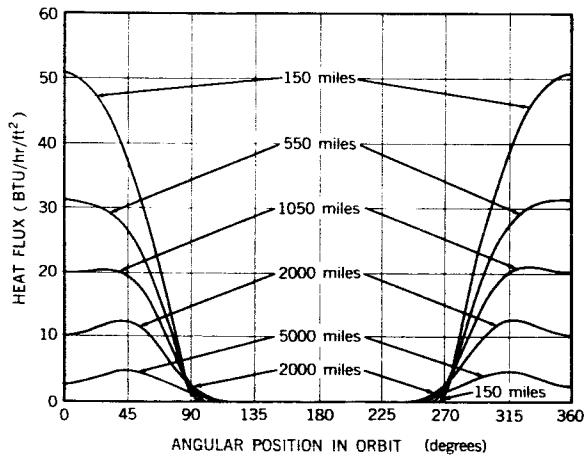


FIGURE 31.—Albedo for minimum sunlit circular orbits, $\delta = 0$ degrees, $\beta = 270$ degrees.

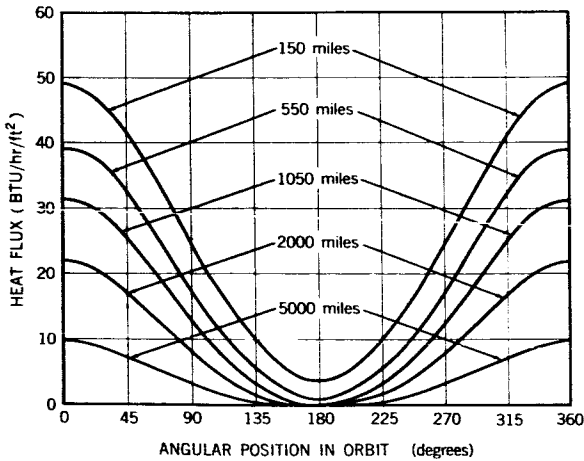


FIGURE 32.—Earth-emitted radiation for circular orbits, $\delta = 0$ degrees, $\beta = 225$ degrees.

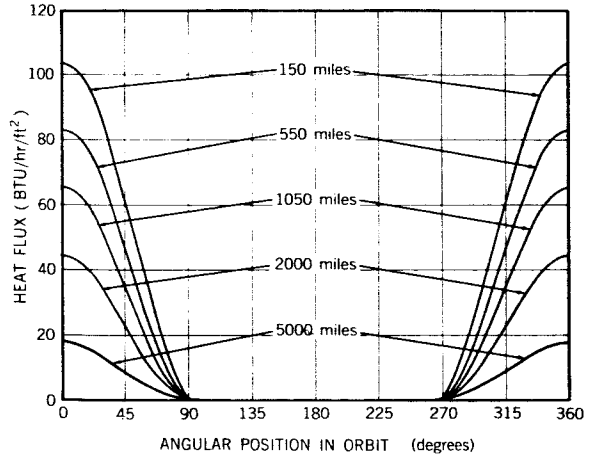


FIGURE 33.—Albedo for minimum sunlit circular orbits, $\delta = 0$ degrees, $\beta = 225$ degrees.

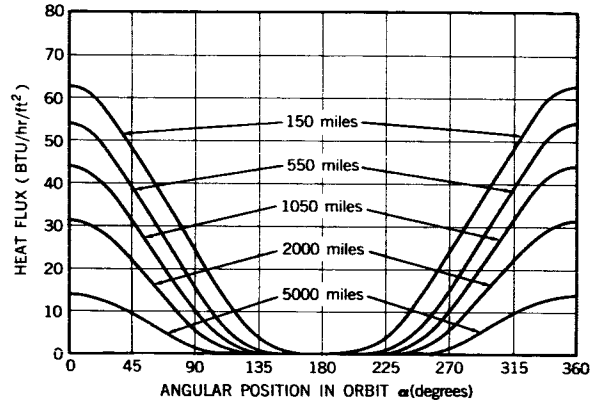


FIGURE 34.—Earth-emitted radiation for circular orbits, $\delta = 0$ degrees, $\beta = 180$ degrees.

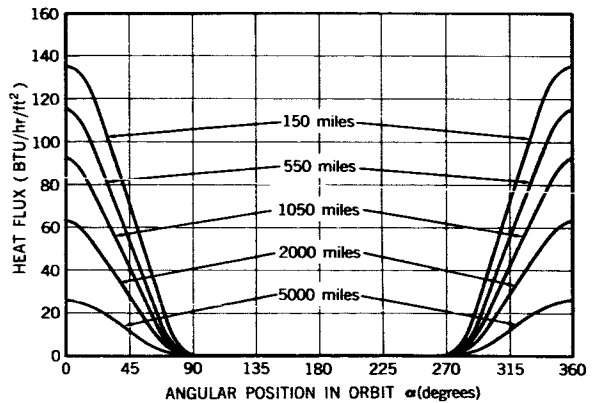


FIGURE 35.—Albedo for minimum sunlit circular orbits, $\delta = 0$ degrees, $\beta = 180$ degrees.

tion of earthshine and albedo.* These heat fluxes represent mean integrated values per rotation at the instantaneous orbital position. The direct solar flux is constant in sunlight for a specific orientation. Appropriate values are presented in Table 1.

ACKNOWLEDGMENTS

The author wishes to express gratitude to Mr. Frank Hutchinson and Mr. Henry Hartley for the IBM 7090 program and the data presentation, respectively.

*The mathematical model assumed the earth cap to be comprised of 512 elemental areas. The curves are based primarily on calculations made at 22.5 degree increments throughout the orbit. Because of this, the peaks in some of the albedo curves were estimated. It should be noted that, because of the configurations chosen, the orbital heat fluxes in several cases are identical except for angular displacement.

Table 1
Mean Solar Heat Flux per Rotation*

δ (degrees)	β (degrees)	$Q_{SR_{av}}$ (BTU/hr/ft ²)
90	90	0.0
90	45	99.0
90	0	140.0
90	315	99.0
90	270	0.0
45	45	311.0
45	0	220.0
45	315	99.0
45	270	0.0
45	225	0.0
0	0	440.0
0	315	311.0
0	270	0.0
0	225	0.0
0	180	0.0

*The data presented represent the instantaneous daylight heat flux impinging upon the rotating faces of the configuration in Figure 5. The sun in all cases is parallel to the X axis; μ , γ , and λ are 90 degrees.

APPLICATION OF EDDY CURRENT CLUTCHES TO TRACKING ANTENNA DRIVE SYSTEMS*

NICHOLAS A. RAUMANN

1. INTRODUCTION

Large microwave tracking antennas often require servo drive systems capable of delivering several hundred horsepower to control the position of the reflector. This is particularly true for antennas that are not equipped with radomes and are required to operate in strong winds. Antenna specifications usually put severe constraints on the characteristics of the servo system. The drive system must have a large dynamic range. It sometimes must be capable of tracking the trajectory of a missile being fired at close range on one hand, and must also be capable of tracking a synchronous satellite on the other. Tracking has to be accomplished with a minimum of error, typically a minute of arc, but in extreme cases the error has to be less than a few seconds of arc. During the tracking operation the servo system has to show reasonable stability and must be relatively insensitive to environmental conditions such as ice, snow, rain, wind, and heat. The drive system must not generate radio frequency interference (RFI) inasmuch as this would interfere with signals received.

Probably 95 percent or more of all tracking antennas in operation today make use of dc motor or hydraulic servo drive systems. This is due to the fact that these systems display most of the desirable characteristics and are also most familiar to system designers. But there are other types of drive systems which may perform as well or exceed the performance of the two popular approaches. The eddy current clutch drive is one. Among others are ac variable frequency drives (Lear Siegler) and "Motron" torque converter drives (W. C. Robinette Co.).

The intent of this report is to study the feasibility of an eddy current clutch drive application to an 85 ft antenna system of the type installed in Fairbanks, Alaska, and Rosman, North Carolina.

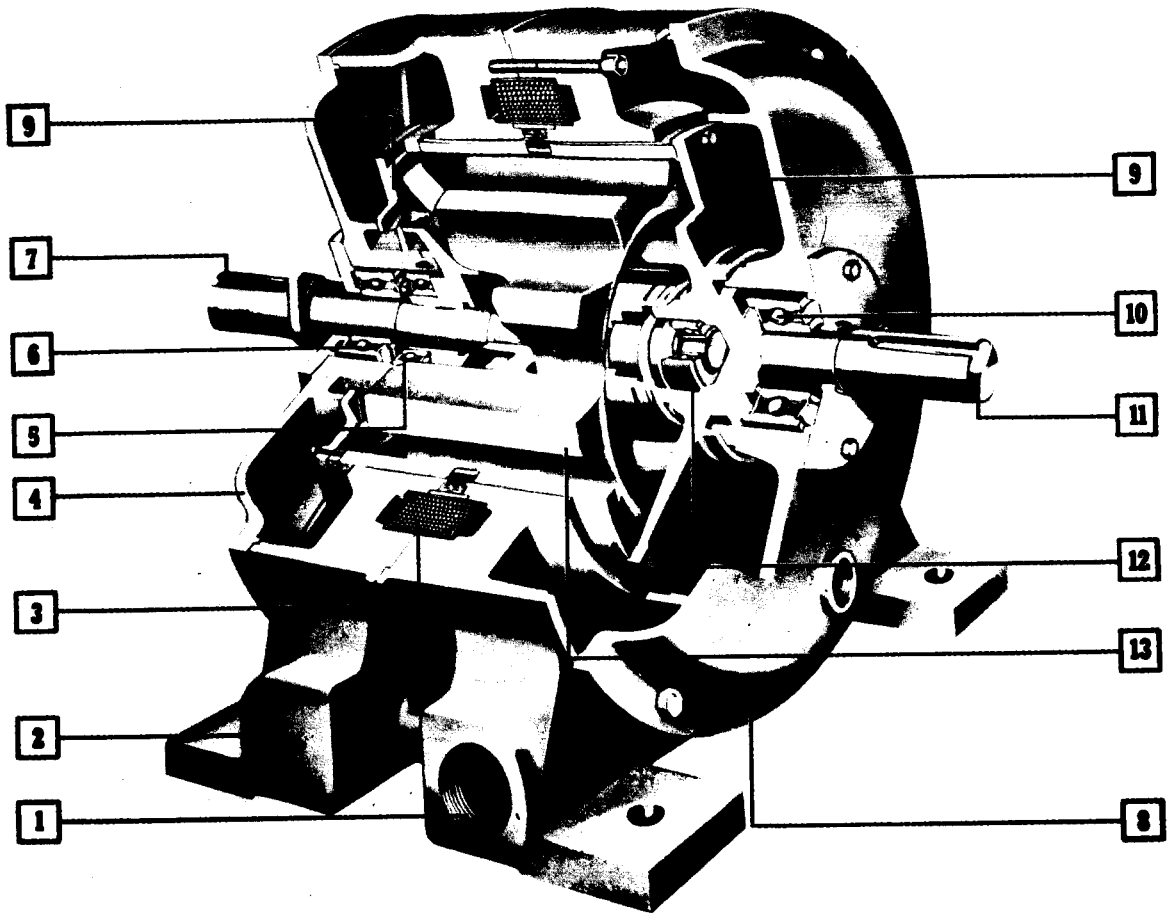
2. CONSTRUCTION AND CHARACTERISTICS OF EDDY CURRENT CLUTCH

Before discussing an eddy current clutch drive system, it may be of interest to discuss the clutch as a component. The clutch can be thought of as an ac squirrel-cage motor which has been modified to permit independent motion of the rotor and stator with respect to each other and with respect to the frame. A dc winding sets up a magnetic field which induces squirrel-cage currents in the rotor when there is relative motion of the rotor with respect to the stator or drum. To obtain relative motion, the drum is driven by a constant speed motor. If the output shaft connected to the rotor is at a different speed, a torque is developed proportional to the flux of the dc field. For discussion purposes a clutch which has been suggested for the drive system of an 85 ft antenna will be used.

2.1 Construction and Magnetic Circuit

Fig. 1 shows a cutaway view of a standard 220 hp eddy current clutch. Basically the coupling consists of three parts: the driven input or drum assembly, the driving output or rotor assembly, and the stationary field assembly. The drum assembly, which is made up of the drum and the input shaft, is supported on two greasable anti-friction bearings, one in the input and the other in the output end bells. In operation, this assembly is the constant speed member driven by a prime mover such as an induction or synchronous motor. The drum is mounted concentrically

*Published as *Goddard Space Flight Center Document X-525-64-203*, July 1964.



stationary parts

- 1** water discharge
- 2** cast frame
- 3** coil
- 4** output end bell
- 8** input end bell

input assembly

Drum [9] and input shaft [11] assembly is carried on two greasable anti-friction bearings, one in input end bracket [10] and other in output end bracket [5].

output assembly

Rotor [13] is mounted on output shaft [7] which is carried on one greasable anti-friction bearing in output end bracket [6] and one pre-lubricated anti-friction pilot bearing [12].

FIGURE 1.—220 hp eddy current clutch (Eaton Manufacturing Co. WCS 216).

between the rotor and field assemblies. The rotor assembly consisting of rotor and output shaft is similarly supported by two greasable anti-friction bearings in the input and output end bells. The

output shaft is usually connected to the load. In this particular design, the rotor is shaped such that twelve pole sections are established with twelve non-magnetic areas between poles. The

field assembly consists of an enclosed toroidal coil and is embedded in the frame of the clutch. A magnetic barrier is placed alongside the coil to prevent a magnetic circuit closure entirely in the field assembly. The advantage of a stationary field is lack of sliprings or commutators which have wear problems and are also a source of RFI.

When the field is excited with dc current, lines of magnetic intensity are set up as shown in Fig. 2. The lines, after leaving one side of the field assembly, pass through one side of the drum assembly to the poles of the rotor and then return through the other side of the drum back to the field assembly. Both sides of the field and drum assemblies are separated by magnetic barriers. Fig. 3 shows

the flux distribution in the clutch. There is a flux concentration in the vicinity of the poles because the flux lines follow a path of minimum reluctance. If relative motion exists between drum and rotor, it can be seen that the flux at any point in the drum changes continuously and thus produces eddy currents in the drum. These currents set up a field which interacts with the field set up by the coil to produce torque.

The interior of the clutch is sealed to permit liquid cooling. Use is made of the constant speed input shaft to circulate the coolant to all critical parts. Normally the coolant is water, but other liquids can be used to accommodate low temperatures.

2.2 Characteristics

The eddy current clutch is a device that can develop torque only in one direction; therefore two clutches are required in systems where load reversal is required. But this consideration is not a disadvantage as far as antenna systems are concerned because opposite or push-pull drives minimize backlash in the gear trains. Push-pull operation also tends to linearize the transfer characteristics of the clutches. Fig. 4 shows the speed-torque curves of a typical clutch. The shape of the curves is very similar to that of ac squirrel-cage motors (NEMA type C or type D operation). Above slip speeds of 600 rpm the curves are virtually flat indicating a lack of damping. Since this is the primary region of operation for antenna application, the eddy current clutch can be considered to be a double integrating device. Dc motors, because of back emf, and hydraulic motors, because of back pressure, can be considered to be single integrators. This makes

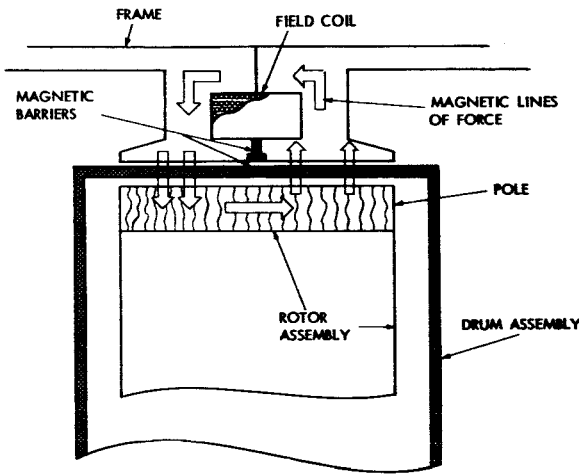


FIGURE 2.—Lines of magnetic intensity.

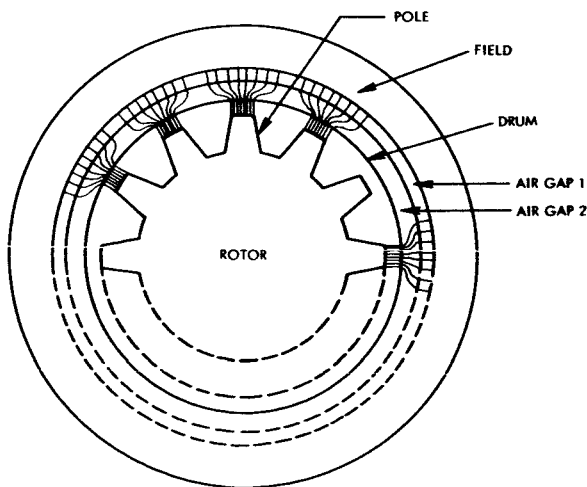


FIGURE 3.—Flux distribution in clutch.

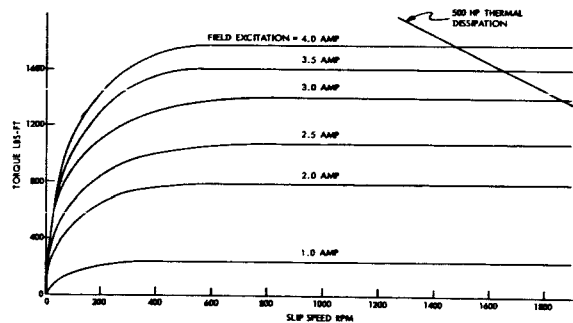


FIGURE 4.—Speed-torque characteristic.

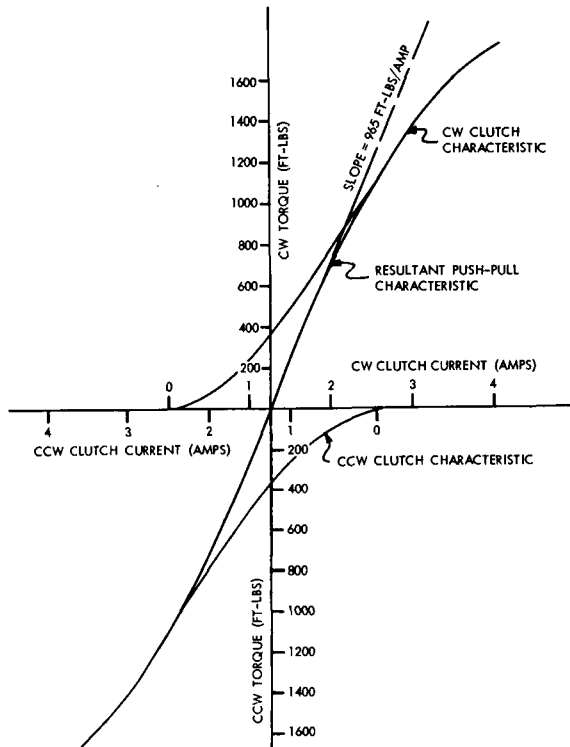


FIGURE 5.—Locked-rotor current-torque characteristic.

dc and hydraulic motors velocity controllers, whereas clutches would be acceleration controllers.

The speed-torque curves (Fig. 4) can be transposed into a current-torque curve (Fig. 5). This curve shows an approximation to a square-law characteristic for the individual clutch, but shows a linear characteristic (over a large band) for the push-pull combination of clutches. The square-law characteristic is exhibited because both dc flux and induced eddy currents increase with an increase in control current. The push-pull linearization effect can also be shown analytically; if

T = torque—ft lbs

I = current in eddy current clutch field—amps

I_0 = bias current in eddy current clutch field—amps (for Fig. 5, I_0 corresponds to the current that produces 20 percent rated torque)

K_i' = torque constant—ft lbs/amp/amp bias

then clutches 1 and 2 develop torques:

$$T_1 = K_i'(I_0 + I)_2 \quad (1)$$

$$T_2 = K_i'(I_0 - I)_2 \quad (2)$$

and push-pull operation will yield the combined torque T :

$$T = T_1 - T_2 = 4K_i'I_0I = K_iI \quad (3)$$

The combined torque T , therefore, is a linear function of field current I for constant torque K_i' and bias current I_0 .

Since two fields are set up in the eddy current clutch, two distinct time constants could be expected: a field time constant, and an eddy current time constant. The inertial time constant prominent in dc and hydraulic motors is absent. Test results on clutches indicate that the eddy current time constant is very small and that the field time constant is not a simple time constant in the sense of representing a corner frequency on the Bode plot. Analytical determination of the field time constant is very difficult because clutch geometry enters into the calculation and requires the solution of complicated field problems. However, if simplifying assumptions are made, i.e., assuming simple geometry and ignoring leakage currents and ampere-turns required to set up flux in the steel, the time constant can be computed as follows:

The inductance L of the field can be determined from

$$L = \frac{N\phi}{I} \text{ henries} \quad (4)$$

where

N = number of field turns

ϕ = flux—webers

I = field current—amps

The flux ϕ can be determined by

$$\phi = \frac{F}{R} \text{ webers} \quad (5)$$

where

$F = 4\pi NI$ = magnetomotive force (mmf)
—pragilberts

R = total reluctance—pragilberts/weber

The total reluctance R of a simple series magnetic circuit is simply the sum of the reluctance of the

steel R_s and the reluctance of the airgap R_a . Assuming that leakage flux can be neglected, the flux in the steel ϕ_s has to equal the flux in the airgap ϕ_a or

$$\phi_s = \phi_a \quad (6)$$

But since the flux in the airgap is

$$\phi_a = \frac{F}{R_a} \quad (7)$$

equations 5, 6, and 7 can be combined to yield

$$\phi = \frac{4\pi NI}{R_a} \quad (8)$$

The reluctance of the airgap R_a is

$$R_a = \frac{\ell_a}{\mu_a A_a} \text{ pragilberts/weber} \quad (9)$$

where

ℓ_a = length of the airgap—meters

μ_a = permeability of the air = 10^{-7} weber/pragilbert—meters

A_a = area of the pole surface—meters²

Combining equations 4, 8, and 9:

$$L = \frac{4\pi\mu_a N^2 A_a}{\ell_a} \text{ henries} \quad (10)$$

The resistance of the field coil can be either measured directly or computed by

$$R = 2\pi r \rho N \quad (11)$$

where

r = mean radius of coil—feet

ρ = resistivity of wire—ohms/ft

In using equations 10 and 11 the time constant of the field T_f can be found:

$$T_f = \frac{L}{R} = \frac{2\mu_a N A_a}{\ell_a r \rho} \text{ sec} \quad (12)$$

For more detailed analysis of the transient behavior of clutches the reader is referred to Reference 2. With the L and the R of the field coil found the transfer function can be determined for a simple inertial load. The system has two

basic equations (in linear range of operation):

$$E = (sL + R)I \quad (13)$$

$$T = K_t I = s^2 J \theta \quad (14)$$

where

E = applied voltage to field—volts

T = torque at output shaft—ft/lbs

J = sum of motor and load inertia—slug ft²

θ = output shaft position—radians

s = Laplace operator

Combining equations 12, 13, and 14 gives the transfer function of the eddy current clutch:

$$\frac{\theta}{E} = \frac{K_t / JR}{s^2 (sT_f + 1)} \quad (15)$$

Another very important characteristic of the coupling is its ability to dissipate heat. Fig. 6 shows the variation of the different power quantities as a function of clutch speed for maximum excitation of the field. The different power quantities can be determined as follows:

$$\text{Input horsepower} = \frac{\text{input speed} \times \text{torque}}{5250} \quad (16)$$

$$\text{Slip horsepower} = \frac{\text{slip speed} \times \text{torque}}{5250} \quad (17)$$

$$\text{Output horsepower} = \frac{\text{output speed} \times \text{torque}}{5250} \quad (18)$$

where: torque = input torque = output torque

slip speed—input speed—output speed

The above expressions are not exact, inasmuch as certain losses have been neglected. These additional losses break down as follows:

Coolant drag loss.....	5%
—with overhauling load.....	10%
Magnetic drag.....	1%
Coolant acceleration loss.....	.75%
Windage and friction.....	.75%
Field excitation loss.....	.06%
—with 5:1 forcing.....	.4%

It can be seen from Fig. 6 that the heat generated in the clutch is appreciable, especially at low

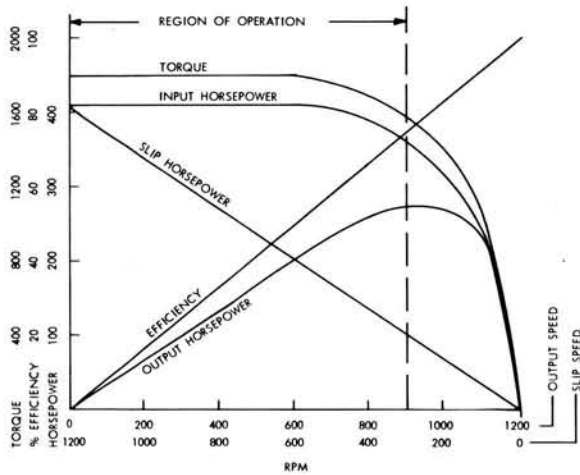


FIGURE 6.—Horsepower performance characteristic.

speeds, and care has to be taken that sufficient coolant is supplied to remove the heat from the critical portions of the clutch. This also implies that, if a pair of opposing clutches is at standstill with no load torques applied and a 20 percent bias is used, $2 \times 0.20 \times 410 = 164$ hp is generated in heat continuously. This is a fair amount of heat to be removed by the coolant and warrants careful economic and reliability type considerations. This heating condition in the clutch represents a major disadvantage of an eddy current clutch system.

The efficiency of the clutch is determined by:

$$\begin{aligned} \text{Efficiency} &= 100 \frac{\text{output horsepower}}{\text{input horsepower}} \\ &= 100 \frac{\text{output speed}}{\text{input speed}} \quad (19) \end{aligned}$$

Consequently, low efficiencies are experienced at low clutch output speeds, a condition most commonly occurring in antenna applications.

3. EDDY CURRENT CLUTCH SIMULATOR

Because the eddy current clutch drive system is a rather unfamiliar device, a scaled-down simulator was built rather than a full-size drive for an 85 ft acquisition antenna. The simulator allows the evaluation of system and component characteristics in the laboratory and shows trends and design requirements for the full-size version. The simulator was built for NASA by Westinghouse

Electric Corporation under contract NAS5-2916 (see Reference 1). The simulator is shown in Figs. 7 and 8. Basically, the simulator consists of a pair of eddy current clutches with drive motors, a gear box, a load, a controller, and a pair of power amplifiers. When torque requirements were scaled down by a factor of 2000 and speeds were scaled up by 65, a reasonably sized simulator was obtained. Table I lists important characteristics of an 85 ft antenna and its corresponding simulator quantities.

3.1 Simulator Clutches

WCS 2062 clutches were selected for the simulator. These are standard 10 hp watercooled units and are driven by 10 hp, 1160 rpm squirrel-cage motors. The clutches are manufactured by

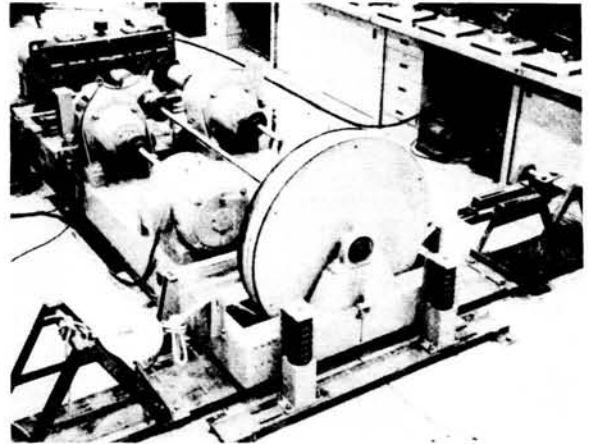


FIGURE 7.—Eddy current clutch simulator (flywheel end)

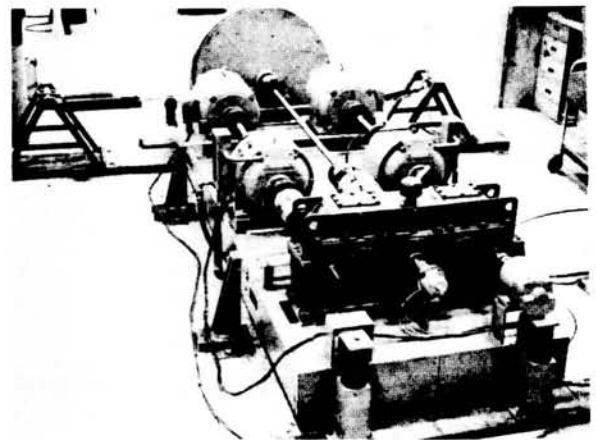


FIGURE 8.—Eddy current clutch simulator (gearbox end).

TABLE I.—System Characteristics

Quantity	Symbol	Units	85 ft Antenna	Simulator
Load inertia.....	J_L	slug ft ²	7.1×10^6	55
Wind torque.....	T_w	ft lbs	1.0×10^6	496
Load velocity.....	$\dot{\theta}_L$	rpm	.5	33.5
Load acceleration.....	$\ddot{\theta}_L$	rad/sec ²	.0872	5.66
Accuracy.....		min	.5	32.5
Bandwidth.....		cps	.7-1	.7-1
Nat. frequency.....	f_n	cps	1.5	1.67
(antenna reflector)				
Nat. frequency.....		cps	5.0	5.55
(antenna tower)				

Eaton Manufacturing Company. To study the performance of these clutches, a series of characteristic curves were obtained.

Current—Torque Characteristic (Fig. 9): The curve shows how torque varies with field current under locked-rotor conditions. For low excitations the square-law characteristic is exhibited and at high excitations the effect of saturation becomes noticeable. The straight line portion of the curve has a slope of 12.5 ft lbs/amp.

Speed—Torque Characteristic (Fig. 10): The curve shows variation of torque with respect to speed for various excitation conditions. Especially of interest is the slope of the curves because

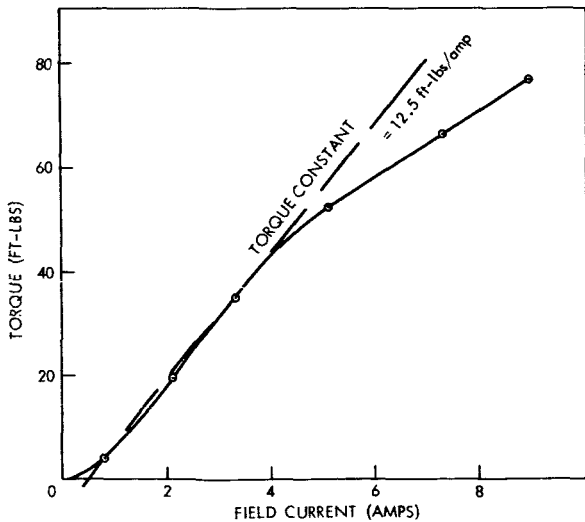


FIGURE 9.—Locked-rotor current-torque characteristic (simulator).

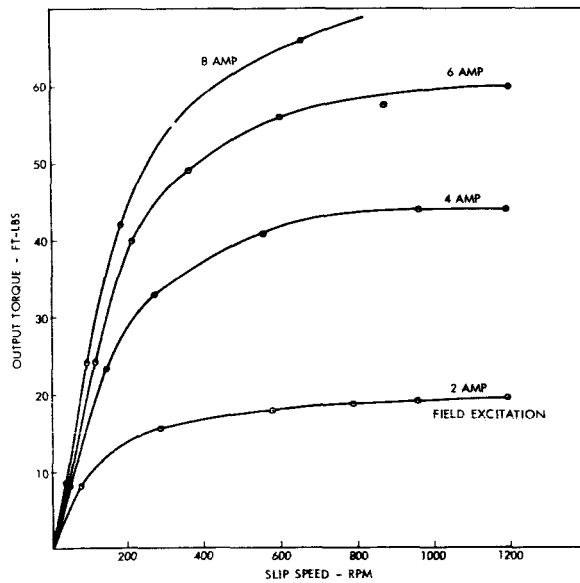


FIGURE 10.—Speed-torque characteristic (simulator).

the slope is a measure of the inherent damping of the clutch. For a given excitation very little damping is obtained at high slip speeds, and high damping is obtained at low slip speeds. Thus the amount of damping is a function of speed. But, if operation is confined to the high slip speed end of the curve—say between 600 and 1200 rpm, virtually constant damping is encountered. If different excitation levels are now considered, a dependency between damping and excitation level is found with damping being proportional to excitation.

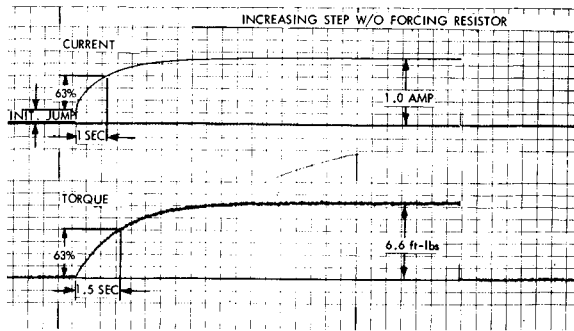


FIGURE 11a.—Transient response-increasing step w/o forcing resistor.

Torque—Time Response (Fig. 11): Fig. 11a shows the current and torque response to an increasing step in voltage. The current jumps to about 21 percent of its final value and then continues to rise exponentially with an approximate time constant (63 percent of final value) of 1.0 sec. The torque does not have an initial jump but rises exponentially with a time constant of 1.5 sec. The linear relationship between torque and current observed in ac and dc machines therefore seems to be missing in eddy current clutches. To reduce the torque time constant, an external forcing resistor can be used. In the case of the simulator 10:1 forcing was utilized; i.e., in series with the 2.3 ohm field of the clutch a resistor of 23 ohms was added. The time response for this condition is shown in Fig. 11b. Now the current jumps to 75 percent of its final value and continues to rise with a time constant of 0.3 sec. The torque response displays a time constant of 0.5 sec. Consequently, field forcing of 10:1 has improved the torque time constant by 3:1 only. Fig. 11c shows the response for a decreasing step in voltage. The current characteristic is very similar to that of the preceding case, but the torque decays faster with a time constant of 0.3 sec. Other tests performed indicated that the time constant does not seem to be affected much by the size of the voltage step applied, and also not by the amount of bias current used.

Torque—Frequency Response (Fig. 12); The frequency response shows approximately the characteristics of a single time constant device with a corner frequency at 5 rad/sec (time constant = .2 sec); however, closer investigation reveals that a rather poor correlation between the amplitude

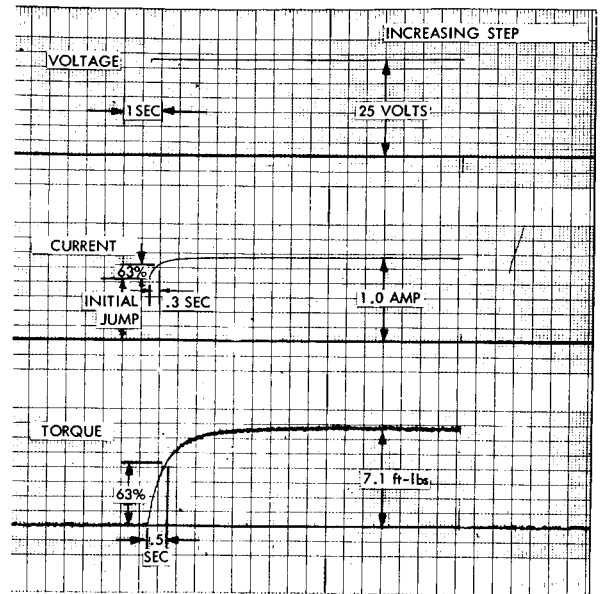


FIGURE 11b.—Transient response-increasing step.

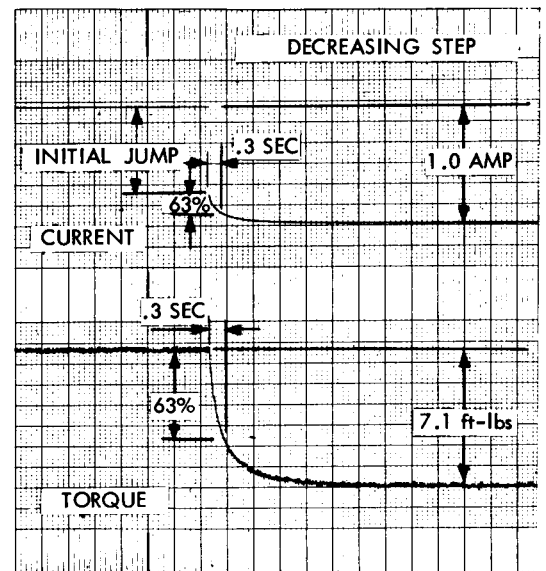


FIGURE 11c.—Transient response decreasing step.

response and the -20 db/decade asymptote is obtained. The corresponding phase shift is larger than that of a single time constant device and shows marked similarity to a non-minimum phase network.

Based on these and preceding observations it can be concluded that the eddy current clutch has several inherent non-linearities which would

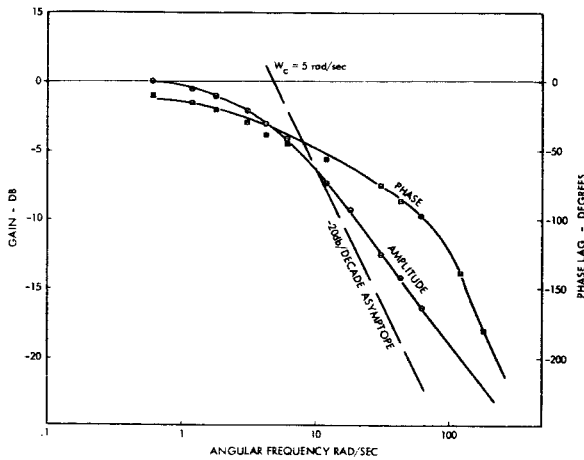


FIGURE 12.—Frequency response of clutch.

rule out linear system analysis like Bode plot design, Root locus method, tec. However, for a first approximation in system design a linearized clutch transfer function of the form given in equation 15 could be used. It is difficult to say whether the non-linearities observed in the clutch tested will become worse in larger sized units and, if so, to what degree.

3.2 Mechanical Load

The load of the simulator consists of a flywheel which represents the scaled down inertia of the reflector of an antenna and a flexible shaft which, together with the flywheel, simulates the natural frequency of the reflector. A block diagram and a simplified equivalent circuit of the mechanical system are shown in Fig. 13. From the equivalent circuit two basic equations can be written:

$$nT = s^2 nJm\theta m + K_s \left(\frac{\theta m}{n} - \theta_L \right) \quad (20)$$

$$0 = K_s \left(\theta_L - \frac{\theta m}{n} \right) + (s^2 J_L + sB_L)\theta_L \quad (21)$$

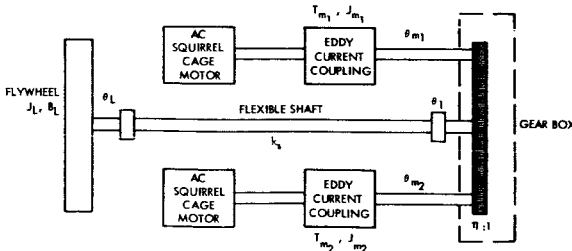
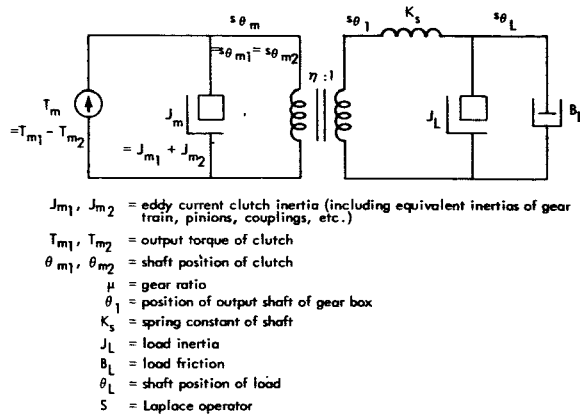


FIGURE 13a.—Block diagram of mechanical arrangement of simulator.



- J_{m1}, J_{m2} = eddy current clutch inertia (including equivalent inertias of gear train, pinions, couplings, etc.)
- T_{m1}, T_{m2} = output torque of clutch
- θ_{m1}, θ_{m2} = shaft position of clutch
- μ = gear ratio
- θ_1 = position of output shaft of gear box
- K_s = spring constant of shaft
- J_L = load inertia
- B_L = load friction
- θ_L = shaft position of load
- S = Laplace operator

FIGURE 13b.—Equivalent mechanical circuit.

Solving these equations simultaneously gives the transfer function of the mechanical system:

$$\frac{\theta_L}{T} = \frac{n/B_L}{s \left[\frac{n^2}{K_s B_L} J_m J_L s^3 + \frac{n^2 J_m s^2}{K_s} + (n^2 J_m + J_L) \frac{1}{B_L} s + 1 \right]} \quad (22)$$

Since the friction in the system is small, the assumption can be made that the friction term B_L approaches zero, which simplifies the transfer function to

$$\frac{\theta_L}{T} = \frac{n}{s^2 \left(\frac{n^2 J_m + J_L}{K_s} s^2 + 1 \right)} = \frac{K_t}{s^2 \left(\frac{s^2}{\omega_n^2} + 1 \right)} \quad (23)$$

where

$$k_t = \text{torque constant} = \frac{n}{n^2 J_m + J_L}$$

$$\omega_n = \text{natural frequency} = \sqrt{\frac{K_s (n^2 J_m + J_L)}{n^2 J_m J_L}}$$

Substituting values for the system constants: $J_m = .329$ slug ft², $J_L = 55$ slug ft², $K_s = 5060$ ft lbs/rad, and $n = 20$, then

$$k_t = .107 \text{ rad/ft lbs}$$

$$\omega_n = 11.4 \text{ rad/sec} = 1.82 \text{ cps}$$

$$\frac{\theta_L}{T_m} = \frac{.107}{s^2 (.0077 s^2 + 1)} \quad (24)$$

Fig. 14a shows the response of the mechanical system to an initial torque input. The natural

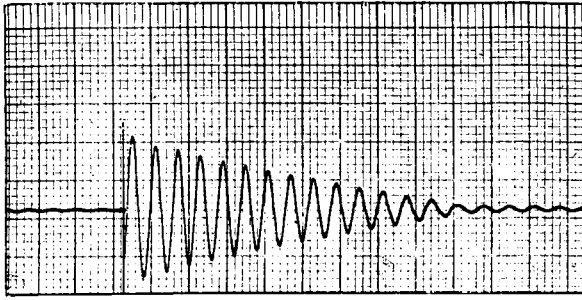


FIGURE 14a.—Natural frequency of simulated reflector.

damped frequency is 1.67 cps, therefore correlating with the value computed above. The response has an exponential decay that corresponds to a damping factor of 0.022. This makes valid the previous assumption that the friction term B_L can be neglected.

To simulate the natural frequency of the tower, mechanical components have been mounted on a spring supported bedplate. The bedplate mass-spring combination gives rise to the natural frequency of the antenna tower structure. Fig. 14b indicates a natural frequency for the mass-spring combination of 5.55 cps with a damping factor of about .007. By properly instrumenting the simulator, the effects of reflector and tower frequencies are summed into the overall system servo loop.

3.3 Controller and Power Amplifier

The controller for the simulator uses operational amplifiers as its basic building blocks. The transfer functions of these amplifiers can be readily shaped so as to give desired open and closed servo loop characteristics consistent with stability and

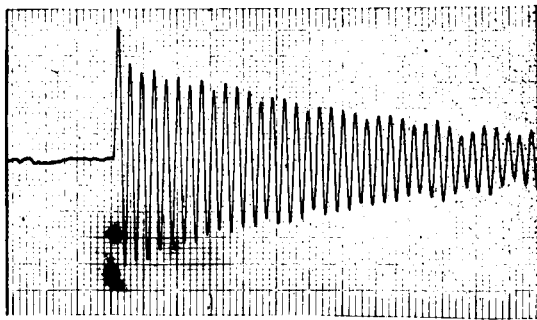


FIGURE 14b.—Natural frequency of simulated antenna tower.

accuracy requirements. The controller is capable of demodulating the synchro position error signal and of accepting velocity and acceleration feedback signals. The controller superimposes a dither signal on its output which decreases system coulomb friction, and therefore reduces breakaway torque. The output of the controller drives the power amplifiers.

The power amplifiers, one for each clutch, are capable not only of exciting the field of the clutches but also their forcing resistors. The power amplifiers are 300 volt, 3 phase silicon-controlled rectifier (SCR) ac to dc converters. From an operational requirement, SCR's are hardly desirable because they generate radio frequency interference (RFI) due to their steep wave fronts and have an inherent delay (transport lag) which is a considerable fraction of a period of the power supply frequency. Replacement of these amplifiers with "true" dc amplifiers should be considered for similar applications.

3.4 Simulator Performance

From the block diagram of the servo system, as shown in Fig. 15, it should be observed that three distinct loops are used for ultimate system stability and accuracy. Because the eddy current clutch has very little inherent damping, the tachometer loop is used mainly to introduce controllable damping. Fig. 16a shows the closed loop frequency response for this loop and indicates a 3 db bandwidth of 35 rad/sec. The somewhat poor correlation between the measured and computed curves is most likely due to the non-linear behavior of the clutch itself. The primary natural frequency of the load is about 10 rad/sec. Since an overall bandwidth of about 6 rad/sec is

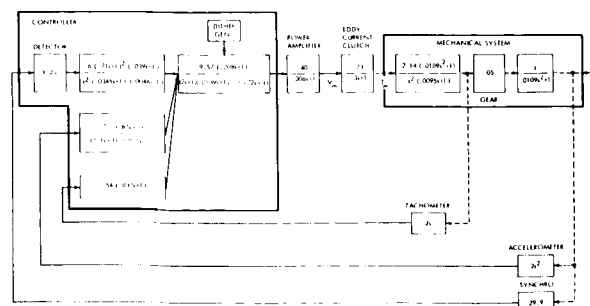


FIGURE 15.—Block diagram of control loop.

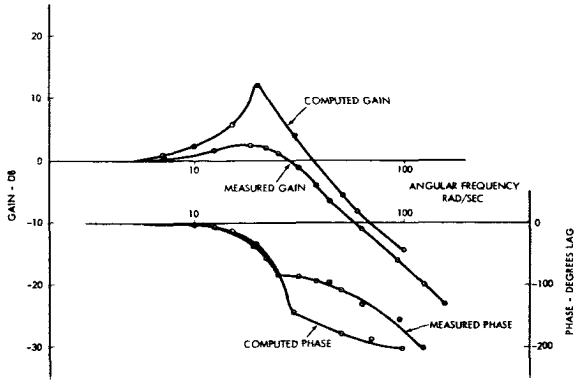


FIGURE 16a.—Bode plot for closed tachometer loop.

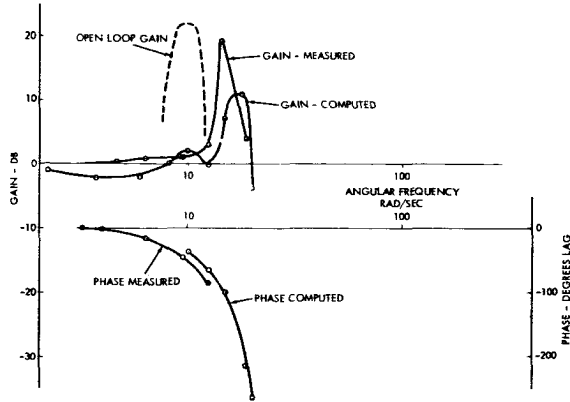


FIGURE 16b.—Bode plot for closed accelerometer loop.

required, stability problems are encountered if the position loop were to be closed. The stability problem arises because the natural frequency is very close to the desired cut-off frequency of the position loop and because the load is highly underdamped. It is for this reason that the accelerometer loop is used. As is evident from the Bode plot for this loop (Fig. 16b), the underdamped mode has been shifted out to about 15 rad/sec and damping has been effectively increased.

Open loop gain and bandwidth of the position loop depend on the type of target to be tracked and the allowable tracking accuracy. If one considers a satellite tracking antenna, then the number of classes of possible input functions to the servo system is rather limited to some well-defined continuous functions. For example, functions with discontinuous steps can be ruled out. One can use a simple graphical curve fitting method as outlined in Reference 5 (pp. 61-65) to obtain

equivalent sinusoidal inputs for velocity, acceleration, and higher derivatives. Table II gives the results of this method for a satellite in a 100-mile circular orbit with an orbital velocity of 4.9 miles/second and a permissible tracking error ϵ of 10 seconds of arc. The choice of the servomechanism type (refer to Reference 4, pp. 2-35 to 2-36) to be used in the position loop depends mainly upon the open loop gain required. For the same conditions as above the open loop gains for different type systems are:

Type 1:

$$K_v = \frac{\dot{\theta}}{\epsilon} = \frac{2.8 \times 3600}{10} = 1010 = 40.1 \text{ db} \quad (25)$$

Type 2:

$$K_a = \frac{\ddot{\theta}}{\epsilon} = \frac{.1 \times 3600}{10} = 36 = 30.5 \text{ db} \quad (26)$$

Type 3:

$$K_j = \frac{\ddot{\ddot{\theta}}}{\epsilon} = \frac{.014 \times 3600}{10} = 5 = 13.9 \text{ db} \quad (27)$$

TABLE II.—Equivalent Sinusoidal Inputs

Quantity	Symbol	Max. Value Given for Antenna (for Simulator $\times 65$)	Equivalent	
			Gain -db	Frequency (rad/sec)
Position.....	θ	$\pm 70^\circ$	88.51	.0256
Velocity.....	$\dot{\theta}$	$2.8^\circ/\text{sec}$	81.55	.0426
Acceleration.....	$\ddot{\theta}$	$.1^\circ/\text{sec}^2$	62.68	.154
Jerk.....	$\ddot{\ddot{\theta}}$	$.014^\circ/\text{sec}^3$	40.20	.307

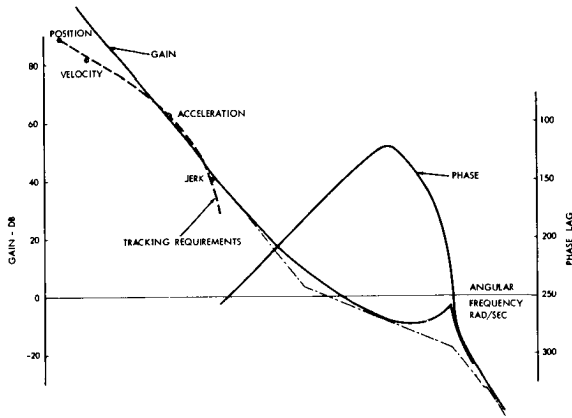


FIGURE 17a.—Bode plot for open position loop.

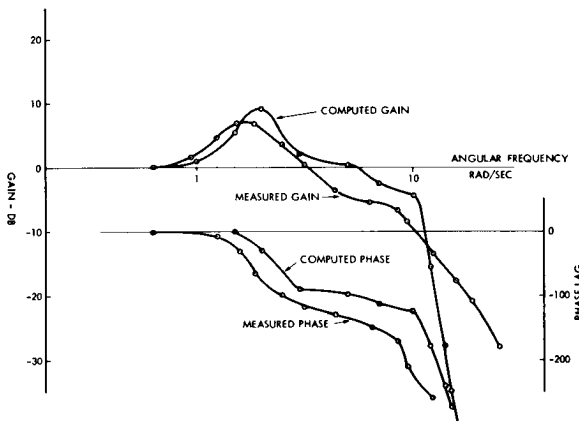


FIGURE 17b.—Bode plot for closed position loop.

This high gain Type 1 system is undesirable from the standpoint of noise problems, and the lower gain Type 3 system has problems in stabilization. Nevertheless, the latter system was chosen because it represented the best compromise. Fig. 17a shows the open and Fig. 17b the closed loop Bode plots for the position loop. The dotted line in Fig. 17a is the graphical representation of results from Table II. This line defines the open loop gains which have to be realized in order to track with the required accuracy. A closed loop 3 db bandwidth of 4 rad/sec was obtained, which is just below the specification requirements of 4.4 rad/sec, but minor adjustments in the loop could remedy this situation.

Fig. 18a, b, and c show the step, impulse, and sine wave responses, respectively, for the position loop. The step response has a 60 percent over-

shoot to a 2.3° step or a peak error of 1.4° . Even though the error exceeds the permissible limit, it is of very little consequence since the step function is an input that would not be experienced normally when tracking a satellite. What is of

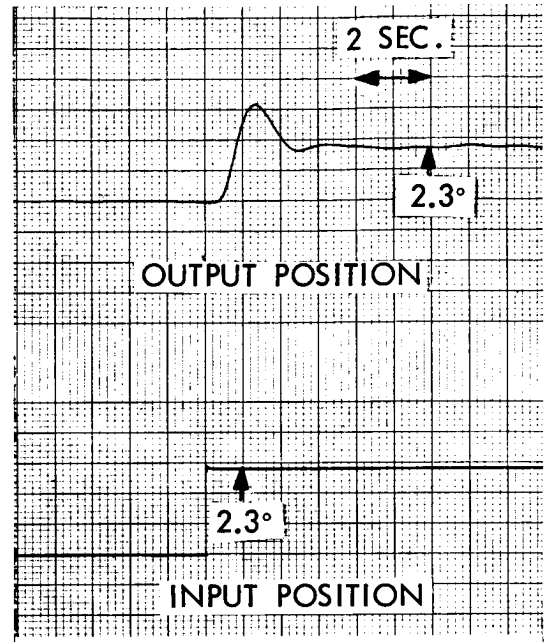


FIGURE 18a.—Step response of simulator.

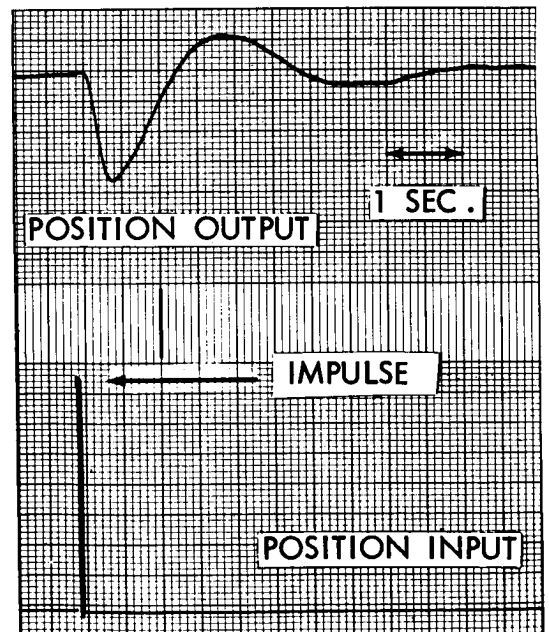


FIGURE 18b.—Impulse response of simulator.

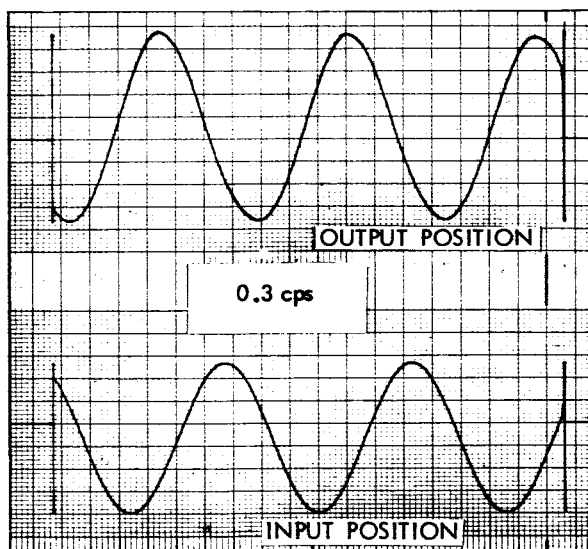


FIGURE 18c.—Sine wave response of simulator.

importance, though, is the fact that after the first overshoot no other overshoots of any sizable amplitude occur, indicating sufficient damping such that oscillations will not be sustained for any large duration. The impulse response, which characterizes the system, shows rapid convergence indicating not only absolute stability, but also the degree of stability. The response to a sine wave indicates that there is hardly a degradation of the output wave but there is a phase shift consistent with the phase frequency response (Fig. 17b).

A heat run was performed in which the simulator was adjusted so that 9.7 hp was dissipated in one clutch continuously. This is close to the 10 hp maximum thermal capacity of the coupling. During this test, the frame temperature rose 32°F above ambient and the water outlet temperature rose 28°F in about 30 minutes. Computations indicate that about 6 hp of heat was removed by the water coolant while the rest was lost by radiation and conduction to the bedplate and to the surrounding air. The result of this test indicates that there is practically no heating problem and temporary overloads could probably be handled readily.

For other test results particularly covering position errors under different conditions, measurement of coulomb friction and threshold of motion Reference 1 should be reviewed.

4. FULL SIZE EDDY CURRENT CLUTCH DRIVE

By using simulator test results and by projecting these to a full-scale version, a full-size eddy current clutch drive can be designed. The drive system could be similar to that proposed by Westinghouse in their report (Reference 1) and shown in Fig. 19. The two opposing couplings are mounted on a common bedplate together with their driving motors. The output shafts of the clutches are connected to independent gear paths, which in turn are connected to the sector or bull gear of either the X or Y axis of the antenna. The size of the couplings is determined from tracking, slewing, and stowing requirements, taking into account windloading, inertia, and friction effects. It is found that an Eaton WCS 216 clutch can be used on the X-axis and a WCS 215 on the Y-axis. These units can deliver nominally 220 hp and 140 hp respectively. These clutches could be standard production units, but most likely would be modified slightly to improve the field time constant. The driving motor would be standard ac squirrel-cage motors compatible with the horsepower requirement of the clutches. Squirrel-cage motors are chosen because of their efficiency, low cost, and lack of commutators or sliprings.

A closed cycle cooling system consisting of a radiator, a fan, and a sump located directly in the bedplate would be used. A closed cycle system has the advantage of economy and lack of extensive piping and filtering requirements. Normally the coolant for clutches is water, but other coolants can be used. To accommodate temperatures below 32°F, either of two things can be done: heaters can be installed to keep the water temperature above 32°F even in a non-operating

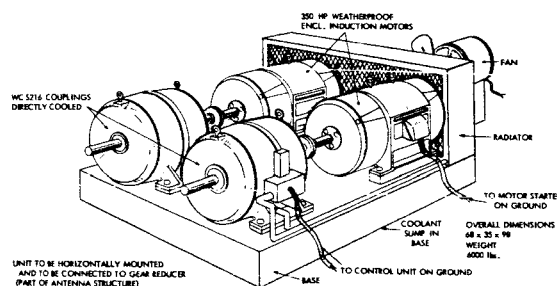


FIGURE 19.—Proposed layout of drive system for full-size antenna.

condition of the antenna, or glycol can be used as a coolant. Either approach has its advantages and disadvantages, the second approach being favored because power failure will not cause damage to the equipment. However glycol, because of its inherent characteristics, will require twice the flow required of water to remove the same amount of heat. The water demand of the clutch is about 1 gpm for each 10 hp slip or twice that rate when glycol is used. The maximum anticipated flow rate is 30 gpm for glycol. For economy coolant metering will be used, permitting higher flows only when the demand exists. The coolant demand can readily be determined from the discharge temperature of the coolant. This temperature rises as the slip horsepower increases. To prevent flooding of the clutch, free discharge to the sump is essential and is maintained easily for the X-axis drive. For the Y-axis, a problem exists since the bedplate is not stationary but has freedom to move $\pm 90^\circ$ from horizontal. This causes a problem with free gravitational flow, and flooding could occur since the clutch rotates end-for-end. Use can be made of bevel gears to change the rotation of the clutches from end-for-end to axial. Under this condition, modifications to the outlet ports could yield satisfactory flow if the sump is modified accordingly. The other alternative is to change the coupling from a direct to an indirect cooled unit. In this method coolant reaches the outside of the drum only, reducing cooling efficiency but permitting rotation of the clutch in any direction. This approach is preferred provided the modified thermal horsepower capacity is not exceeded.

Excitation for the fields of eddy current clutch could be provided from several devices such as dc generators, amplidynes, rototrols, magnetic amplifiers, SCR amplifiers, or transistor amplifiers. Rotating amplifiers have the disadvantage of limited brush and bearing life and are also notorious for RFI generation. SCR amplifiers have the same problem. Even though they have no moving parts, they are a source of RFI because of their steep wave fronts. Of course, attempts could be made to filter out RFI by using low-pass filters which do not affect signal frequencies. Filtering will probably be efficient, since the amplifiers are located in the control building rather than on the antenna structure itself. Magnetic

amplifiers are very similar in their operation to SCR's, but do not display the steep wave fronts. The disadvantage of magnetic amplifiers is their relatively larger time constant. Ideally suited for this application would be a linear transistor amplifier if it were not for the high power output requirement of over 1 kw. This amount of power is difficult to handle by a linear dc amplifier, although it is not impossible to build one.

The controller for this application would be very similar to that of standard controllers for this type of antennas. The controller could consist of operational amplifiers connected so as to give the desired transfer characteristics. The controller should have a clutch temperature monitor in addition to the ordinary protection devices, such as velocity, position, and torque limits. Because of the high slip horsepowers involved, an automatic temperature monitor should be included which either reduces the power to the antenna drive or shuts it down completely when predetermined limits are exceeded. This precautionary device will prevent costly burnouts of clutches.

It is expected that the dynamic response of the full-size drive will be similar to that of the simulator. Performance of course will be degraded somewhat by factors not taken into account in the simulation. But it is also anticipated that performance will be as good as that of comparable hydraulic drives.

5. COMPARISON

It is difficult to make a flat statement that the eddy current clutch drive is better than the more conventional drives. On the other hand, it is as difficult to say that any other drive is better than the eddy current clutch. Each type of drive has its advantages and disadvantages, and these have to be weighed against each other when considering a particular application. Even then a clear-cut decision is often hard to come by. An attempt will be made here to point out the advantages and disadvantages of some drives and how they compare with the eddy current clutch.

5.1 DC Motor Drive

For high performance dc drives armature-controlled motors are used exclusively. Three types of motors can be considered for control

applications: the standard NEMA frame motor, the more specialized servo motor, and the torque motor. Torque motors are low speed, permanent magnet field motors. Their construction is such that they can be mounted on an antenna axis directly without gearing and therefore without backlash that accompanies gearing. Smooth and linear operation can be obtained even at very low speeds. Unfortunately their maximum torque capability is limited to about 10,000 ft lbs on commercially available units, which rules out their application on radomeless 85 ft antennas.

NEMA design motors can be used up to about 200 hp requirements. Above that the inertia of these units becomes the limiting factor. NEMA design motors have a limited dynamic speed range, higher inertia, poorer linearity, larger ripple torque, and higher brake-away torque than their torque motor counterparts. Servomotors are designed to meet control applications better than NEMA design motors do, but are available only up to 10 hp.

All three types of motors can be excited by vacuum tube, solid state, or magnetic amplifiers for the lower hp ranges but almost always require rotating amplifiers for higher hp applications. The most commonly used connection with rotating amplifiers is the Ward-Leonard loop. A

simplified schematic is shown in Fig. 20. Often the motors and generators are modified to include series fields for copper loss compensation, interpole fields for reduction of armature reaction, and current feedback to reduce the generator field time constant. The following equations can be written for the Ward-Leonard loop:

$$V_f = (R_f + sL_f)I_f \tag{28}$$

$$V_g = K_g I_f = (R_a + sL_a)I_a + V_m \tag{29}$$

$$V_m = sK_v \theta \tag{30}$$

$$T_m = K_t J_a = s^2 J \theta \tag{31}$$

These equations can be combined to yield the transfer function of the dc drive:

$$\frac{\theta}{V_f} = \frac{\frac{K_g}{K_t R_f}}{s \left(s^2 \frac{L_a J}{K_t K_v} + s \frac{J R_a}{K_t K_v} + 1 \right) \left(\frac{L_f}{s R_f} + 1 \right)} \tag{32}$$

The quadratic factor has a natural frequency ω_n and damping factor ζ of:

$$\omega_n = \sqrt{\frac{K_t K_v}{L_a J}} \tag{33}$$

$$\zeta = \frac{R_a}{2} \sqrt{\frac{J}{K_t K_v L_a}} \tag{34}$$

5.2 Hydraulic Drive

The prime mover in a hydraulic drive could be a linear or rotary actuator for low horsepower applications or a hydraulic motor. Actuators can be connected directly to the load without gearing and have the advantage of low inertia and fast response. In applications where continuous rotation of high horsepower is required, hydraulic motors have to be used. Prime movers can be excited by either a servo valve or a variable displacement pump. The valve has the advantage of faster response over the pump but has the disadvantage of low power efficiency and of suffering degradation in performance due to contamination of the hydraulic fluid. Pumps which are preferred in the higher horsepower ranges very often require a secondary valve-piston combination to control the yoke of the pump. But his secondary loop can operate at lower pressures and much

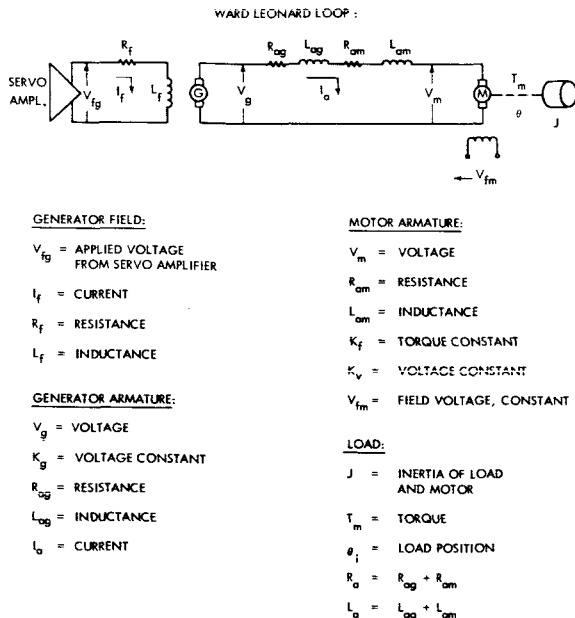
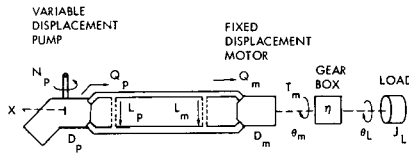


FIGURE 20.—Electrical drive system.



NOTATION:

PUMP

X	=	TOTAL STROKE
X_v	=	STROKE TO GIVE VELOCITY TO LOAD
X_a	=	STROKE TO GIVE ACCELERATION TO LOAD
X_c	=	STROKE TO SUPPLY COMPRESSIBILITY FLOW
N_p	=	PUMP SPEED
Q_p	=	PUMP FLOW
L_p	=	PUMP LEAKAGE
D_p	=	DISPLACEMENT OF PUMP
B	=	BULK MODULUS
V	=	VOLUME OF OIL UNDER COMPRESSION
K_v	=	VELOCITY GRADIENT
	=	$\frac{D_p N_p}{D_m}$
K_t	=	TORQUE GRADIENT
	=	$\frac{D_p D_m N_p}{L_p + L_m}$

MOTOR

θ_m	=	MOTOR SHAFT POSITION
Q_m	=	MOTOR FLOW
L_m	=	MOTOR LEAKAGE
D_m	=	MOTOR DISPLACEMENT
J_m	=	MOTOR INERTIA

LOAD

J_L	=	LOAD INERTIA
J	=	$J_m + \frac{1}{N^2} J_L$

FIGURE 21.—Hydraulic drive system.

lower power level than the primary pump-motor circuit. A simplified schematic is shown in Fig. 21. To develop the transfer function relating shaft position to pump stroke, equations can be written for parts of the total pump stroke and then the parts added:

$$X_v = \frac{s}{K_v} \theta_m \quad (35)$$

$$X_a = \frac{s^2 J}{K_t} \theta_m \quad (36)$$

$$X_c = \frac{s^3 V J}{D_m^2 K_v B} \theta_m \quad (37)$$

$$X = X_v + X_a + X_c = s \left(s^2 \frac{V J}{D_m^2 K_v B} + s \frac{J}{K_t} + \frac{1}{K_v} \right) \theta_m \quad (38)$$

By rearranging terms, the transfer function is obtained:

$$\frac{\theta_m}{X} = \frac{K_v}{s \left(\frac{V J}{D_m^2 B} s^2 + \frac{K_v J}{K_t} s + 1 \right)} \quad (39)$$

where the natural frequency and damping factor is:

$$\omega_n = D_m \sqrt{\frac{B}{V J}} \quad (40)$$

$$\zeta = \frac{L_p + L_m}{2 D_m} \sqrt{\frac{B J}{V}} \quad (41)$$

5.3 Comparison with Eddy Current Clutch Drive

Comparisons will be made between dc, hydraulic, and eddy current clutch drives at two power levels: at the simulator level or 10 hp range, and at the 85 ft antenna level or 200 hp range. To make the comparison meaningful, certain basic assumptions have to be made. Sizing of the driver will be done by using load characteristics from Table I and substituting these, together with driver characteristics, into the following equations to determine power requirements:

$$T = T_a + T_w + T_f \quad (42)$$

$$T_a = \text{acceleration torque (ft lbs)} = \frac{(J_L + 2N^2 J_m) \ddot{\theta}}{\quad} \quad (43)$$

$$T_w = \text{wind torque (Table I)}$$

$$T_f = \text{friction torque (ft lbs)} = 0.1 (T_a + T_w) \quad (44)$$

$$P = \text{required load power (hp)} = T \dot{\theta} / 5250 \quad (45)$$

where

$$J_L = \text{load torque (Table I)}$$

$$J_m = \text{torque of driver (ft lbs sec}^2\text{)}$$

$$N = \text{gear ratio}$$

After the driver is sized, the dynamic response will be determined by using equations 33 and 34 for dc, 40 and 41 for hydraulic, and 15 for eddy current clutch drives.

Table III shows comparisons between the various systems. Care has to be exercised in using numbers from this table, inasmuch as the numbers have to be interpreted correctly and should not be used indiscriminantly as an absolute measure.

For low horsepower ranges several drive types have been considered, but in the high horsepower

TABLE III—Comparison

Characteristic	Simulator Size					Full Size	
	DC-NEMA Motor	DC Torque Motor	Hydraulic Pump-Motor	Hydraulic Valve Motor	Eddy Current Clutch	Hydraulic Pump-Motor	Eddy Current Clutch
PRIME MOVER							
Manufacturer.....	GE	Inland	Vickers	Vickers	Eaton	Vickers	Eaton
Part Number.....	504	T-18004-A	Model A	Model B	WCS 2062	MFA 120	WCS 216
Power—hp.....	15	11.4	9	14.2	10	200	220
Speed—rpm.....	300	60	12,500	10,000	1200	1800	1200
Inertia—slug ft ²	1.43	3.1	1.56×10^{-5}	4.73×10^{-5}	.155	.292	1.92
Torque to Inertia Ratio— ft lbs/sec ²	98,000	323,000	7,900	15,400	16,800	8,230,000	634,000
Weight—lbs per unit.....	1915	600	1.7	3.3	270	600	1900
POWER GENERATOR							
Type.....	Amplidyne	Amplidyne	Pump	Valve	AC Motor	Pump	AC Motor
Manufacturer.....	GE	GE	Vickers	Moog	Westinghouse	Vickers	
Part Number.....	624 F	622 F	Model A	Series 32	284 US	PVA 120	
Power—hp.....	20	13.5	9	4.7	10	200	350
Speed—rpm.....	1750	1750	12,500	N/A	1160	1200	1160
Weight—lbs.....	825	600	1.7	.75	290	1150	3100
SYSTEM							
Weight—lbs.....	4655	1200	5.1	7.4	830	2,350	10,000
Natural Frequency.....	61.4	36.8	200	412	5.78	122	7.1
Damping Factor.....	.145	.416	.0276	.038	.292	.0528	.355

class only two practical drives are mentioned. Dc motors, because of their high inertia, cannot be used unless special designed motors are developed with smaller rotors.

When the prime mover characteristics are considered a variation of horsepower levels is noticed. For the simulator size drives this variation is mainly caused by steps in component ratings. For example, in the dc motor drive about 12 hp are required, but the available size closest to this is a 15 hp unit. The valve-motor drive seems to require a larger unit than the pump-motor drive. This is not because a larger load is encountered but because the smaller motor would require a system pressure of over 3000 psi, which is generally undesirable. For full-size drives the eddy current clutch requires a higher horsepower rating than the hydraulic drive because of its high acceleration torque requirement due to higher inertia of the clutch.

Inertias seem to vary considerably for different drives, but so do the shaft speeds of these drives. To make a better comparison, inertias that are referred to a particular speed should be considered. Such a comparison is made graphically in Fig. 22. As can be seen from this figure, eddy current clutches have higher inertias than their dc motor counterparts up to a rating of 100 hp; above that the situation reverses. Hydraulic motors have a decided advantage in inertia over both motors and clutches.

A figure of merit in the motor is the torque to inertia ratio or the torque-squared to inertia ratio. The latter is sometimes preferred because it remains constant at both ends of a gear reducer. At simulator power levels, the dc motors and in particular the dc torque motor are the preferred units, while the eddy current clutch runs second. At the high power level, though, the hydraulic motor has an advantage over the clutch.

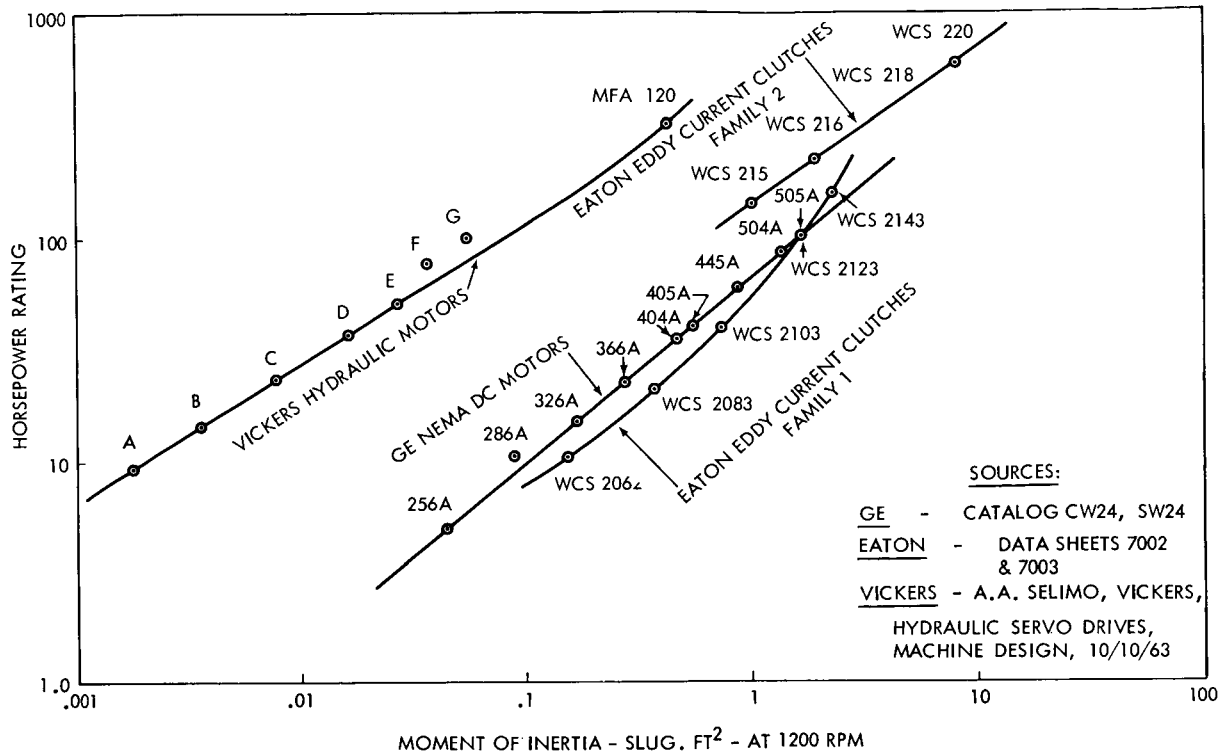


FIGURE 22.—Comparison of inertias.

From a weight standpoint the hydraulic motor is by far the preferred unit, but the eddy current clutch is lighter than the dc motor. These figures, however, are somewhat misleading as far as dc motors are concerned. The torque motor requires no gearing, but all other units do. Consequently a weight for the appropriate gear reducer should be considered. Also a low speed NEMA motor was chosen for the comparison. If weight is a major system design criteria, a 3500 rpm motor could be chosen in place of a 300 rpm unit, which reduces the weight from 1915 lbs to 285 lbs. The figures in the table also do not take into account that special designs could be arrived at that would demonstrate much improved characteristics. Power generators also make hydraulics more attractive than the eddy current clutch and dc motors as far as weights are concerned.

Total system weights given in the table are the sum of prime mover and power generator weights. These figures neglect weights of electronic controllers, gearing, mounting hardware, reservoirs and piping for hydraulic systems, cooling system

for eddy current clutch systems, etc. All of these factors would have to be considered in arriving at a total system weight. Indications are that the hydraulic system still will be lighter than the clutch or dc motor systems.

To study the transient response, it is sufficient for a preliminary check to determine the predominant natural frequency and damping factor. For the dc and hydraulic drives this can be done simply by using equations 33, 34, 40, and 41. The clutch drive because of the type of its transfer function, does not lend itself to direct comparison. To put all systems on an equal basis, a tachometer feedback will be introduced with an arbitrarily chosen loop gain of 10. This addition modifies the original natural frequency ω_n and damping factor ζ for dc and hydraulic drives as follows:

$$\omega_n/\text{modified} = \omega\sqrt{11} \tag{46}$$

$$\zeta/\text{modified} = \frac{\zeta}{\sqrt{11}} \tag{47}$$

The transfer function of the clutch drive G_c is modified by the feedback H_c to give a new G_c' :

$$G_c = \frac{K_t/JR}{s^2\left(\frac{L}{sR} + 1\right)} \quad (15)$$

$$H_c = sKg \quad (48)$$

$$G_c' = \frac{G_c}{1 + G_c H_c} = \frac{1/K_g}{\left(s^2 \frac{LJ}{K_g K_t} + s \frac{RJ}{K_g K_t} + 1\right)} \quad (49)$$

where

$$\omega_n/\text{clutch} = \sqrt{\frac{K_g K_t}{LJ}} \quad (50)$$

$$\zeta/\text{clutch} = \frac{R}{2} \sqrt{\frac{J}{K_g K_t L}} \quad (51)$$

$$K_g = 10 \frac{JR}{K_t} \quad (52)$$

Substituting values into equations 46, 47, 50, and 51, the predominant natural frequency and its damping factor is obtained and entered in Table III. It is assumed here that the predominant natural frequency is the sole determining factor for bandwidth. Consequently this frequency is somewhat hypothetical because other time constants have been neglected which could influence bandwidth. But, if a further assumption is made that all of these other effects are present in all drive systems to approximately the same degree, then the natural frequency does represent an index of performance in the comparison of various systems.

Among the low horsepower drives the hydraulic valve-motor system has the highest natural frequency of 412 rad/sec, with the pump-motor system second with 200 rad/sec. One characteristic of hydraulic drives is that a high loss in piping is encountered. For example, if the piping length were doubled for the pump-motor system, the natural frequency would drop to 128 rad/sec. This equivalent effect is not present in dc motor or eddy current clutch drives; doubling lead length will not affect performance significantly. Also, the hydraulic drives are highly under-damped; but adjustments of the tachometer loop could easily improve this situation. Dc motor drives have a smaller bandwidth than hydraulic but exceed clutch drives by a good margin. The fact that the clutch drive has a natural frequency of

only 5.78 rad/sec is not necessarily a limiting factor for system design; it just makes compensation more difficult. For the full-size drive the hydraulic system is preferred over the clutch system because of a much larger bandwidth.

So far, in comparing drive systems, use has been made of the numbers in Table III; and the hydraulic systems seem to be superior. However, to complete the comparison, other factors have to be considered—factors that do not lend themselves readily to being put into numbers.

The first of these factors is reliability. The clutch drive has a limitation on life because of wearing mechanical parts, namely bearings and coolant seals. Besides bearings and oil seals, the hydraulic drive has more moving parts, i.e., pistons, yokes, and other parts which tend to lower the reliability. The reliability is affected also by possible contamination of the fluid. Contamination is caused by introduction of impurities like sand, dirt, metal filings, etc. during assembly and during operation of the system due to wear of hydraulic components. Chemical changes of the fluid due to extreme temperatures and pressure changes tend to add sludge and gum. A conservative estimate of mean-time-between-failure (MTBF) for a 200 hp drive is 16,500 hours for the hydraulic system and 36,500 hours for the clutch system, thus giving the latter system a better than 2:1 margin. Dc motor systems should have a MTBF similar to that of the clutch system except that brush and commutator wear has to be taken into account, thus lowering the MTBF by some amount.

Maintenance is very closely related to reliability; thus the hydraulic system requires more maintenance than the clutch drives. In the clutch drive maintenance consists mainly of greasing bearings and checking the coolant level. Hydraulic system maintenance requires regular oil changes, system cleaning, greasing of bearings, tightening of fittings, changing of filter elements, etc.

Another factor not considered in Table III is RFI. Hydraulic systems by nature do not generate RFI. Eddy current clutch drives, because of their construction, do not have slip rings or commutators and thus also do not generate RFI, the only exception being the fact that a relative large dc field excitation is required on the clutches.

If this excitation is derived from a commutating source such as SCR's, thyratrons, or dc generators, RFI may appear. This can be prevented however by choosing the proper amplifier or by using RFI filters and good shielding practices. Dc motor drives are inherent RFI generators, but by proper design successful drive systems can be developed.

5.4 Other Clutches

Before leaving the topic of comparisons, it may be in order to mention other types of clutches and how they compare with the eddy current clutch. Other types are friction disk, crystal, loudspeaker, hysteresis, magnetic particle, and magnetic fluid clutches.

Friction disk clutches are either mechanically or electrically operated devices similar to ordinary automobile clutches. These units can transmit high torques, have a relative low durability, and are basically suited for on-off operation only. Friction disk clutches are not suited for antenna drive applications. Crystal and loudspeaker clutches, even though utilizing different principles, are on-off devices. Their low power rating puts them into the instrument servo class rather than a power drive class.

Lately a lot of interest has been shown in the application of hysteresis clutches for instrument servos and low horsepower systems. However, no information is available on developments of higher horsepower units.

Magnetic particle and fluid clutches are available with characteristics close to those of the eddy current clutch. Basically, these clutches consist of two disks separated by an airgap which is filled with magnetic powder or magnetic particles suspended in fluid. A magnetic field causes the particles to form a more or less rigid bond between the plates, depending upon the excitation level. This bond is capable of transmitting torque independent of the speed of the clutch. Magnetic particle and fluid clutches up to about 100 hp ratings are available commercially. Similar to the eddy current clutch, these clutches are subjected to high slip power causing heat dissipation problems. Excessive heat could chemically affect the magnetic particles to cause a change in their characteristics. Magnetic particles also cause sealing and wear problems and pose a difficult

design requirement, which necessitates that the particles be distributed evenly in the airgap under all operating and non-operating conditions.

6. CONCLUSION

On the basis of comparisons between hydraulic, dc, and eddy current clutch drives, it can be concluded that the clutch drive has one major advantage: That is in reliability and consequently decreased maintenance requirements. Disadvantages are relatively high inertia and therefore a poor torque to inertia ratio, predominant field time constant which limits bandwidth, non-linear operation, and heat dissipation problems which cause poor efficiency. But despite these disadvantages it is possible to design and build a satisfactory control system that assures required stability and tracking accuracy.

Historically, the eddy current clutch has been used for many years in automatic machinery, but very few applications are known to antenna systems. One of NASA's antennas on Wallops Island had an eddy current clutch drive. This antenna was designed and built by General Bronze, a company which has since discontinued the manufacture of antennas. The drive system of this particular antenna has a history of burned-out clutches due to excessive heat dissipation and poor control ability. However the general opinion is that this situation was not caused by inherent problems with the clutch but rather by choosing an underrated unit because of lack of information on and experience with clutches. Eddy current clutch drives were proposed for the 250 ft antenna of JPL and the 600 ft Sugar Grove antenna of the Navy. The latter antenna system was to utilize clutch drives but the whole project was cancelled before completion, and consequently the drive system could not be tested and evaluated.

To prove conclusively the advantages or disadvantages of the eddy current clutch drive system over the hydraulic drive, it is necessary to build a full-size clutch drive. Only then, after several months of operations under realistic conditions, can a firm opinion be formed. While certain characteristics of the hydraulic drive are superior to those of the eddy current clutch as has been pointed out here, either drive appears capable of meeting all dynamic performance specifications so that these differences are of secondary

importance in the application under consideration. In view of the down-time experience with the present hydraulic drives the reliability factor becomes the most important characteristic when making a comparison. If the clutch system should prove itself more reliable by a reasonable factor, and should no unusual problems develop such as excessive heating, poor control, etc., it will be well worthwhile to choose eddy current clutches for future applications and even to consider modification of present antennas. Reliability of the drive system is very important since it is expensive to have down-time on a multimillion dollar antenna system because of a hydraulic leak or a clogged servo valve.

REFERENCES

1. W. O. OSBON, R. F. LEEPA, R. F. SMITH, "Electric Drive System Study for 85 ft Data Acquisition Antennas," Westinghouse Electric Corp., March 1963.
2. A. M. FLIDLIDER, "Transient Responses in the Magnetic Circuits of Electromagnetic Clutches," *Automatika i Telemekhanika*, Vol. 20, No. 1, pp. 31-44.
3. P. N. KOPAI-GORA, "Concerning Some Properties of Ferromagnetic Clutches," *Automatika i Telemekhanika*, Vol. 19, No. 4, pp. 366-375.
4. J. G. TRUXAL, "Control Engineer's Handbook," McGraw-Hill Book Co., 1958.
5. H. CHESTNUT and R. W. MAYER, "Servomechanisms and Regulating System Design," Vol. 2, John Wiley and Sons, 1955.

A METHOD FOR COMPARING THE REFLECTING POWER OF CONCAVE GRATINGS IN THE SOFT X-RAY REGION*

D. H. TOMBOULIAN AND W. E. BEHRING

The equations needed for calculating the intensity of a spectral line produced by a concave grating are developed for a source of given monochromatic radiant power. These equations are then applied to give a comparison of the reflectance of gratings tested with the same source and with the other conditions also similar. Another application for the purpose of comparing the reflectance of the same grating in different orders of the same line is also described. The results are given for the determination of the reflectance of a particular Siegbahn glass grating.

I. INTRODUCTION

The purpose of this report is to describe a method for evaluating the spectral response of a grating under test by comparing its response with that of a second grating whose efficiency is known. The particular approach is designed as a substitute for the laborious procedure by which the performance of the given grating may be measured directly for the particular mounting. The calibration is best carried out by using an intense and dense line spectrum such as the one produced in a capillary spark discharge.

II. DEFINITION OF SYMBOLS

In the list below we define most of the symbols to be used in various relations needed for the reductions. The geometrical aspects of the grazing incidence mounting are seen in Fig. 1.

- ρ = radius of curvature of grating blank,
- σ = grating space,
- m = order of diffraction,
- λ = wavelength of radiation,
- w = width of entrance (Rowland) slit,
- W_g = illuminated width of grating,
- w' = width of magnified image on Rowland circle,
- α = angle of incidence measured from normal,
- β = angle of diffraction measured from normal,

ϕ_0 = angle of grazing incidence measured from surface,

x = distance from central image to position of spectrum line as measured along Rowland circle,

$(\phi_0 + x/\rho)$ = angle of diffraction measured from surface,

r = distance from source to pole of grating,

r' = distance from pole to line image,

d = distance from source to entrance slit,

$2l$ = height of illuminated rulings,

z = height of point source above median plane,

z' = height of image above median plane.

The following relations are also included for reference:

$$r' = \rho \cos \beta; r = \rho \cos \alpha + d.$$

The intensity maxima are given by the grating equation:

$$m\lambda = \sigma \left[\cos \phi_0 - \cos \left(\phi_0 + \frac{x}{\rho} \right) \right] = \sigma [\sin \alpha - \sin \beta]. \quad (1)$$

By geometry, the image width along the Rowland circle is found to be

$$w' = \frac{w}{\sin \left(\phi_0 + \frac{x}{\rho} \right)} = \frac{w}{\cos \beta}. \quad (2)$$

*Published in *Applied Optics*, 3(4): 501-505, April 1964.

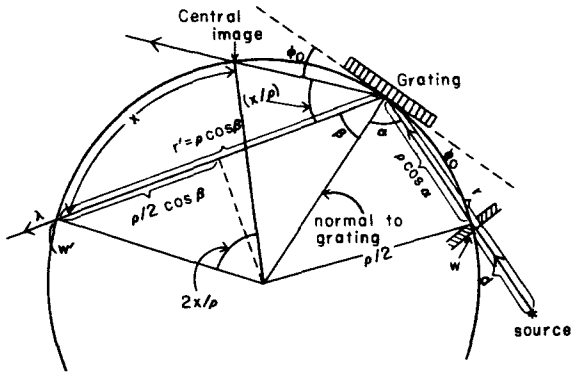


FIGURE 1.—Geometrical features of the grazing incidence mounting.

III. RADIANT ENERGY INCIDENT ON GRATING

Let $J(\lambda)$ denote the monochromatic radiant power emitted in a given direction per unit solid angle and unit area of source. The radiant power incident on the grating is calculated by considering the solid angle subtended by the grating at a given point in the source and by considering the

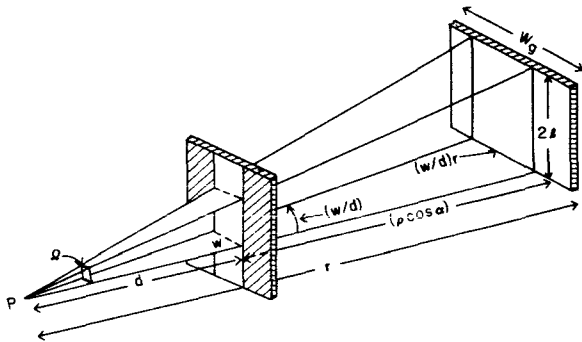


FIGURE 2.—Solid angle subtended at the source P . The angle is determined by the width W of the entrance slit and the height of rulings $2l$.

effective area of the source. By referring to the diagram in Fig. 2, it is seen that Ω , the solid angle subtended at a source point is given by

$$\Omega = \left(\frac{2l}{r}\right)\left(\frac{w}{d}\right) = \frac{2lw}{rd} \tag{3}$$

The solid angle is determined by w , the width of the entrance slit, and by $2l$, the height of the illuminated portion of the rulings. For a $25\text{-}\mu$ slit, source to grating distance of 20 cm, $d = 10.5$ cm, and $2l = 1.0$ cm, Ω is nearly 1.2×10^{-5} sr.

Next we wish to calculate the effective area of the source. The procedure is indicated schematically in the sketch of Fig. 3. Starting with the projected width of the grating $W_g \cos \alpha$, one finds the useful width of the source to be $(W_g/\rho)d$, provided (W_g/ρ) represents the smallest angular width γ for a given experimental arrangement.

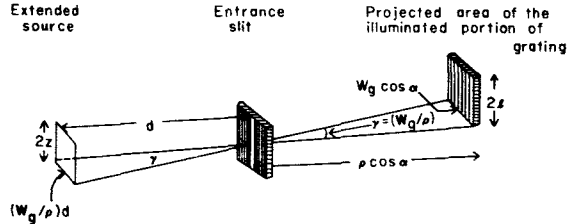


FIGURE 3.—Effective area of source. Starting with the projected width of the grating the angle γ is obtained from $\gamma = (W_g \times \cos \alpha / \rho \cos \alpha) = (W_g / \rho)$. For the useful area of source this gives $2z$ times γd or $(2zW_g d / \rho)$.

It could be that the effective width of the source is the actual width of the source, or if a post-slit of width W_b is located at a distance b from the Rowland slit, the value of angle γ as given by (W_b/b) may be smaller than (W_g/ρ) . In this case the effective width of the source would be $(W_b/b)d$. If $2z$ is the height of the source which is effective in illuminating the full height of the rulings, the effective area A_{eff} of the source for the present choice of γ becomes

$$A_{eff} = 2z(W_g/\rho)d \tag{4}$$

If $S(\lambda)$ denotes the radiant power incident on the grating for the case of a line radiation of wavelength λ , then, utilizing the results given by relations (3) and (4), $S(\lambda)$ may be represented as

$$S(\lambda) = J(\lambda) \left(\frac{2lw}{rd}\right) \left(\frac{2zW_g d}{\rho}\right) \tag{5}$$

provided that $J(\lambda)$ is isotropic and uniform over the source. Furthermore, when the spectrum is recorded photographically, the radiant energy incident on the grating during an exposure time T is

$$E_i = S(\lambda)T = J(\lambda)T \left[\frac{4lwW_g z}{r\rho}\right] \tag{6}$$

If the radiant power is detected electronically, Eq. (5) may be modified to denote the incident photon flux.

IV. ENERGY IN IMAGE OF SPECTRAL LINE

From a microphotometer trace of the spectrogram and with the aid of the relative intensity calibration of the particular emulsion used, one can arrive at a quantity M which is a measure of the radiant energy per unit area of the line image. The measurement of the radiant energy associated with the entire image of the spectral line involves a knowledge of the area of the astigmatic image.

Recently Namioka¹ has re-examined the theory of the concave grating and has pointed out certain errors in the treatment of the problem as given by Beutler.² However, Namioka states that Beutler's treatment of astigmatism is correct. We begin with Eq. (36) in Beutler's paper. A point source at a height z above the median plane and at a distance r is imaged at a height z' and at a distance r' (see Fig. 4). The coordinate l is the

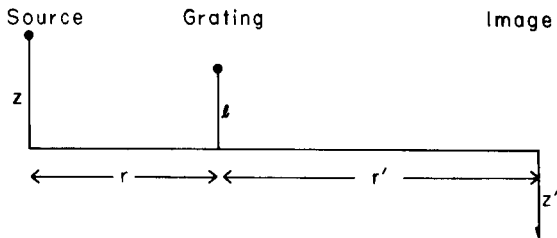


FIGURE 4.—Distances involved in the condition for calculating the vertical height of an astigmatic image.

length of rulings measured from the center. The light paths lie in vertical planes which include z , z' , and l at the angles α and β . Solving Beutler's formula for z' , we find

$$z' = -z \left(\frac{r'}{r} \right) + l \left[\left(\frac{r'}{r} \right) - (\cos \alpha) \left(\frac{r'}{\rho} \right) + 1 - (\cos \beta) \left(\frac{r'}{\rho} \right) \right]. \quad (7)$$

Replacing r by its equal $(\rho \cos \alpha + d)$ and r' by $\rho \cos \beta$ and upon simplification, we find

$$z' = -zg(\alpha, \beta, \rho) + lf(\alpha, \beta, \rho), \quad (8)$$

where the angular functions are represented by

$$g(\alpha, \beta, \rho) \equiv \left\{ \frac{\rho \cos \beta}{\rho \cos \alpha + d} \right\} \quad (9)$$

and

$$f(\alpha, \beta, \rho) \equiv \left[\frac{\rho \cos \beta}{\rho \cos \alpha + d} - \cos \alpha \cos \beta + \sin^2 \beta \right]. \quad (10)$$

In passing we note that, for an extended source of height z focused on the entrance slit ($d=0$), z' reduces to

$$z' = -z \frac{\cos \beta}{\cos \alpha} + l \left[\sin^2 \beta + \frac{\sin^2 \alpha \cos \beta}{\cos \alpha} \right] \quad \text{for } d=0, z \neq 0, \quad (11)$$

and for the more special case of a point source at $d=z=0$, we have

$$z' = l \left[\sin^2 \beta + \frac{\sin^2 \alpha \cos \beta}{\cos \alpha} \right], \quad (d=z=0). \quad (12)$$

The expression in (12) is the one commonly cited in elementary treatments. It gives the largest z' for given values of α , β , and ρ . In the cases represented by (11) and (12) the astigmatic effect is seen to be independent of ρ , the grating radius. On the other hand, relation (8) which connects the lengths of source, image, and rulings does involve ρ .

The area of the astigmatic line image is deduced from the diagram in Fig. 5. The quantity z' is

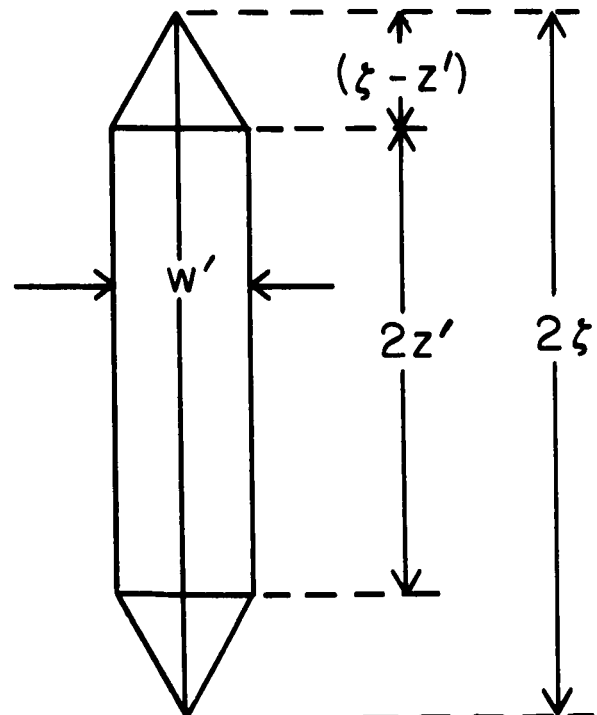


FIGURE 5.—Area of astigmatic image.

given by (8), while ζ , as given by Beutler, has the form

$$\zeta = +zg(\alpha, \beta, \rho) + lf(\alpha, \beta, \rho), \quad (13)$$

where g and f have already been defined by (9) and (10). The area as pictured in Fig. 5 will be assumed to be built of the rectangular portion $2w'z'$ plus the area of two triangular caps equal to $w'(\zeta - z')$. From (8) and (13), one obtains

$$(\zeta - z') = 2zg(\alpha, \beta, \rho). \quad (14)$$

Hence A , the image area, is

$$A = 2w'z' + w'(\zeta - z') = 2lw'f(\alpha, \beta, \rho), \quad (15)$$

where w' is the width of the line and $2l$ is the height of the illuminated portion of the ruling. Consequently, for the radiant energy diffracted into the entire image of a particular line in a specified order m , we may write

$$E_{r,m} = MA = 2Mlw'f(\alpha, \beta, \rho), \quad (16)$$

where M is a measure of the radiant energy per unit area of the line image in question.

V. CALCULATION OF REFLECTANCES

We are now in a position to formulate an expression for the comparison of grating efficiencies. Considering the radiant energy associated with a line radiation of wavelength λ , we define the reflectance $R(\lambda)$ as the fraction of the incident radiant energy which appears in the entire image of the diffracted line. In symbols

$$R(\lambda)_m = \frac{E_{r,m}}{E_i}, \quad (17)$$

where the subscript m refers to the order of the diffracted line. With the aid of previously developed relations (6) and (16), $R(\lambda)_m$ may be written as

$$R(\lambda)_m = \frac{\rho(\rho \cos \alpha + d)[Mw'f(\alpha, \beta, \rho)]_m}{2TJ(\lambda)wW_\rho z}. \quad (18)$$

In computations, the quantities in the bracket appearing in the numerator of (18) must be appropriate to the diffracted line of order m . The

expression in (18) applies to a *given* grating and is a measure of its spectral response. To compare the responses of two *different* gratings in a given spectral order, it is necessary to write down the ratio $R_1(\lambda)/R_0(\lambda)$, where $R_1(\lambda)$ is the response of the grating to be tested and $R_0(\lambda)$ is the previously determined reflectance of the grating which is taken as a standard.*

By the application of (18) we find

$$\frac{R_1(\lambda)}{R_0(\lambda)} = \frac{\left\{ \frac{[M_1 w_1' f_1(\alpha, \beta, \rho)](\rho_1 \cos \alpha_1 + d_1) \rho_1}{2T_1 J_1(\lambda) w_1 W_{\rho_1 z_1}} \right\}}{\left\{ \frac{[M_0 w_0' f_0(\alpha, \beta, \rho)](\rho_0 \cos \alpha_0 + d_0) \rho_0}{2T_0 J_0(\lambda) w_0 W_{\rho_0 z_0}} \right\}} \quad (19)$$

where the subscripts 0 and 1 refer, respectively, to quantities associated with the standard and test grating characteristics or measurements. The ratio in (19) is written for the comparison of first-order reflectances.

It may be possible to arrange the experimental conditions so that the same sources, exposure times, Rowland slit-widths, and the same illuminated areas of gratings are used. This means that in (19)

$$J_0(\lambda) = J_1(\lambda); z_0 = z_1; T_0 = T_1; w_0 = w_1; W_{\rho_0} = W_{\rho_1}.$$

With such an arrangement and recalling that $w' = w/\cos \beta$, relation (19) reduces to

$$\frac{R_1(\lambda)}{R_0(\lambda)} = \frac{M_1 \cos \beta_0 f_1(\alpha_1, \beta_1, \rho_1)(\rho_1 \cos \alpha_1 + d_1) \rho_1}{M_0 \cos \beta_1 f_0(\alpha_0, \beta_0, \rho_0)(\rho_0 \cos \alpha_0 + d_0) \rho_0} \quad (20)$$

Rearranging (10) we obtain

$$f(\alpha, \beta, \rho) = \cos \beta \left\{ \left(\frac{1}{\cos \alpha + d/\rho} \right) - \cos \alpha + \sin \beta \tan \beta \right\}. \quad (21)$$

*The grating referred to as "standard" is a dense glass grating ruled lightly by R. W. Wood. Its radius of curvature is 1.53 m with a spacing of 8480 Å. Its response in the soft x-ray region was determined by illuminating the grating by the sensibly monochromatic emission bands of C, Be, Al, and suitable radiations from discharge sources. The calibration process involved the measurement of the radiant power input and also the radiant power associated in the diffracted beam in a particular order. The calibration was then used to analyze the spectral distribution of the synchrotron radiation and was successful in predicting the theoretically known shape of the spectral power distribution from this source.

Substituting (21) into (20), the working equation for the comparison of responses in the *same order* for *two different gratings* becomes

$$\frac{R_1(\lambda)}{R_0(\lambda)} = \frac{M_1 \left\{ \left(\frac{1}{\cos \alpha_1 + d_1/\rho_1} \right) - \cos \alpha_1 + \sin \beta_1 \tan \beta_1 \right\} (\cos \alpha_1 + d_1/\rho_1) \rho_1^2}{M_0 \left\{ \left(\frac{1}{\cos \alpha_0 + d_0/\rho_0} \right) - \cos \alpha_0 + \sin \beta_0 \tan \beta_0 \right\} (\cos \alpha_0 + d_0/\rho_0) \rho_0^2} \quad (22)$$

If it is desired to compare reflectances in various orders, one of two procedures may be followed. One may take the grating under test, and calibrate it in the first order against a standard. From the second- to first-order line intensity ratios produced by the test grating, one may proceed to calculate the second-order reflectance. A second

approach would be to compare second-order lines in the spectrum of the test grating with corresponding lines of the spectrum produced by the standard. Below, we express analytically the working equation for the calculation of second-order reflectance $R_{II}(\lambda)$ by direct comparison with the first-order reflectance $R_I(\lambda)$ of the *same grating*. Thus, from (18) it follows that

$$\frac{R_{II}(\lambda)}{R_I(\lambda)} = \left(\frac{M_{II}}{M_I} \right) \frac{\left\{ \frac{1}{\cos \alpha + d/\rho} - \cos \alpha + \sin \beta_{II} \tan \beta_{II} \right\}}{\left\{ \frac{1}{\cos \alpha + d/\rho} - \cos \alpha + \sin \beta_I \tan \beta_I \right\}}, \quad (23)$$

where the Roman numerals I and II have been used to refer to first- and second-order quantities associated with a single spectral record produced by the same grating.

VI. PROGRAMMING

The computations are handled easily by first selecting a set of values for λ , thereby obtaining $\sin \beta$, the sine of the angle of diffraction, from (1). We recall that $\sin \alpha$, m , and σ are constants. The remaining trigonometric functions appearing in (22) or (23) are readily tabulated. The calculations are repeated for the same first-order lines produced by a second grating or for the comparison of first- and second-order intensities produced by the same grating.

VII. EXPERIMENTAL

The experimental approach involves the use of two similar spectrometers. The standard grating is permanently mounted in one of them at a known angle of grazing incidence. The grating to be studied is similarly mounted in an auxiliary spectrometer and is focused at a predetermined angle of incidence, not necessarily the same as that of the standard spectrometer. As we have seen,

the comparison of grating efficiencies involves a number of parameters such as: the size of source, its distance from the entrance slit, the width of the entrance slit, the angle of incidence, the area and the height of the ruled portion of the grating illuminated by the source, the radius of curvature of the grating blank, and the grating constant.

Using a convenient source, and keeping operating conditions of this source as constant as possible, one obtains spectrograms with each of the two spectrometers. From the densitometer trace of each spectrogram and from the intensity calibration curve of the emulsion one arrives at a measure

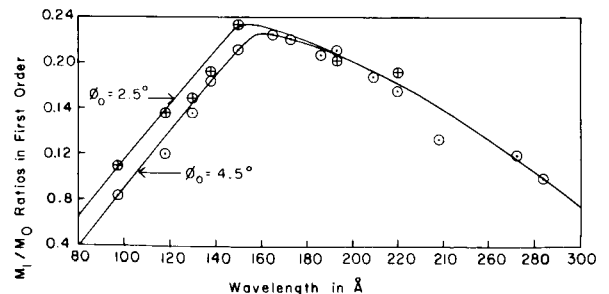


FIGURE 6.— (M_1/M_0) ratios for two angles of grazing incidence in the first order. Siegbahn grating W1-576-314.

of the radiant energy density in each of the line images.

VIII. RESULTS

In the way of sample calculations we present results obtained on Siegbahn grating W1-576-314. Figure 6 gives the (M_1/M_0) ratios for two different angles of incidence. Here the subscript zero refers to the first order of the standard grating which was calibrated against the radiation from a synchrotron, the intensity of which was calculated by Tomboulian and Hartman.^{3,*} The subscript one refers to the Siegbahn grating.

The scatter in the experimental points is due mostly to photometric errors. No reduction of data was contemplated when the initial records were obtained so that the usual refinements were not introduced in the registration of the spectra. For calculating the reflectivities plotted in Fig. 7, the values of the (M_1/M_0) ratio were taken from

*If the image of the source is focused on the entrance slit, Eq. (18) is still valid if one sets the slit to source distance, d , equal to zero. It can be shown that Eq. (6) is likewise still valid with $d=0$ (hence, $\tau = \rho \cos \alpha$). Therefore, these equations may be used with all of the usual mountings of the concave grating. If the source produces continuum rather than line radiation, Eq. (18) can be easily modified by replacing $w'/J(\lambda)$ by $m\rho/J_c(\lambda)\sigma \times \cos\theta$ where $J_c(\lambda)$ is the $J(\lambda)$ per unit wavelength interval.

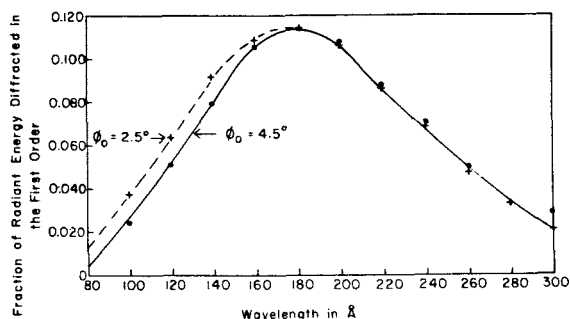


FIGURE 7.—Response curves for two different angles of grazing incidence ϕ_0 . Siegbahn grating W1-576-314.

the smoothed out curves, and individual experimental points were not used. Figure 7 indicates that at either of the two angles of incidence, the grating response attains its peak value of 11 percent near 180 Å. Below this wavelength the grating efficiency is higher at the smaller angle of grazing incidence.

REFERENCES

1. T. NAMIOKA, *J. Opt. Soc. Am.* **49**, 446 (1959).
2. H. G. BEUTLER, *J. Opt. Soc. Am.* **35**, 311 (1945).
3. D. H. TOMBOULIAN and P. L. HARTMAN, *Phys. Rev.* **102**, 1423 (1956).

A NARROW-BAND TRACKING FILTER*

RONALD WOODMAN†

The characteristics of a servo phase-lock tracking filter is described herein. The filter tracks a 100 cps carrier with a bandwidth of either 0.3 cps or 0.03 cps. The advantages of this type filter and its use is also discussed.

This report is to supplement the NASA TN D-1419 report, "A Phase-Locked Phase Filter for the Minitrack System," to cover the improved prototype built by the Network Engineering Branch, Data Systems Section.

After due consideration, the name "Phase-Locked Phase Filter" has been changed to "Narrow Band Tracking Filter." This name implies its function and the technique it uses.

The block diagram and system theory is essentially the same as described in the above Technical Note. There were two different approaches that could have been taken: to shift the reference and phase-lock it to the signal or to shift the signal and phase-lock it to the reference. The first approach has been taken since the output is a clean 100 cps signal whose phase can be measured and digitalized with existing Minitrack equipment.

The system has been designed for two different bandwidths: 0.3 cps with 10 sec. acquisition time and 0.03 cps with 100 sec. acquisition time. The dynamics of the system are slightly different than the dynamics of the system described in the above mentioned TN report. The open loop transfer function is still of the form.

$$G(s) = \frac{K(s+z)}{s^2(s+p)} \quad (1)$$

but the ratio p/z has been increased from 5 to 9. A ratio of nine generates a root locus as shown in Figure 1, where for a particular gain ($K=p/3$) the three poles of the close loop transfer lie on the same point of the real axis. This results in a critically damped second order servo with 40

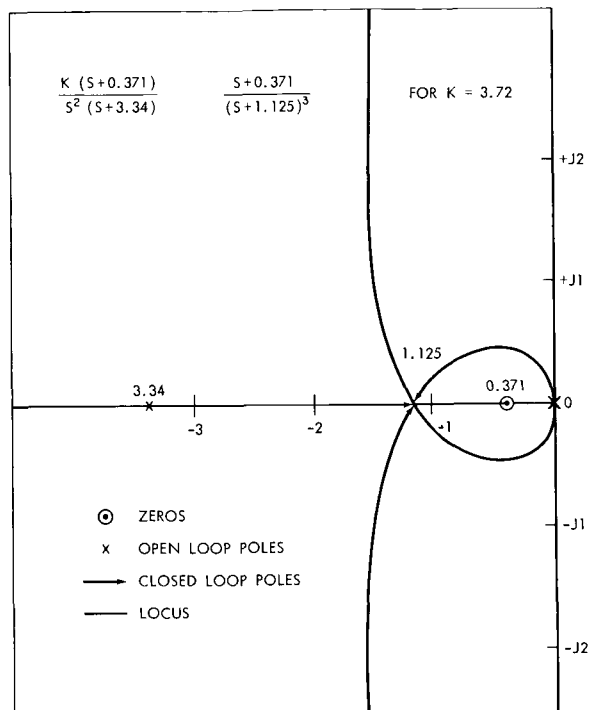


FIGURE 1.—Open and closed loop root locus.

db/sec roll-off. A value for z equal to 0.371 or 0.0371 has been chosen to give a bandwidth of 0.3 or 0.03 cps and a acquisition time of 10 or 100 seconds. The resulting closed loop transfer function is

$$\frac{\phi_u(s)}{\phi_t(s)} = \frac{0.261(S+0.371)}{(S+1.125)^3} \quad \text{for 0.3 cps BW} \quad (2)$$

and

$$\frac{\phi_0(s)}{\phi_1(s)} = \frac{0.00261(S+0.0371)}{(S+0.1125)^3} \quad \text{for 0.03 cps BW} \quad (3)$$

*Published as Goddard Space Flight Center Document X-531-64-9, January 1964.

†Instituto Geofísico Del Peru, Lima, Peru.

with a frequency response as plotted in Figure 2 and a time response as shown in Figure 3. The frequency response shows a peak of 3 db as compared with 4 db with the former system and a step input response with 25 percent overshoot as compared to 30 percent (Figure 3).

It has been mentioned in the report that the steady state error for this filter is zero for a step as well as for velocity input. If the input has some acceleration there is some error that should be corrected at the time of processing the information. In most cases this acceleration is so small that it can be neglected. Nevertheless for an input of the form

$$\phi_i = a + bt + ct^2 \quad (4)$$

the output would be in error by

$$\phi_i - \phi_0 = \frac{2C}{K_a} \quad (5)$$

where

$$K_a = s^2 G(s) \Big|_{s=0} = \frac{3.72(s+0.371)}{(s+3.34)} \Big|_{s=0} = 0.414$$

for 0.3 cps, and

$$K_a = 0.0372 \frac{(s+0.0371)}{(s+0.334)} \Big|_{s=0} = 0.00414$$

for 0.03 cps. The error constant is proportional to the square of the bandwidth; fortunately the acceleration is proportional to the square of the maximum phase rates. If the 0.3 cps is limited to phase rates of the order of 0.5 lobes/sec (0.05 lobes/sec for the 0.03 cps), corrections, if any, of only a few counts are necessary. (One count equals .001 of 360° electrical degrees or .36°).

The phase shifter in the new design has been improved for optimum linearity. The outputs of both rotor windings are being summed exactly 90° out of phase and equal amplitude by a high gain operational amplifier with unity feedback. It was noticed that the linearity of the phase shifter was affected by any slight distortion of the signal. Distortion has been kept as low as possible and the signal is being filtered from harmonics before phase detection by a 15 cps band pass filter. A measured linearity of 0.05 percent has been achieved.

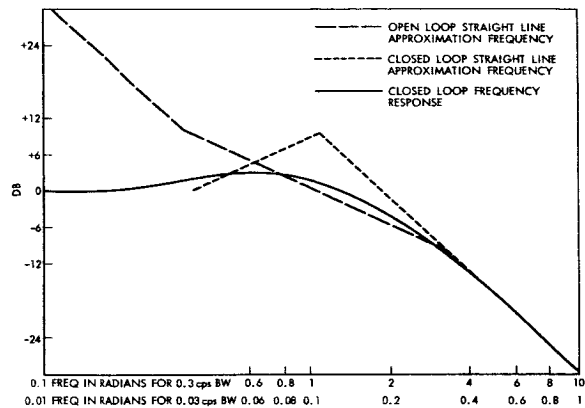


FIGURE 2.—Frequency response.

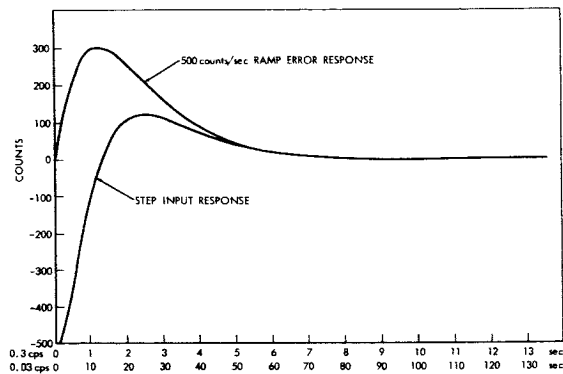


FIGURE 3.—Time response.

The AC operational amplifier was especially designed for high gain (90 db) at 100 cps and stable operation with feedback for unity gain. The phase shifter design is being used for other applications and will be described in a separate report.

A limiter stage has been introduced in the signal path before phase detection to get rid of amplitude variations. This limiting produces some noise quieting and generates a square wave for the phase detector.

The phase detector is the diode switching type. The signal square wave switches the 100 cps reference. The output of the detector after being filtered by the first operational amplifier is essentially the cross-correction of both inputs. The function of the DC Operational amplifiers are essentially the same. The values of some of

the elements have been changed to conform with the new transfer function.

The Chopper, Servo Amplifier and Motor-Generator is a Diehl integrated package. The Motor-Generator is a 60 cps, 3,000 rpm maximum speed type with 7.5 in-oz stalled torque. A 10 in-oz² flywheel was clamped on the motor shaft to increase its mechanical time constant. A larger mechanical to electrical time constant ratio allows stronger feedback in the velocity loop, thereby reducing the effects of mechanical friction and increasing its dynamic range. With the flywheel, it was found that the amount of feedback was not limited by the loop stability but by amplifier saturation with generator residual noise. Therefore, the flywheel inertia effect is not being fully realized. A flywheel inertia of 3 in-oz² should be sufficient.

A problem has been encountered which was not noticed before. For a particular phase input signal the system oscillates very slowly, as much as ± 2.5 count. This occurs only in the 0.03 cps. The oscillations occur for certain values of the fixed phase input which correspond to a particular position of the motor-generator rotor. If the system is tracking at a normal slow rate the problem does not develop since the rotor is never at the critical position long enough for the oscillations to build up. The only time, in practical use, that a fixed phase is fed into the system is during calibration. It is possible to adjust the phase shifter so that the calibration phase corresponds to a stable motor-generator rotor position, or calibrate in the 0.3 cps *BW* to avoid the oscillations.

The cause of the problem is the in-phase, position dependent, fundamental residual voltage of the motor-generator. The output of an ideal generator is zero for zero rpm. Actual generators have an output at 0 rpm which is called "residual voltage." The residual voltage is the sum of random noise and signals of the fundamental frequency and harmonics. The fundamental residual voltage has an "in-phase" and a "quadrature" component. It is only the in-phase component that causes any dynamic effects in the velocity servo loop. The in-phase residual voltage has a fixed and a position dependent component. The only effect of the fixed component is to cause "creeping" of the motor-

generator, and in a type 2 position servo loop is automatically biased out. The in-phase, position dependent, fundamental residual component varies sinusoidally as a function of the position of the rotor, with as many cycles as number of poles in the generator. It is this last component that causes the oscillations, being position dependent, and when fed back behaves as a position feedback either degenerative or regenerative, depending on what part of the sinusoid the rotor is positioned. In the 0.03 cps mode, this regenerative feed is of the same order as the desirable position feedback of the loop causing the system to be unstable and oscillate. The magnitude of the oscillations are as large as the negative slope portion of the sinusoid, i.e., one fourth of a turn for a two pole generator. In this system there is a 50 to 1 gear reduction from the motor-generator to the resolver, therefore, a one fourth turn peak to peak oscillation corresponds to 0.005 turns peak to peak oscillation at the resolver or ± 2.5 counts. If a 100:1 reduction were used the oscillations would be reduced to an acceptable ± 1.25 counts with a reduction in maximum tracking rate from 0.5 lobes/sec to 0.25 lobes/sec which still would take care of satellite passes with normal phase rates. If 400 cps were used, the oscillations would be reduced by a factor of 0.15 ($= 60/400$) for corresponding tracking rates.

In order to get rid of the oscillations completely, it has been calculated that a motor-generator with a maximum in-phase, position dependent, residual voltage, e_m , less than:

$$e_m \pm \left(0.003 \frac{v}{\pi p}\right) \text{ volts}$$

is necessary, where, v = output volts per 1000 rpm

$$p = \text{number of poles}$$

For a two pole unit and 6 $v/1000$ rpm

$$e_m = \frac{1}{2} \left(0.003 \frac{v}{\pi p}\right) = \frac{0.003 \times 6}{2 \times 3.14 \times 2} = 0.0014 \text{ volts}$$

would be necessary. Diehl Manufacturers state that they are in the development stage of a 60 cps unit that would meet such specifications.

Another solution for the problem is the use of a DC tachometer which will be tried in the laboratory. Increased friction and ripple on the output are new problems that would have to be solved.

The simplest solution and the one recommended for future units is to increase the gear reduction from 50 to 100 to 1.

A synchronous amplitude detector has been installed in the system with good results. The amplitude of the 100 cps signal input is synchronously detected using the output of the filter shifted 90° as the reference. The advantage of a synchronous detector over an envelope detector is that no noise is introduced in the process of detection and its threshold is determined by the

post-detection bandwidth rather than the pre-detection bandwidth. Post-detection bandwidth can be set to either 0.3 or 0.03 cps.

Another advantage of the detector is that it detects the amplitude of the 100 cps output of the Minitrack receivers. This amplitude depends only on the magnitude of any coherent signal or noise going into both front ends of the receiver and discriminates against any noncoherent noise such as the one generated in each front end or the one coming from the sky background.

The detected amplitude could be used to control the AGC of the receiver. A well balanced product detector would not produce any 100 cps ripple which would be a problem when a fast AGC is desirable.

VIII. ORBIT DETERMINATION

DIGITAL OHMMETER*

JOHN SEMYAN

This paper describes a unique method of digitizing resistive sensor outputs for telemetry. In the system proposed, the resistance to be measured comprises the unknown leg of a wheatstone bridge which is automatically nulled by sequentially adding the required binary values of resistance in series in the known leg. The states of the reed relay switches which control the binary balancing resistance then constitute a parallel digital word which represents unknown resistance. An absolute accuracy of ± 0.05 percent is easily achieved.

A design example that was used to prove the practicality and accuracy of the system is presented in detail. Some consideration is given to possible improvements in power consumption, range of resistance measurement, and operating speed.

INTRODUCTION

An ohmmeter is a device that measures resistance in ohms. The *digital ohmmeter* described herein is one whose readout is inherently digital, rather than analog. Basically, the digital ohmmeter is a self-balancing bridge circuit. It may be compared to a standard servo system in that an error voltage is compared with a reference, after which corrective action is taken. Through a series of successive approximations, the bridge is brought close to balance.

The advantages of automatically nulling a bridge using the system described herein are numerous compared with an electromechanical system which uses servo motors, precision gear boxes, and precision potentiometers. The latter approach, common in high-accuracy ground systems for many years, has the following disadvantages:

1. Lubrication is required for many moving parts
2. Controlled environment is desired for efficient, long life operation
3. Precision required in manufacture and assembly of mechanical components is costly and time consuming
4. The final package is comparatively bulky and heavy
5. Power requirements are relatively high
6. An analog-to-digital converter is needed to supply data to digital telemetering systems
7. System redundancy is impractical

The digital ohmmeter virtually eliminates most of the above-mentioned disadvantages. The only moving parts in the proposed system are the reeds in hermetically sealed reed relays of proven reliability. Final packages may be small in size and weight, and consume little power. Direct digital readout is an inherent feature. Dynamic range is easily controlled, and accuracy of readout is readily predictable. Simple switching circuitry results in a long operating life.

The system design example shown in this paper demonstrates an absolute accuracy capability of better than 0.05 percent, the primary limitation being the number of bits provided for readout. Higher orders of accuracy may be obtained by providing a greater number of bits and altering the bridge ratio.

The main limitation of the proposed system is its relatively low encoding speed. In the conservative system described, about 2 seconds is required to null the bridge and generate a single 8-bit resistance reading. Another limitation is the range of resistance which can be accommodated. Relay contact resistance can produce an error when very small resistances are switched, and maximum resistance is limited by stray capacitance effects. These limits depend on desired readout accuracy and dynamic range.

DESCRIPTION OF SYSTEM OPERATION

As stated previously, the method employed here in measuring an unknown resistance is to digitally null a bridge automatically. This is

*Published as *NASA Technical Note D-2003*, October 1964.

done by the successive approximations method, employing a binary adjustable resistance in the known leg of the bridge. Each resistor comprising this step-adjustable resistance is sequentially sampled and controlled, from the largest to the smallest binary value. When all have been sampled, the bridge is in a balanced condition within an easily defined percentage of error.

Reed relays are used as switches across each resistor in the known leg of the bridge. Before the sequential sampling process begins, a reset circuit *opens* all switches so that the binary adjustable resistor assumes a maximum value. This condition produces an error signal out of the bridge to an amplifier, and thence to a phase detector. Serial pulses, derived from a clock signal, are used to sequentially close the switches across the binary resistors. During the time that each binary resistor is being tried, the error signal out of the bridge is amplified and compared in phase with the bridge-drive reference signal. The output of the phase detector is then applied to circuitry that controls the condition of the respective switch. After the last—and smallest—binary resistor is sampled, the status of the switches constitutes a parallel digital word which represents the value of unknown resistance.

DESIGN EXAMPLE

Preliminary Considerations

The following design example illustrates the practicality of the concept. Before detailed design can proceed, a group of somewhat arbitrary design goals must be specified:

Unknown resistance range.....900 to 1000 ohms
 Allowable power dissipation in
 unknown resistance.....1 milliwatt max.
 Number of bits per data word.....8
 Encoding speed.....approx. 2 sec (4 cps clock)
 Available supply voltages.....±6 volts d.c.

An 8-bit binary word provides 2^8-1 , or 255, discrete coding levels, corresponding to about 0.4 percent of range uncertainty. The 10 percent measurement range of unknown resistance therefore should permit an absolute measuring accuracy approaching ±0.02 percent.

The long encoding time was chosen for system demonstration, since the logic operation is easily followed and visual readout is possible while the

coder is cycling continuously. A 400 cps drive for the bridge was selected. This frequency must be considerably higher than the clock rate to allow effective carrier filtering at the phase detector output, but not so high that stray capacitance will be troublesome. A square waveform was chosen to permit use of a simple logical AND gate as a phase detector.

The maximum permissible power dissipation in the unknown resistance is specified because of self-heating effects which would have to be considered in a typical sensor application. A platinum wire temperature sensor, for instance, should have negligible self-heating to minimize error. Since 1 milliwatt has been specified as the maximum permissible power dissipating in the 1000 ohm sensor, the square wave driving voltage appearing across the sensor must not exceed the amplitude shown in the following calculation:

$$P = \frac{E_s^2}{R_s}$$

where

P = power dissipation,
 E_s = sensor voltage (rms),
 R_s = sensor resistance;

or

$$E_s = \sqrt{P R_s}$$

For 1 milliwatt power dissipation and 1 kilohm resistor resistance,

$$E_s = \sqrt{(10^{-3}) (10^3)} = 1 \text{ volt rms.}$$

Figure 1 shows a block diagram of the complete system.

Bridge Drive Circuits

The free-running bridge drive multivibrator used in the breadboard model (Figure 2) operates between ±6 volts dc at a frequency of approximately 400 cps. The output waveform is symmetrical to within 5 percent.

Perhaps the most obvious bridge drive configuration uses a balanced transformer drive [Figure 3 (a)] and produces a single-ended output. The transformer is undesirable in this application

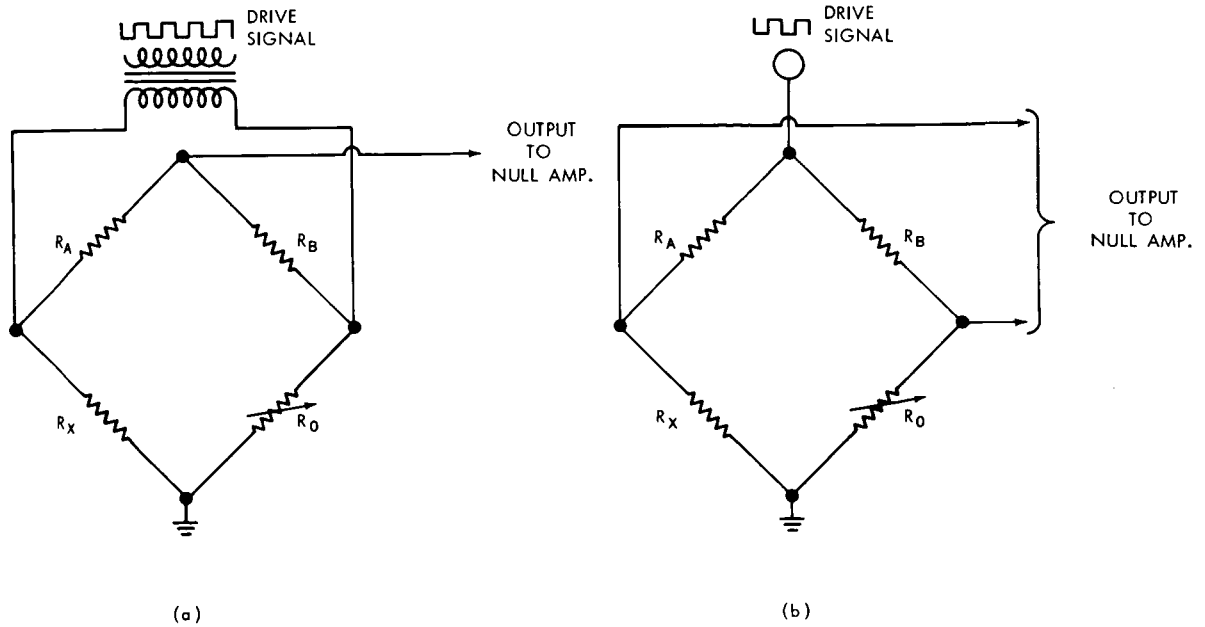


FIGURE 3.—Two bridge drive configurations (a) Bridge connected for balanced drive and single-ended output. (b) Bridge connected for single-ended drive and balanced output.

TABLE 1.

Binary Values of Resistance (N=8 bits)	Binary Resistor Value (ohms)	Equivalent R in Sensor (ohms)
2 ⁰	2	0.4
2 ¹	4	0.8
2 ²	8	1.6
2 ³	16	3.2
2 ⁴	32	6.4
2 ⁵	64	12.8
2 ⁶	128	25.6
2 ⁷	256	51.2

an average power dissipation of less than 1 milliwatt in the sensor for the 5:1 drive ratio selected. The differential signal out of the bridge

produced by a 0.4 ohm change in the sensor resistance, or a 2 ohm change in the X5 leg, may be calculated approximately as follows:

$$V_{out} = V_{in} \frac{\Delta R_0}{R_0 + R_B} = 7 \frac{2}{(5 \times 10^3) + (25 \times 10^3)} = 0.47 \text{ Mv P - P.}$$

where

$$V_{in} = 7vP - P \text{ (bridge driving voltage),}$$

$$R_0 = 5 \times 10^3 \text{ ohms (known leg of bridge),}$$

$$R_B = 25 \times 10^3 \text{ ohms [see Figure 3 (b)].}$$

Thus, the error signal out of the bridge could range from 0 to a maximum of 0.47 Mv peak-to-peak when the 2⁰ binary resistor is sampled.

Error Amplifier

The nominal gain in the error amplifier should be about ten times greater than that required to operate the phase detector from a 0.47 Mv input, to insure the validity of the least significant bit. This allowance also provides a margin for safe operation with fairly large, environmentally induced gain variations.

The minimum voltage gain required in the error amplifier may be computed as follows, allowing a

factor of 10 excess of gain over that required to just operate the least significant digit:

$$A = \frac{KV_0}{V_{in}} = \frac{10 \times 1.2}{0.47 \times 10^{-3}} = 26,000,$$

where

V₀ = 1.2 v, the peak-to-peak voltage necessary to saturate the AND-gate switch in the synchronous demodulator circuit;

V_{in} = 0.47 Mv as calculated for the smallest incremental bridge output;

K = 10, a nominal gain margin as explained above;

A = minimum gain required.

A high common-mode rejection (CMR) ratio in the error amplifier is necessary because the common-mode signal output would have the same effect on the synchronous demodulator as a

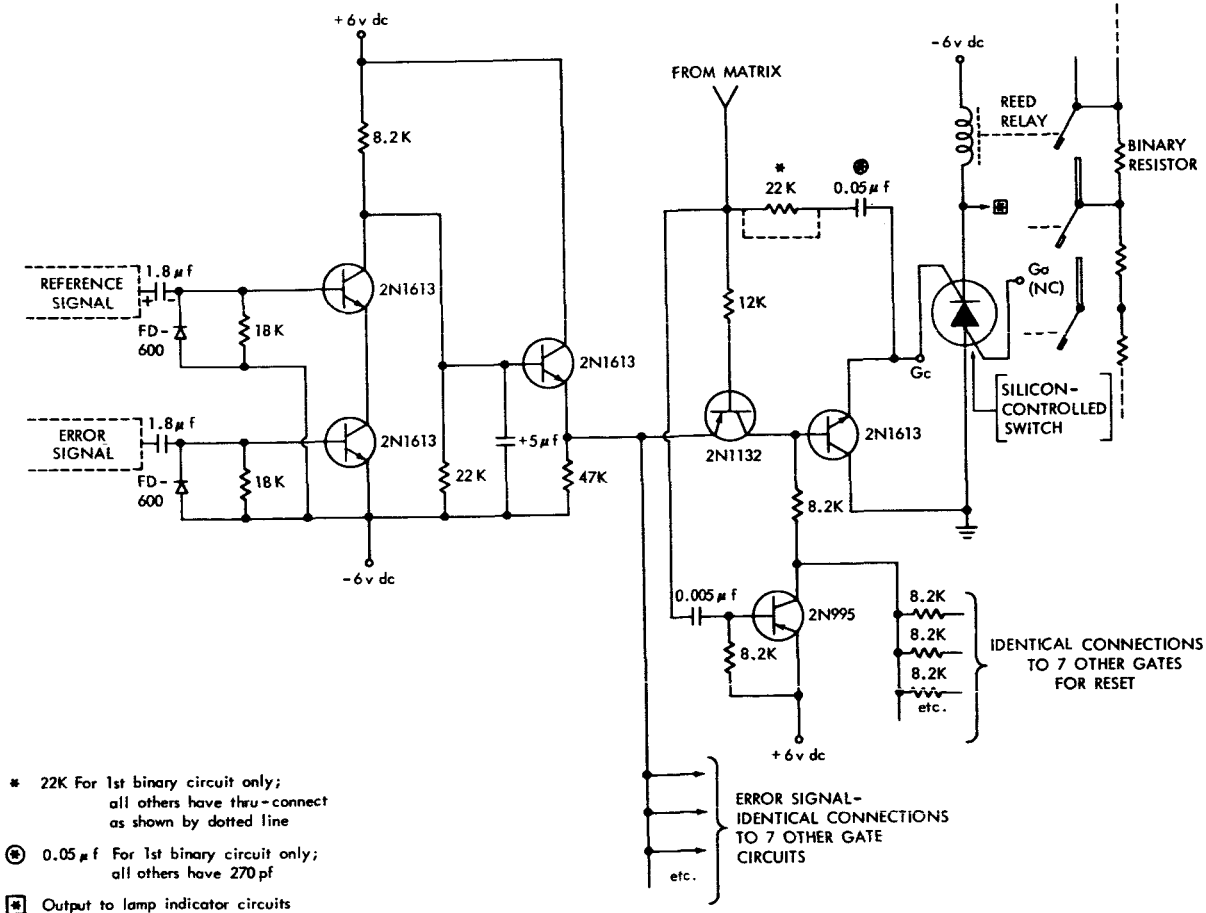


FIGURE 4.—Synchronous demodulator, information gate, and silicon-controlled switch.

differential error signal from the bridge. To insure a negligible error contribution, the equivalent common-mode input signal should be a factor of 10 below the differential input signal equivalent to a 1-digit error. Thus, for an amplifier gain of 26,000, the CMR ratio required would be

$$\frac{1.2 \times 26,000 \times 10}{1.2}$$

or about 108 db. Measured CMR ratio and gain for the amplifier shown in Figure 2 are 102 db and 30,000 respectively. These are considered entirely adequate for system test purposes.

Since a square wave signal is being amplified, the effect of phase shift—within limits—can be negligible. If operating frequency were to be changed greatly, then the frequency characteristics of the amplifier would require alteration. Frequency response of the amplifier design, as shown in Figure 2, is about 20 cps to 6.5 kc.

Synchronous Demodulator

The output of the error amplifier is capacitor-coupled to the lower transistor in the synchronous demodulator. The reference signal from the astable multivibrator is capacitor-coupled to the upper transistor. Refer to Figure 4. When both signals are *in phase* with each other, a negative output appears at the collector of the upper transistor. If the reference and error signals are *out of phase*, no negative voltage is produced at the output of the demodulator. A voltage divider adjusts the dc voltage level at the collector of the upper transistor to approximately +1.8 volts dc when no output signal is present. When error and reference signals are in phase, the filter circuit produces a negative dc output voltage that maximizes at approximately -5 volts dc. The output of the synchronous demodulator is fed through an emitter follower to the eight gate inputs.

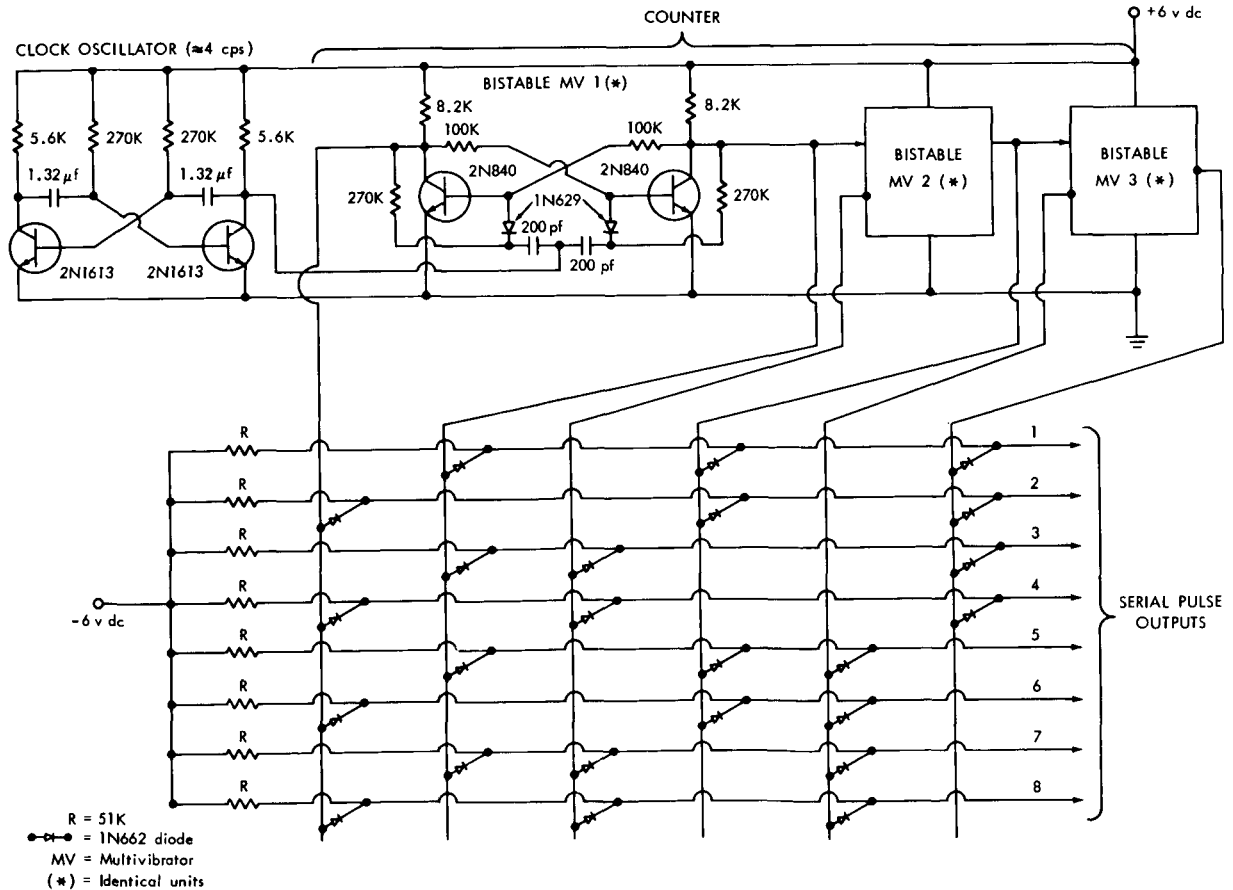


FIGURE 5.—Clock, counter, and matrix circuits.

Clock Oscillator and Serial Pulse Generator

In the present system, an astable multivibrator or clock oscillator having an operating frequency of approximately 4 cps drives a counter as shown in Figure 5. The counter consists of three cascaded bistable multivibrators, which count down to a final output frequency of approximately 0.5 cps. The counter outputs connect to a twenty-four diode AND matrix. The diode cathodes are connected in a binary configuration to eight 51 kilohm resistors. At this point of connection the matrix supplies eight serial outputs that are used as sequential gate trigger signals. The 51 kilohm resistors are connected to a -6 volt source. Because of the loading effects of the gate, the swing of the matrix output pulses is approximately +2 volts to -0.2 volt. Figure 6 shows a timing diagram of the system.

Control Circuits

The matrix outputs are connected to the bases of the eight transistors; also, they are fed through a differentiating network to the cathode gates of

the silicon-controlled switches. The matrix output to the largest binary digit control circuit also is fed through a differentiating network to a reset circuit that is common to all eight binary digit control circuits. The reset pulse places all the silicon-controlled switches in an *off* condition; hence all binary resistors are series-connected in the bridge circuit (see Figure 4).

The same pulse that the reset circuit receives from the matrix is applied to another differentiating circuit, with a longer time constant than that employed to gate the reset circuit *on*. This negative-going pulse switches the silicon-controlled switch to an *on* condition so that the reed relay switch is closed, and the largest binary resistor is shorted out of the bridge circuit. This sampling of the largest binary resistor produces an error signal out of the bridge circuit that has a phase relation with the reference signal. If a positive signal is applied to the emitter of the gate transistor, the negative matrix output pulse turns the gate transistor *on* and applies the positive error signal to the base of a switching transistor which virtually shorts the cathode gate of the silicon-controlled switch to ground. Thus, the silicon-controlled switch is opened and the largest binary resistor is again functional in the bridge circuit. If the error signal is negative, the gate never reopens and the largest binary resistor remains shorted out of the bridge. This process is then repeated for the remaining 7 binary digits.

The reset circuit is actuated only at the start of the sampling pulse for the largest binary resistor. The other control circuit operations take place during the sampling period for each binary resistor.

EXPERIMENTAL EVALUATION

The unit described herein performed as expected. A decade resistance box with 0.1 ohm increments of resistance was used as the unknown. A lamp indicator system, shown in Appendix A, was used as an output display. A plot of binary numbers versus resistance proved system performance and accuracy. Absolute accuracy was better than ± 0.1 percent, the tolerance of the bridge resistors used.

The 4 cps clock frequency used in the example was experimentally increased to about 20 cps, without appreciable change to other parts of the

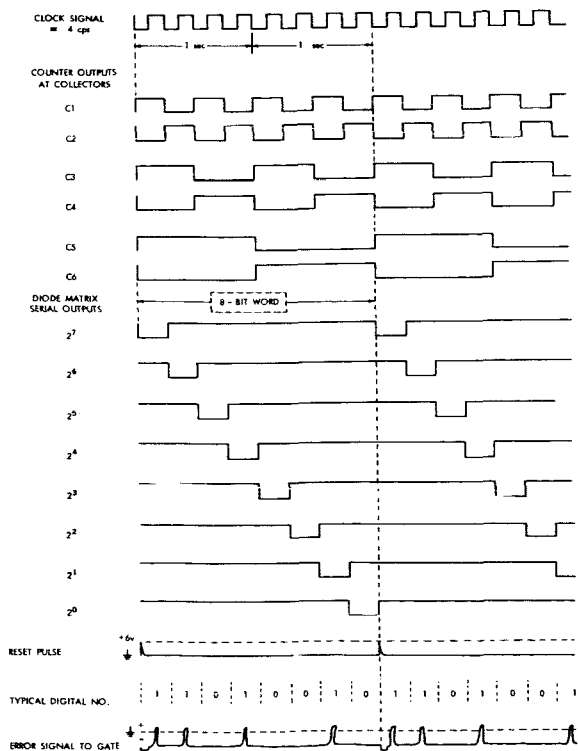


FIGURE 6.—Timing diagram for digital ohmmeter.

system. Increasing bridge drive frequency should allow further increase in coding speed.

A ring counter circuit employing modular bistable elements was tried in place of the counter and diode matrix circuits. It operated very efficiently and would be preferred in a reliable flight package. A ninth serial pulse is easily provided for more reliable "reset" action and for a "transfer" pulse to a telemetry system or word buffer.

It is expected that a shunt-type gating system for the error signal circuit would permit a greater range of parameter variation and possibly simplify the control circuit. Also, the silicon-controlled switch might be replaced by a more easily controlled level-sensitive bistable circuit.

CONCLUDING REMARKS

No great amount of circuit refinement effort has gone into the development of this system.

The design example presented here proves feasibility only. It readily can be seen that the demonstrated method is valid and that it may be adapted to many systems concerned with resistance measurement. The direct digital output makes the unit quite compatible with modern digital telemetry systems. It may be concluded that proper application of the basic scheme can result in a system of outstanding accuracy, coupled with high efficiency and inherent reliability.

ACKNOWLEDGMENTS

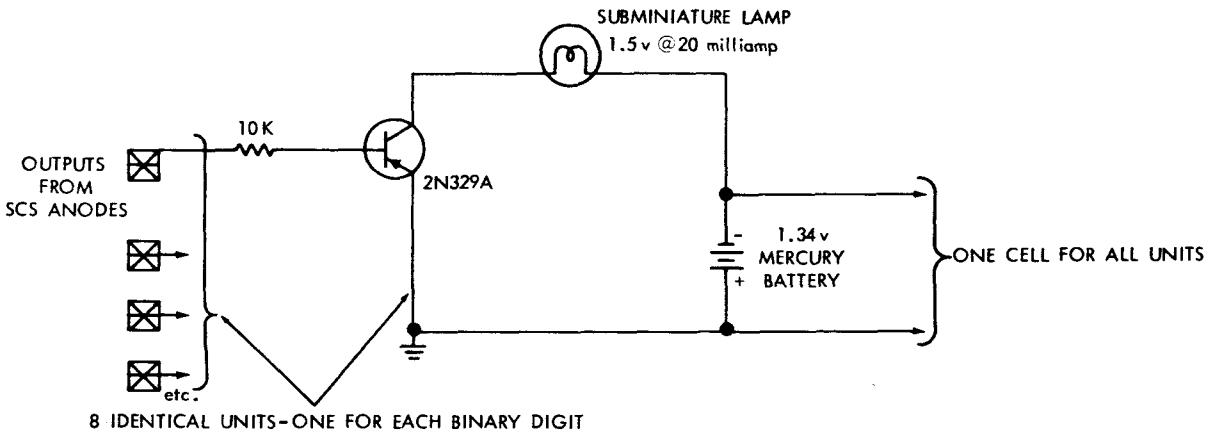
The author wishes to cite Mr. Marvin S. Maxwell, Goddard Space Flight Center, as the originator of the concept for the digital ohmmeter. Many thanks are also due Mr. Maxwell G. Strange of GSFC for his technical advice in the development of the demonstration model.

APPENDIX A

LAMP INDICATOR SYSTEM FOR BENCH CHECK OF DIGITAL OHMMETER

To facilitate bench checkout and demonstration of the system, miniature lamps were used to indicate the state of each bit. These are driven from the reed relay coils through separate

amplifiers as shown in Figure A-1. A lighted lamp indicates a binary "1" and an *off* lamp, a "0". Actual unknown resistance is easily calculated as shown for the sample readout.



	2 ⁷	2 ⁶	2 ⁵	2 ⁴	2 ³	2 ²	2 ¹	2 ⁰	
	256	+ 128	+ 64	+ 32	+ 16	+ 8	+ 4	+ 2	→ ACTUAL VALUE OF BINARY R's
	DIVIDE BY 5								
	51.2	+ 25.6	+ 12.8	+ 6.4	+ 3.2	+ 1.6	+ 0.8	+ 0.4	→ Equivalent FOR SENSOR
	IF LIGHTS ARE AS SHOWN ABOVE								
	0	1	0	1	0	0	1	0	
ADD	0	+ 25.6	+ 0	+ 6.4	+ 0	+ 0	+ 0.8	+ 0	+ 898 = 930.8 ohms

·FIGURE A-1.—Lamp indicator circuit for bench checkout system.

LOCAL ERROR CONTROL IN NUMERICAL INTEGRATION THROUGH OPTIMIZING THE ORDER OF THE INTEGRATOR*

W. S. SOAR

THE NATURE OF NUMERICAL INTEGRATION

Numerical integration consists of replacing the function to be integrated by another of simpler form, usually a polynomial, which is then integrated by some standard formula. This is true whether the integrand is a function of the independent variable only, as in the case

$$y' = f(x),$$

or is a function of both independent and dependent variables, as in the differential equation

$$y' = f(x, y).$$

The first case is the problem of the integral calculus, and is solved by numerical quadrature. Here, for a given x we can evaluate y' , and can substitute values of y' , or finite differences of y' , into appropriate formulas to obtain the integral y over some domain.

In the second case we require knowledge of the integral y for a given x before we can evaluate y' . This is like having to know the answer before we can ask the question. A common approach to this problem is to estimate y at the specified x , substitute it into the differential equation to obtain an estimate of y' , after which y is computed as in ordinary quadrature. This procedure is called the predictor and corrector method, of which numerous forms exist. If the first computed y is not good enough, it may be substituted into the differential equation exactly as was the original estimate of y . An improved y' is then found, which produces a better y .

*Published as *Goddard Space Flight Center Document X-547-64-295*, October 1964.

We shall assume throughout that y' is approximated by a polynomial. Then the success of the numerical integration depends upon how well y' is represented by the polynomial which replaces it, for the formulas integrate the polynomial exactly.

HAZARDS OF POLYNOMIAL APPROXIMATION

It is well known that polynomials, under appropriate circumstances, can be made to approximate almost any function to arbitrary accuracy. The Weierstrass approximation theorem assures us that if $f(x)$ is a real function continuous on the closed interval $[a, b]$, then for every positive number ϵ there exists a polynomial $P(x)$ of degree n ,

$$P(x) \equiv \sum_{k=0}^n a_k x^k,$$

such that $|f(x) - P(x)| < \epsilon$ for all x in $[a, b]$.

But this theorem is widely misinterpreted. To begin with, the fact that the desired polynomial exists does not imply that we can find it. That does not suggest we lack skills or methods. We may be unable to find the polynomial of arbitrary accuracy simply because we cannot get enough information about it.

Of one thing we may be sure. The existence of the desired polynomial does not mean it is of low degree. But here is something often misunderstood: A polynomial of higher degree does not necessarily give a better approximation.

Consider the example

$$f(x) = \frac{1}{1+x^2}.$$

Values of $f(x)$ are given for equal intervals in x .

TABLE 1

x -----	-3	-2	-1	0	1	2	3	4
$f(x)$ ----	.1	.2	.5	1	.5	.2	.1	.05882 352

We now interpolate among these points to find the functional value $f(.5)$, which is the same thing as saying we approximate the function by a polynomial which we evaluate for $x=.5$. Since $f(x)$ is of simple form, we can evaluate $f(.5)$ and know the correct answer in advance.

$$f(.5) = \frac{1}{1 + (.5)^2} = .8$$

To represent $f(x)$ by a polynomial of degree n we pass the polynomial exactly through $n+1$ known values of $f(x)$. The Weierstrass theorem assures that a polynomial exists which approximates $f(x)$ uniformly over the domain in question, but let us be realistic and compute our polynomial from information obtained as close as possible to the one value we are interested in. Thus, in the table below, when $n+1$ is even the points are symmetric with respect to $x=.5$; when $n+1$ is odd one extra point is taken before $x=.5$.

Consider the table, then, which gives the values of $P(.5)$ where $P(x)$ is a polynomial of the degree specified.

TABLE 2

Degree n of Polynomial	Placement of Points		$P(.5)$	See Text
	Before	After		
1-----	1	1	.75	.75
2-----	2	1	.875	.80961 540
3-----	2	2	.8	.80192 309
4-----	3	2	.85625	.80018 986
5-----	3	3	.81406 25	.80008 631
6-----	4	3	.85015 625	.80000 002
7-----	4	4	.82136 030	.80000 156

Contrary to common belief, the attempt to approximate $f(x)$ by polynomials of higher degree leads to failure. Why do not those of degree

greater than 3 do as well as the cubic? Does this example disprove Weierstrass?

The difficulty, aside from the fact that a rational function is not easily fitted by a polynomial, is that no attention was given to the interval $[a, b]$ over which the function was to be approximated. For the polynomial of degree one the interval used was $[0, 1]$; for degree 2 the interval was $[-1, 1]$; for degree 7 the interval was $[-3, 4]$. A final column is added to the table showing interpolated values of $f(.5)$ when all information consists of $n+1$ points equally spaced within the interval $[0, 1]$.

It is clear, then, that increasing the degree of the approximating polynomial does not always guarantee a better approximation. It should be clear also that changing to an integration formula of higher order does not necessarily assure a more accurate integral.

CONTROLS IN NUMERICAL INTEGRATION

If we seek to increase the accuracy of the integral

$$\int_a^b f(x)dx$$

by using a formula of higher order, then, in general, we require more information about $f(x)$ in $[a, b]$, which is to say the integration interval $\Delta x = h$ from one point to the next must be shortened.

And here we have two important controls to govern the accuracy of numerical integration:

1. We may select the order of the formula.
2. We may specify the step size h .

They are intimately related; and both are related to the third control:

3. We may decide the number of significant digits to be retained in the computation. (This last decision usually consists of directing the use of single or double precision mode of machine operation).

Very well, you may say, but how do we apply the controls?

Ordinarily when one undertakes a job of numerical integration his first choice is the integration formula. He can decide later what step size to use, and whether to run in single or double precision. But how does he pick the formula?

FROM INTERPOLATION TO INTEGRATION

First let us consider what an integration formula is and how we get it. Although it is not the only way, one common way to obtain an integration formula is to integrate formally an interpolation formula. The interpolation formula is usually an equation into which we may substitute either known values of the given function or their finite differences, together with an arbitrary value of the independent variable for which we wish to know the corresponding value of the function. The formula automatically fits a polynomial exactly to the given points and evaluates it at the arbitrary point without our ever knowing explicitly what the polynomial is.

If we did know the interpolating polynomial, we could integrate it term by term between selected limits and obtain an approximation to the definite integral of the given function between those limits. But by integrating the interpolation formula we obtain an integration formula which is general. Suppose we choose the Lagrange interpolation formula, which employs functional values directly: we get an integration formula in terms of functional values. Suppose we choose Stirling's formula, which employs central differences of the functional values: we get an integration formula in terms of central differences, as we shall see later.

In general the integration formulas expressed in differences are no better or worse than those expressed in ordinates. Often one form is readily derived from the other. For certain purposes, however, one may be better suited. For our immediate purpose of selecting the order of formula to do a specific job, we find an advantage in the formulas expressed in differences.

DERIVATION OF THE "COWELL'S" FORMULA

Suppose we are called upon to integrate the equations of motion of an artificial earth satellite, in rectangular coordinates. We have here a system of three second-order equations which must be integrated simultaneously. They are of the form

$$\ddot{x} = f(x, y, z; t),$$

with corresponding equations for y and z . The first derivative \dot{x} does not appear explicitly.

We digress briefly to consider the derivation of Taylor's formula with remainder, which we shall use in obtaining the integration formula.

Making all necessary assumptions about the continuity of $f(x)$ and the existence and continuity of its derivatives of all orders, we can construct the formula by integrating the n -th derivative n times from x_0 to x .

$$\int_{x_0}^x f^{(n)}(x) dx = f^{(n-1)}(x) - f^{(n-1)}(x_0)$$

$$\int_{x_0}^x \int_{x_0}^x f^{(n)}(x) dx dx = \int_{x_0}^x f^{(n-1)}(x) dx - \int_{x_0}^x f^{(n-1)}(x_0) dx$$

Since the integrand $f^{(n-1)}(x_0)$ is a constant, we have

$$\int_{x_0}^x \int_{x_0}^x f^{(n)}(x) dx dx = \left[f^{(n-2)}(x) \right]_{x_0}^x - (x-x_0)f^{(n-1)}(x_0)$$

$$= f^{(n-2)}(x) - f^{(n-2)}(x_0) - (x-x_0)f^{(n-1)}(x_0).$$

This equation can be rewritten

$$f^{(n-2)}(x) = f^{(n-2)}(x_0) + (x-x_0)f^{(n-1)}(x_0)$$

$$+ \int_{x_0}^x \int_{x_0}^x f^{(n)}(x) dx dx.$$

Although we have performed only two integrations, already the Taylor's formula has begun to take shape. Each additional integration will provide another term in the series. But this is far enough. Let $n=2$. The zero-th derivative of $f(x)$ is simply $f(x)$, and the last equation becomes

$$f(x) = f(x_0) + (x-x_0)f'(x_0) + \int_{x_0}^x \int_{x_0}^x f''(x) dx dx. \quad (1)$$

Using the notation $y = f(x)$, we evaluate $y_1 = f(x_1)$ and $y_{-1} = f(x_{-1})$ by means of (1).

$$y_1 = y_0 + hy'_0 + \int_{x_0}^{x_1} \int_{x_0}^x y'' dx dx$$

$$y_{-1} = y_0 - hy'_0 + \int_{x_0}^{x_{-1}} \int_{x_0}^x y'' dx dx$$

Add these equations and obtain

$$y_1 - 2y_0 + y_{-1} = \int_{x_0}^{x_1} \int_{x_0}^x y'' dx dx + \int_{x_0}^{x_{-1}} \int_{x_0}^x y'' dx dx. \quad (2)$$

It will be observed that the left side of this equation is $\delta^2 y_0$, the second central difference in Sheppard's notation. Further, the equation does not contain the first derivative, a significant fact since we are seeking a method for integrating second-order differential equations where the first derivative is absent.

To dispose of the two integrals on the right side of this equation we make use of Stirling's interpolation formula, which is expressed in central difference notation.

$$y_s = y_0 + \sum_{n=1}^{\infty} \left\{ \mu \delta^{2n-1} y_0 \cdot \frac{s(s^2-1^2)(s^2-2^2)\dots[s^2-(n-1)^2]}{(2n-1)!} + \delta^{2n} y_0 \cdot \frac{s^2(s^2-1^2)(s^2-2^2)\dots[s^2-(n-1)^2]}{(2n)!} \right\} \quad (3)$$

Here

$$\begin{aligned} \delta y_0 &= y_{\frac{1}{2}} - y_{-\frac{1}{2}} \\ \mu y_0 &= \frac{1}{2}(y_{\frac{1}{2}} + y_{-\frac{1}{2}}) \\ s &= \frac{x - x_0}{h} \end{aligned}$$

where h is the length of equal intervals in x .

It is clear that the factors

$$\frac{\mu \delta^{2n-1}}{(2n-1)!} y_0$$

and

$$\frac{\delta^{2n}}{(2n)!} y_0$$

in (3) are constants, so that the formula can be integrated with respect to the variable s . Let us, then, represent y'' in (2) by Stirling's interpolation formula, and integrate twice as (2) requires.

Stirling's formula is an infinite power series. If we expand it to some point and truncate the series, we have simply a polynomial, which we can integrate term by term.

Upon carrying out the integration and adding the two integrals as called for in (2), we find that all terms in $\mu \delta^{2n-1} y''_0$ vanish; and, after considerable arithmetic involving the addition of common

fractions, which are kept as common fractions for exactness, we finally obtain

$$\delta^2 y_n = h^2 \left(y''_n + \frac{1}{12} \delta^2 y''_n - \frac{1}{240} \delta^4 y''_n + \frac{31}{60480} \delta^6 y''_n - \frac{289}{36 \cdot 28800} \delta^8 y''_n + \dots \right). \quad (4)$$

This is the formula used by Cowell and Crommelin to integrate the orbit of Halley's comet. Nowadays it is most often used in a modified form with values of y'' instead of central differences of y'' , but for our purpose the difference form has special advantages, as was suggested earlier.

The first advantage is that (4) itself is an infinite series. How many terms appear depends upon how many were retained and integrated in Stirling's formula. But the important thing is that if we wish to use a formula of order higher than (4), we simply compute the coefficient of the next-higher term without changing the lower-order terms. We shall see later that it has another advantage.

To use (4) as written, we compute a central difference table to ninth differences $\delta^9 y''_n$. Extrapolate values of $\delta^9 y''$ beyond $\delta^9 y''_n$, assuming that $\delta^9 y''_{n+i} = \delta^9 y''_n$, and sum across the table until we have the necessary values to compute $\delta^2 y_n$ from the formula. Since

$$\delta^2 y_n = y_{n+1} - 2y_n + y_{n-1},$$

we have

$$y_{n+1} = \delta^2 y_n + 2y_n - y_{n-1}.$$

Here, then, is an estimate of the integral y at the $n+1$ -th step, from which we obtain an estimate of y''_{n+1} by substituting into the differential equation. With this y''_{n+1} compute the line of differences out to $\delta^9 y''_{n+1}$, which we had formerly estimated by extrapolation. With these corrected differences, use the formula to find a better $\delta^2 y_n$. Compute a better y_{n+1} . Iterate until y_{n+1} converges.

USE OF DIFFERENCE TABLE IN OPTIMIZING

Take a good look at (4). Suppose we have the indicated differences and a value for h . Then we can compute the contribution of each term, and can see readily whether the k th term affects the

last significant digit retained in the computation. Here all three controls come together. For a given h we may see that $h^2 a_k \delta^k y''_n$ is less than 10^{-3} and may be omitted in single precision operation, but is greater than 10^{-15} and must be included for double precision.

We can control the magnitude of $h^2 a_k \delta^k y''_n$ in two ways: by increasing k or by decreasing h . In (4) we see that coefficients a_k of successive terms grow smaller, and it may be shown that this trend would continue if more terms were computed. It is also true in general—although not always, and the exceptions constitute a dangerous trap*—that higher-order differences tend to zero. Hence, by taking more terms we make both of the last two factors, and their product, smaller.

Suppose, on the other hand, that we take a shorter step size h , say cut h in half. Immediately we multiply each term by one fourth because of the h^2 factor; but, beyond this, the differences also are reduced. First differences are approximately half as large as before; second differences, one fourth as large; n th differences, $\frac{1}{2}^n$ as large. Therefore, cutting h makes the first and third factors, and their product, smaller.

All this would be nice if we could obtain the difference table over the domain of integration. We could find an optimum combination of h and k , step size and order of formula, to squeeze out the last bit of significance at each integration step.

In the case of simple quadrature this actually can be done if the quadrature formula is expressed in differences. But for differential equations we are not likely ever to have a difference table of the actual quantities we wish to integrate. We obtain estimates of these quantities only by substituting into the differential equation the estimates of the integral at successive steps.

Of course even here we can look at the successive terms of the formula and see which actually are contributing to the solution, and how much. But this study can be made only after we have picked some formula, and some h , and some number of digits to carry.

In the problem we assigned ourselves earlier, that of integrating the equations of motion of an artificial earth satellite, we can do better. We

can form a difference table from values of acceleration based upon the two-body solution. Admittedly, the perturbed orbit is not the unperturbed orbit; but in most cases the greatest perturbing force, that due to the nonsphericity of the earth, is less than 10^{-3} times the central attracting force. By observing how the unperturbed values of y'' change with time we can get some notion of how the actual values change.

If we consider only the order of the integration formula, we probably should assume that a rapidly changing integrand will require a formula of higher order than a slowly changing one. It might be wise, then, to look at the table of differences in those parts of the orbit where accelerations are changing most rapidly as well as where they are changing slowly.

We can find the values of true anomaly corresponding to the maximum and minimum rates of change of acceleration as a function eccentricity. But this really is not much help because we are integrating not the total acceleration but its rectangular components, and where one component is changing rapidly another may be changing slowly. It would seem wise, therefore, to pick some likely value of h , say one minute, and compute the difference table for two-body accelerations from perigee to apogee.

Table 3A corresponds to a 20-step section in the orbit of an actual satellite, but without perturbations. The quantities labelled "Maximum in Block" are the largest among all the k th differences for the component specified. Thus for \ddot{x} there were 14 values of 6th differences, of which the largest was $.12600 \times 10^{-3}$. Here h is one minute of time, or .07436 492 canonical time units.

During the part of the orbit covered by this table the magnitude of the x -component of position reached a maximum of $1.09663 +$ canonical units of length. In single precision this is represented as $.10966389 \times 10^{+1}$; in double precision, as $.10966389829835 \times 10^{+1}$. The last three columns of the table are k th terms in the formula (4). Clearly terms in $\delta^6 \ddot{x}$ and higher are not significant in single precision operation: they could add nothing to x as represented above. But all terms listed are significant for double precision.

Perhaps a term of order higher than $\delta^{12} \ddot{x}$ might also be significant. Where do we stop? We have seen integration formulas which actually used the

*Form differences of the values in Table 1. The sixth difference is -7.2 ; the seventh, $11.8 +$.

TABLE 3A

$h = 1 \text{ min.} \cong .07436 \text{ 492 canonical time units}$
 $h^2 = .55301 \times 10^{-2}$

k	$ a_k $	Maximum in Block			$h^2 a_k \delta^k \bar{x}$	$h^2 a_k \delta^k \bar{y}$	$h^2 a_k \delta^k \bar{z}$
		$ \delta^k \bar{x} $	$ \delta^k \bar{y} $	$ \delta^k \bar{z} $			
2.....	$.83333 \times 10^{-1}$	$.16354 \times 10^{-1}$	$.10209 \times 10^{-1}$	$.47243 \times 10^{-2}$	$.75366 \times 10^{-5}$	$.47047 \times 10^{-5}$	$.21771 \times 10^{-5}$
4.....	$.41666 \times 10^{-2}$	$.71380 \times 10^{-3}$	$.98164 \times 10^{-3}$	$.51168 \times 10^{-3}$	$.16447 \times 10^{-7}$	$.22619 \times 10^{-7}$	$.11790 \times 10^{-7}$
6.....	$.51257 \times 10^{-3}$	$.12600 \times 10^{-3}$	$.14600 \times 10^{-3}$	$.10473 \times 10^{-3}$	$.35715 \times 10^{-9}$	$.41385 \times 10^{-9}$	$.29686 \times 10^{-9}$
8.....	$.79641 \times 10^{-4}$	$.44652 \times 10^{-4}$	$.22886 \times 10^{-4}$	$.10397 \times 10^{-4}$	$.19666 \times 10^{-10}$	$.10080 \times 10^{-10}$	$.45791 \times 10^{-11}$
10.....	$.13898 \times 10^{-4}$	$.67087 \times 10^{-5}$	$.12880 \times 10^{-4}$	$.66286 \times 10^{-5}$	$.51561 \times 10^{-12}$	$.98992 \times 10^{-12}$	$.50945 \times 10^{-12}$
12.....	$.26014 \times 10^{-5}$	$.56675 \times 10^{-5}$	$.22767 \times 10^{-5}$	$.18297 \times 10^{-5}$	$.81533 \times 10^{-13}$	$.32753 \times 10^{-13}$	$.26322 \times 10^{-13}$

TABLE 3B

$h = \frac{1}{2} \text{ min.} \cong .03718 \text{ 246 canonical time units}$
 $h^2 = .13825 \times 10^{-2}$

k	$ a_k $	Maximum in Block			$h^2 a_k \delta^k \bar{x}$	$h^2 a_k \delta^k \bar{y}$	$h^2 a_k \delta^k \bar{z}$
		$ \delta^k \bar{x} $	$ \delta^k \bar{y} $	$ \delta^k \bar{z} $			
2.....	$.83333 \times 10^{-1}$	$.90116 \times 10^{-3}$	$.12224 \times 10^{-2}$	$.93323 \times 10^{-3}$	$.10382 \times 10^{-6}$	$.14083 \times 10^{-6}$	$.10752 \times 10^{-6}$
4.....	$.41666 \times 10^{-2}$	$.17614 \times 10^{-4}$	$.96060 \times 10^{-5}$	$.31525 \times 10^{-5}$	$.10146 \times 10^{-9}$	$.55333 \times 10^{-10}$	$.18159 \times 10^{-10}$
6.....	$.51257 \times 10^{-3}$	$.23060 \times 10^{-6}$	$.25858 \times 10^{-6}$	$.25020 \times 10^{-6}$	$.16341 \times 10^{-12}$	$.18324 \times 10^{-12}$	$.17730 \times 10^{-12}$
8.....	$.79641 \times 10^{-4}$	$.15199 \times 10^{-7}$	$.10168 \times 10^{-7}$	$.34307 \times 10^{-8}$	$.16735 \times 10^{-14}$	$.11195 \times 10^{-14}$	$.37773 \times 10^{-15}$
10.....	$.13898 \times 10^{-4}$	$.66474 \times 10^{-9}$	$.59098 \times 10^{-9}$	$.54233 \times 10^{-9}$	$.12772 \times 10^{-16}$	$.11355 \times 10^{-16}$	$.10420 \times 10^{-16}$
12.....	$.26014 \times 10^{-5}$	$.44138 \times 10^{-10}$	$.61070 \times 10^{-10}$	$.22333 \times 10^{-10}$	$.15874 \times 10^{-18}$	$.21963 \times 10^{-18}$	$.80319 \times 10^{-19}$

equivalent of twelfth differences in (4). But at some point it would seem reasonable to cut the step size and use fewer differences, i.e., a formula of lower order.

Table 3B shows what happens in the same part of the orbit when h is cut to half a minute, or .03718 246 canonical time units. Already the terms in fourth differences are insignificant for single precision, while those in eighth differences barely affect the last digit in double precision.

It is unnecessary, or course, to form the products as given in Table 3. Instead, we can consider the first two factors only, and decide how large the k th difference must be for significance. In double precision, for example, if x is approximately one canonical unit of length, as above, the 15th significant digit is of the order of 10^{-14} , and the product of the three factors must be no less than

this quantity. Thus, for $h = 1$ minute, $\delta^{12} \bar{x}$ must be no less than $.27805 \times 10^{-5}$.

$$h^2 a_{12} \delta^{12} \bar{x} = (.13825 \times 10^{-2})(.26014 \times 10^{-5}) \delta^{12} \bar{x} \geq 10^{-14}$$

$$(.35964 \times 10^{-8}) \delta^{12} \bar{x} \geq 10^{-14}$$

$$\delta^{12} \bar{x} \geq .27805 \times 10^{-5}.$$

AN AUTOMATIC FORMULA SELECTOR

We have known all along that a computer can do anything we can tell it how to do. It seems not unreasonable to suppose we might tell it how to select its own integration formula in a particular instance.

Suppose we wish to integrate an orbit of such and such properties. The computer might integrate one revolution with a formula of high order and short step size. Then, by examining the differences of various orders, it could decide

that terms in the formula beyond a certain k add nothing to the solution. On the other hand, it might decide that because the k it selected is smaller than some predetermined value, the step size should be lengthened.

To be sure, the differences would not be differences of the actual accelerations we want to integrate, but they would be differences of the only accelerations we ever can integrate. And they would include the perturbing forces.

Having selected the order of the formula and the step size, the actual computation might then be carried out using an equivalent formula in terms of the accelerations directly instead of differences of accelerations, thereby avoiding further computation and storing of the difference table.

The orbital parameters which govern what order of formula should be used, and what step

size, are the semimajor axis a and the eccentricity e . It is reasonable to think that if some study is given to the results of the automatic formula selection, what formula goes with which a and e , we might then change the integration program so that it makes its selection on the basis of the initial input alone.

Whether it is desirable to change the formula or step size within one revolution of the satellite would certainly depend upon a and e . For any a , but $e=0$, there seems little need for changing either the formula or the step size. For large a and large e , however, it might be wise to change either formula or step size, or both, within one revolution. Again, since the rate of change of acceleration depends upon these two parameters, the computer could decide, on the basis of the initial conditions, whether to make changes and, if so, where.

NASCOM REAL-TIME SYSTEM GSFC 490 ELECTRONIC DATA SWITCHING CENTER INSTALLATION*

DALE L. SOLLARS

FOREWORD

In 1958 the National Aeronautics and Space Administration established a world-wide communications network which was used to successfully funnel the radar and telemetry data to the Goddard Space Flight Center computers, as well as the Cape Kennedy (Cape Canaveral) computers, to support the early space missions, especially the Project Mercury manned flight series. The network was based entirely on teletype speed transmission in the 60 to 100 word per minute range. By 1962 it was recognized that the missions were becoming more frequent, more complex and would involve high speed data sources outside the Continental United States which the slow speed teletype network could not service. Therefore, a modernization program was initiated with the intent of establishing a system of high speed solid state switches strategically located to provide the highest degree of automation; allow the use of a minimum of expensive long haul high speed circuits and give the network a real-time communication capability to stay abreast of the future projects including Gemini and Apollo.

The GSFC was the first switching center to undergo the modernization process. A Univac 490 Real-Time Switching System was chosen for the main switching functions and was installed during the month of December 1963. An extensive study of the network requirements had been in progress for approximately one year preceding the equipment installation. These requirements were interpreted into the programmable 490's and in March 1964 the system went into operation. Following is a description of the system as it is in present operation.

*Published as *Goddard Space Flight Center Document X-571-64-259*, August 1964.

ADDITIONAL CONTRIBUTORS

G. BROWN; H. CROWLEY; W. EIKER; E. PARKINSON; J. STRATTON and R. WETZEL.

The contributors listed above are employees of the NASA Communications Division and were closely associated with the project.

1.0 SYSTEM ARRANGEMENT

In order to accommodate the network requirements it was determined that the following equipment was best suited for the job (see Figure 1). It consisted of the following:

- 2—Univac 490 Real-Time Central Processors with 16K core
- 2—Univac FH 880 Magnetic Drums
- 6—Univac IIIC Tape Servos
- 1—Univac 1004 Card Processor

A system of input multiplexing of the data circuits was chosen because of the large number to be serviced. Each circuit enters the system through its individual line terminal which is a Univac CLT (one character buffer). Both the five level baudot code and the eight level ASC II codes are being used in the Center. When a circuit becomes busy and an input buffer has received a complete character, a series of interrupts are triggered in the logic of the 490 Central Processor. A multiplexer has been assigned an address in the main core of the 490 CP for that character and it is brought into the system. The CLT is then ready for the next character. The transfer is accomplished in a total of 20 microseconds which is much faster than the input circuits can operate. This rapid scan time of the multiplexer insures that all inputs will be serviced in a real-time mode, even when they are all busy at the same time.

At the present time, each multiplexer occupies one of the 14 channels into the Central Processor

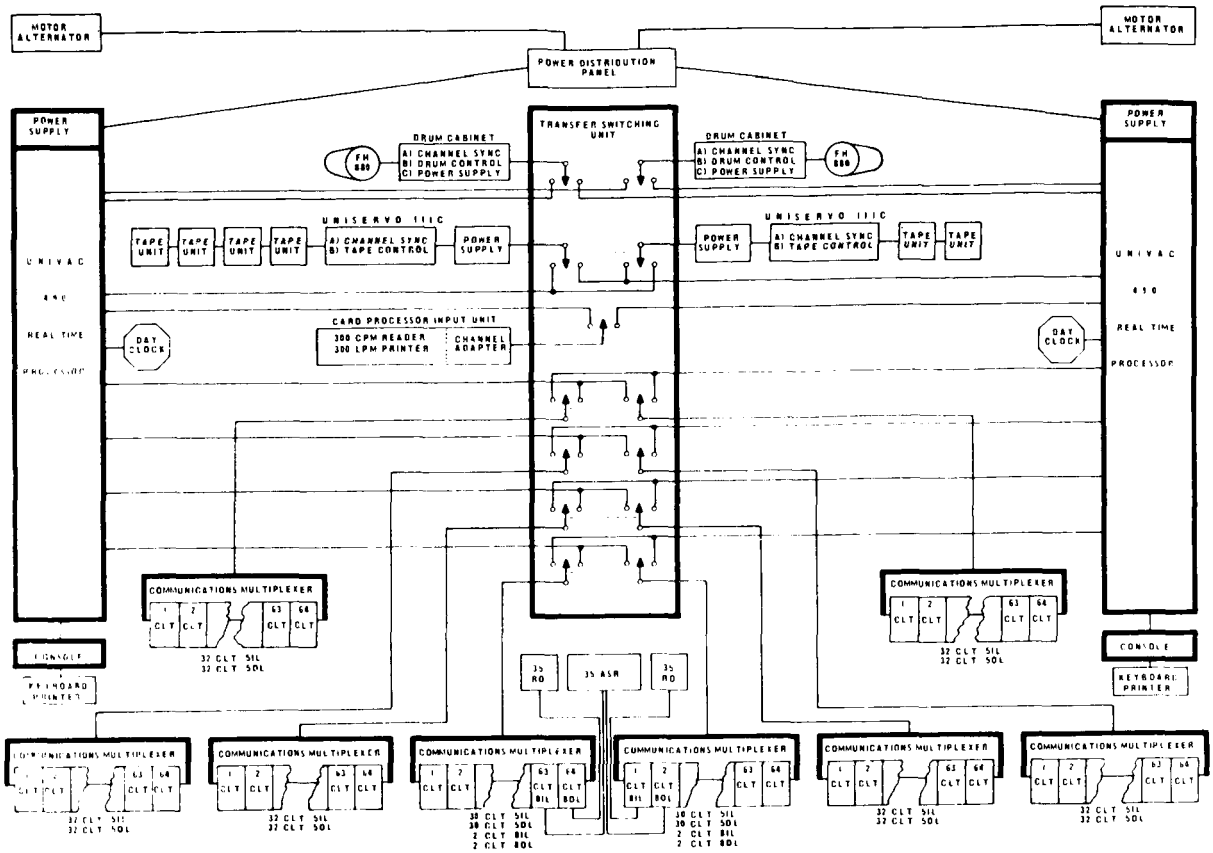


FIGURE 1.—Equipment arrangement.

core; therefore, four channels are used to service the four multiplexers. Each multiplexer has the capability to service 32 data inputs or circuits in a full duplexed mode for a total of 128 circuits. This is by no means the capacity of the system, as the multiplexers can be cascaded onto each other should the network grow to that extent. In a peak load condition with all circuits busy at 100 wpm speeds, each input would necessarily be serviced each 1000 micro-seconds. It requires 20 micro-seconds for each busy CLT or a total of 1000/20 or 50 busy circuits maximum for each channel. Assuming a logical 25 percent loading of the circuits at any particular time would allow $4 \times 50 = 200$ circuits per channel, which is more realistic when thinking in terms of total capacity.

As stated before, each CLT has been assigned its own small area of core in the Central Processor where incoming data is stored for that particular

circuit. This is one of the more important features of the 490 Central Processor for this application in that it requires no internal machine cycle time to service the inputs and outputs once the input or output mode has been established. The data is stored in its own location within the Central Processor under control of programmed hardware.

Once inside the CP the data is packed into five character words and transferred to the FH 880 drum in a location which is remembered for future use. When sufficient characters have been received to completely identify the sending and receiving addresses, the receiving circuit is status checked. If available, the message is sent in real-time as it comes in. If not available, a notation is made in the "things-to-do" table and at the first opportunity the message is sent and the notation is removed from the "things-to-do" table.

The data is not removed from the drum at this time as there are several uses for it in other programs. It must be logged, journalized, stored for future reference or recall, and so forth. In order to provide a system of indiscriminate sending and receiving messages into an on-line device, systematic layout of the drum was designed as follows: It was divided into five time frames of eight hours each. These time frames are somewhat related to five rotating sections on the drum, each section containing a complete eight hours of traffic. At the end of eight hours, of traffic on section 1, section 2 is used for the next eight hours; section 3 and 4 utilized next in turn. However, when section 4 is being filled, a comparative search is made of section 1. Any message found to have been sent is transferred to tape and erased from the drum. Any message which is found not to have been sent is transferred to section 5 of the drum and all references to the drum address of this message are changed to the new location wherever it might be stored in memory such as in the "task" table. This procedure insures that one section of the drum is always available for new traffic in advance of the filling of the last section to insure that a saturated condition cannot exist on the drum.

The output of data from the drum, through the Central Processor core, through the multiplexer and to the correct CLT is a reverse of the input cycle. The message is retrieved from the drum, unpacked from the 5 character words and placed in the correct output buffer for transmission to its destination. The output mode is established between the CP and the CLT and the message is sent, character by character to the receiving site.

A representative sample of the CLT and multiplexer is shown on Figure 2 and the equipment arrangement is shown on Figure 1.

2.0 SYSTEM FEATURES

The 490 system of traffic switching accepts inputs and outputs the same as other teletype networks, except that once the message is entered into the 490, it is read, interpreted, directed, corrected, if necessary, and sent to its destination automatically providing enough information is received to determine the addressee(s). If a trouble condition exists in a message, it is interpreted by the programs and a print-out which

states what is wrong with the message, is sent to the proper service equipment. These print-outs are called Advisory Messages which call for some action by the responsible person. The print-outs call attention to such things as:

- a. Invalid SOM (start of message)
- b. Invalid Precedence
- c. Invalid EOM (end of message)
- d. Message Sent
- e. Message Held (with drum address)
- f. Garbled Message

These features are all designed to help the operations personnel service the network. There are also many side benefits which are primarily of more interest to the engineering planning and analysis personnel. Many of the basic parameters for network engineering and analysis have been incorporated into these programs, such as:

- a. Message Count
- b. Word Count
- c. Busy Circuit
- d. Traffic Backlog
- e. Circuit Outage
- f. Circuit Assurance
- g. Network Utilization
- i. Network Loading

These parameters are built into the programs and can be recalled in any desirable form. They are mostly associated with the slow speed TTY circuits, but in the very near future the center will be switching high speed circuits. Different techniques are required for the necessary handling of these new circuits. They require different programs in order to identify various parts and pieces of messages, as well as different parameters which are covered by the data.

High speed circuits are to be used to carry traffic between the major NASCOM switching centers, and also to connect to the various computers at GSFC and outlying locations. The center will connect to three IBM 7094's in the installation, with a transfer rate of 40.8 kilobits per second. Each computer in the system presents a new and different set of inter-facing programming problems.

The 7094 interfacing programs are required to accept the regular NASCOP formats and perform the necessary message accounting. With the knowledge and familiarization of the 7094 functions we are progressing on these programs.

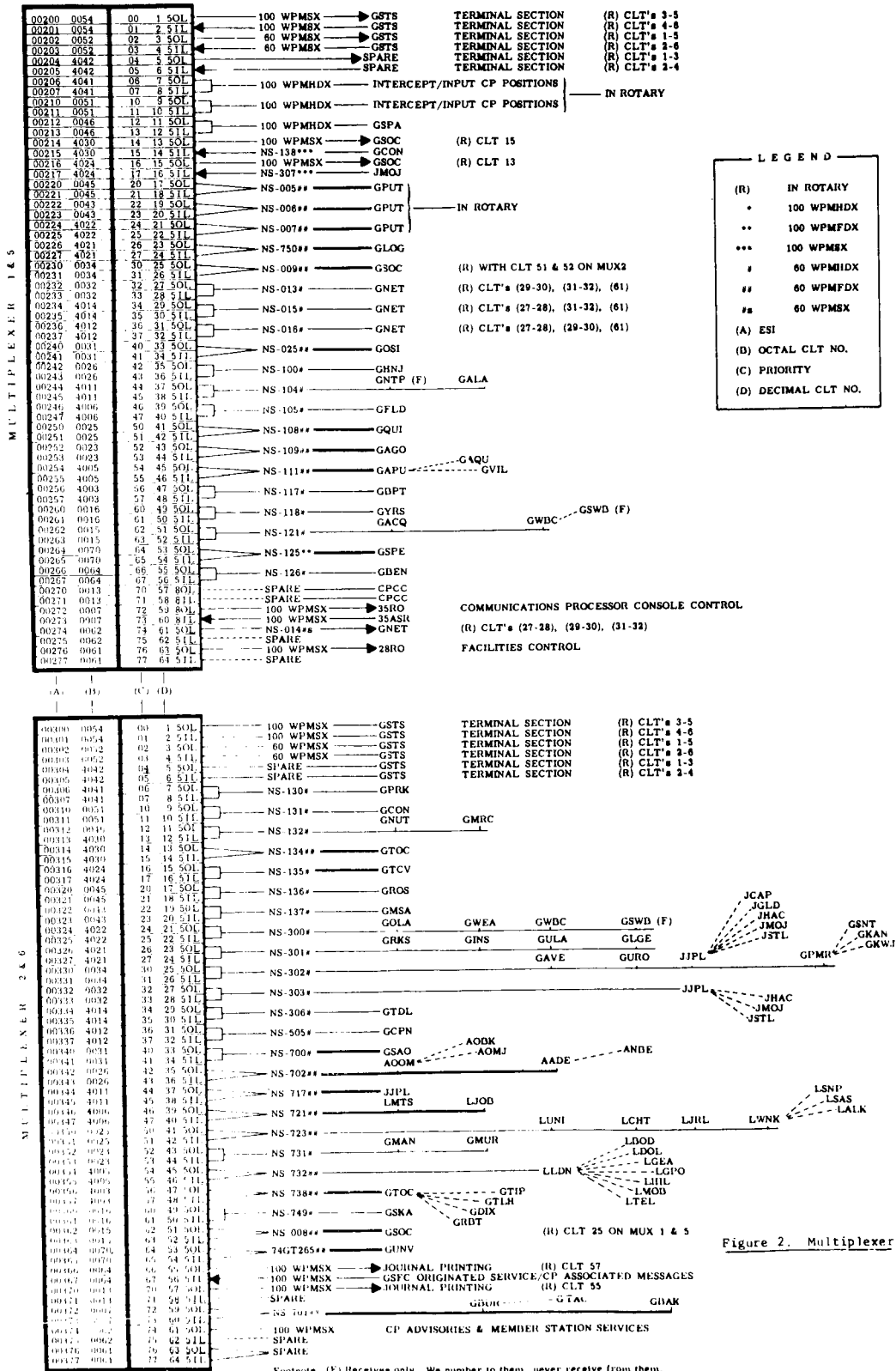


Figure 2. Multiplexer

FIGURE 2.—NASCOM CLT/circuit multiplexer assignments.

There are plans at GSFC for interfacing other computers which do not conform to the 7094 type of task. These are ultra-high-speed computers which will be used for data reduction associated with the various scientific satellites such as Nimbus, Tiros, Relay, SYNCOM, and so forth. The 490's will be required to store data for varying lengths of time and on pre-arranged schedules send this data to the using computers at very high speeds. New techniques are needed for these tasks as the requirements are finalized. New types of data handling programs are being investigated which will assure better reliability.

3.0 SYSTEM DESIGN

The basic premise for the application of the 490's to the NASCOM network was 100 percent reliability and a flexible input handling technique. The reliability is assured by using two 490's in the system, either of which can perform the complete switching job, and also execute the important programs in such a manner that only solid state circuitry is involved in the switching chain. Although a 1004 Card Processor and tape servos are a part of the overall 490 system, only the main frame logic and core storage plus an FH880 Magnetic Drum are used in the main switching chain. As long as one main frame and one drum is operable, the system can perform its basic job of switching the traffic. The system can remain in this status for extended periods of time under emergency conditions.

Flexibility of inputs has been achieved through an intensive programming effort coupled with a communications oriented processor. To achieve the most flexibility in the system, a study of other networks was necessary. In some cases involving the other networks studied it was found that strict adherence to pre-planned formats, message lengths, acceptance times, and so forth is required; otherwise, the messages are not accepted. In other cases, polling of input circuits determines when a message will be accepted. This imposes varying degrees of restrictions and regulations on the operators in the field and will not support a real time network as used in NASCOM.

The 490's in this system are programmed in such a manner that all inputs are accepted without restriction in real-time. There is a recognized format which is thoroughly covered in the NASCOM

operating manual. Any input will be accepted and forwarded by the 490 programs regardless of any non-conformance to the recommended format, but such non-conformance is noted on the intercept position. In some cases, insufficient information may exist to determine the addressee, therefore, making it impossible to forward the message. This situation has been given consideration, and a notation at the intercept position will state the drum address where the message is stored, enabling it to be recalled for any required action by the operator at GSFC.

4.0 WORKING PROGRAMS

Instructions are grouped together to form individual programs which perform their specific job. Some of these programs are long and complicated (validation), while others are short and quite simple (various advisories). Each has its place in the overall picture and is called upon to act when it has a job to do. For example, the advisory to inform the service area of an excessively long message (which could indicate a stuck tape or an open circuit) is only called upon when this situation arises. The input program, on the other hand, is involved with every character entering the 490. The assignment of the programs, when they are needed, is under the direct control of the Executive Program, "COMEX."

4.1 Comex

In order for the system to be in a position to perform at maximum efficiency in a real-time mode, an overall supervision of the individual jobs to be accomplished is required. There are a myriad of details involved in supporting a network such as NASCOM and each of them must be performed as soon as possible, however, only one job can be performed at a time (an exception to this is simultaneous input and output). Also, some jobs are more important than others. For example, it is much more important in relation to machine time to service a busy input than it is to write a message on tape for use some time in the future. The message is available for transcribing for many hours, but the input would be lost in, say 100 milliseconds, even less on a high speed input.

A parable to the comex control or executive system is the "job jar" which is present in many

households. A jar is kept in the house and as tasks present themselves, a notation on a slip of paper is placed in the jar. Whenever time presents itself, a task slip is taken from the jar and assigned to the person responsible. This process continues until the jar is empty. In fact, in an efficient arrangement, there should never be more than one task in the jar at any particular time. Experience shows, however, that this is not the case.

In the 490's, with the many tasks to be accomplished, approximately 50 individual worker programs have been written, each with its own specialty. One inputs the data, one outputs, one packs the characters, one stores to the drum, etc. Comex, as master of the "job jar" assigns the tasks to the responsible worker program. Upon completion of the tasks, the worker program erases the task from the task table and returns control of the system to Comex. Many times completion of the task involves parameters which are beyond the control of the worker program. For example, transmitting a multiple address message to all addresses at the same time requires that all of the circuits involved are idle at that time. Sometimes this is not the case. Therefore, the worker program must send the message to the addressees whose circuits are available and make a notation or "job-slip" for those remaining. The notation is placed on the 490 "task-table" (job-jar) with an associated priority. When time presents itself, and it does in the 490, the worker program checks the circuits again for availability. If the job is completed the task slip is erased. If not, it is left on the task table.

Comex always knows where a worker (program) is located. When Comex checks the task table, the task is assigned to the address which is associated with the task. In other words, reverting to the household job-jar, if the pipe is in need of repair in the basement, the task slip is sent to the basement. If the leak were in the kitchen, the slip would be sent there, etc. In the 490's, the responsible program is always at the correct location and prepared to act.

Comex devotes the majority of attention to inputs and the remaining time is spent on the myriad of details involved in the servicing of a message through the network. For the most part, the decisions made by the 490 are hidden

deep inside the various programs, but the results of the decisions are readily apparent. They are covered in this report under "System Features."

5.0 NEW PROGRAM IMPLEMENTATION

The most strict control possible is maintained to insure that any change to the programs does not adversely affect the operation programs. This control consists of verification, test, and simulation methods. When a change is made on any of the circuits, sites, or NASCOP procedures the detailed requirements are submitted to the programming group. The task is assigned to a specific programmer for implementation.

Upon completion of the coding, assembly and debugging, it is given to another programmer for his verification and concurrence. It is then submitted to the programming analyst to assure that the change will actually perform and will integrate into the total system. At that point, the change will be made on a simulated off-line basis in the back-up 490 system. It is run for the required period of time and the results are analyzed from an individual program basis and system basis to reassure that none of the important information storage locations have been inadvertently modified. When everyone concerned has agreed that the change is valid, it will be incorporated into the system programs for on-line operation. This system of documentation will insure that the most minor network change will not cause a series of events which would disrupt the operational programs.

Along with the actual work, a filing system is kept on magnetic tape. This file lists all the pertinent information concerning the individual circuits and also contains the actual configuration of the various modules which are used by the 490 to describe the individual circuit parameters. Its main use is to keep an up-to-date network file as it actually exists. When a circuit parameter changes, and the program is modified accordingly, this file is printed on the 1004 either in whole or in part. A complete network print-out may be taken at any time and varied as to particular needs. All modifications to the circuit parameters are made from this master file; when the modification is finally incorporated into the NASCOM system programs, the file is up-dated also. A system of verification by program, insures

that there are no errors in restoring the up-dated file back onto the tape. These verification programs are covered in this report under "Utility Routines."

6.0 CONSOLE OPERATION

Prime function of the control console is to activate and initiate a worker program held in memory. This is done by manual key or via a 35 ASR. The console has the capability to perform various functions, such as:

- a. Open and close stations
- b. Set stations operational or non-operational
- c. Set or clear alternate routing indicators
- d. Circuit status reports
- e. Reset channel numbers
- f. Initiate message recall
- g. Request specific table listings
- h. Initiate alternate routing
- i. Initiate normal and emergency transfer of control from one processor to the other
- j. Restart validation

A standard format must be observed when addressing the CP via the 35 console. All characters keyed into the CP will be disregarded until a start of message function (\$) is received. It is necessary to give information to activate a particular program following the SOM. The information must be in the form of 3 alfa character identification code which describes the function to be performed. This code may or may not be followed by elucidating parameters. Following these parameters is an optional 3 alfa character code which is significant only with particular entries. Arguments may be present and must follow in an alfa-numeric sequence (A-Z, 0-9) or will be rejected. Upon completion of all necessary entry information, an end of message (".") must be sent; then the CP will take action on the request presented.

Each entry will generate a disposition message. This message gives the current derived time from the day clock, the function identification code and a message acknowledgment which indicates the entry was made or an error response which states why the entry was rejected.

The console has the ability to backspace and delete the last character entered. Three consecutive backspaces will terminate the message being entered.

7.0 SERVICE AREA

The service section plays an important role in the operation of the Communications Processor 490 Real-Time System. Although messages are not held, according to a recent survey, 15 percent of all incoming traffic will go to the service area. The service area section consists of the following:

- a. Four teletype Model 28 RO's
- b. Four teletype Model 28 ASR's

Two Model 28 RO's are used for Journal Listings. There are two types of Journal Listing; normal and recall. These two RO's are on rotary; if one is busy you can use the other. If necessary, everything which is received on the RO's can be put on tape for future use. The information appears on the normal Journal Listing as input channel designator, input channel letter, input time word and date, output channel designator, output channel letter, output time word and date, the Julian date, the internal message number and the entire message heading to the first letter Z. The information appears on Recall Journal Listing as input channel designator, input channel letter, input time word and date, output channel designator, output channel letter, output time word and date, the date, internal message number, recall request message, supplement heading letters Z and original message heading letters Z.

The other two rotary RO's are used for Communications Processor generated Advisory Messages. Most of the messages received on these RO's will not require immediate action for they are primarily for information purposes and will not be numbered, journalized, or stored. There are two types of discrepancies and operational messages directed to non-operational stations. The advisory informs operating personnel that no other errors are involved in the message, and the message is forwarded to all addresses. The information as it appears on the RO's is; input channel and letter designator, input date-time word, invalid channel number, last valid number received, an operational message directed to a non-operational station, input message through the first double line feed.

Two Model 28 ASR's on rotary are used for intercept traffic. All messages on these machines are available for normal recall from the drum. They are also recorded on tape, but are not available for recall from the tape because no journal

entries are made for this traffic. All messages of the type which require action on the part of switching center personnel appear on these ASR's. All traffic directed to intercept by the validation routine is placed on a common queue for delivery to the intercept position. Some of the station discrepancies that cause the CP to generate a message to the Intercept ASR's are:

- a. Invalid channel number
- b. Invalid start of message
- c. Invalid precedence
- d. Invalid routing indicator
- e. Operational traffic to non-operational station
- f. Invalid end of message

The print-out on the ASR is; input channel and letter designator, input date-time word, and type of error followed by sent or held. When the message is forwarded with an error in the precedence or routing indicator, the station to which it is sent is listed following the word SENT. When the message is forwarded with an invalid start of message or operational to non-operational error the word SENT is present but not the station listing. When the error is an invalid channel number without last valid number received, the word HELD is present and no stations are listed, but the drum address where it is stored, and the input message through the first double line feed does appear.

If there is an invalid end of message on a line other than GSCP, the message is sent to the intercept ASR, the end of message (FIGS H LTRS) is inserted at the end of the message and action is taken for a valid ending.

Another Model 28 ASR is used for all traffic addressed to GSPA. It is available for normal recall from drum or tape.

The last Model 28 ASR is used for all transmitted traffic originated from GSPA; this includes intercept traffic corrections and is also available for normal recall from drum or tape.

Intercept corrections use VALIDATION, and FIX-IT routines. The VALIDATION routine validates the routing indicators, when it gains control from the FIX-IT routine, and lists the drum address of the supplementary heading on the proper output queues. The FIX-IT routine utilizes the supplementary heading and drum address (supplied in a message from GSPA to

GSCP) to forward previously intercepted messages to their proper addresses.

8.0 FACILITIES CONTROL INTERFACE

8.1 CP Access to Distant Stations

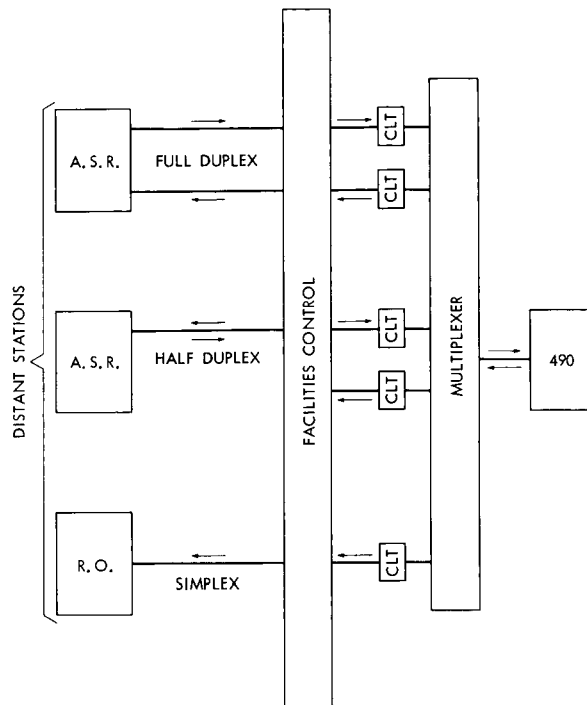


FIGURE 3.—Facilities control interface.

8.1.1 Multiplexing

Communications transactions are handled by the CP on a character multiplexing basis. Character multiplexing is defined as one character being sent to the CP from one communications line terminal (hereafter referred to as CLT) before control is sent to another active CLT for a one character transmission. Some CLT's being of a higher priority due to character speed, are sampled by the multiplexer more frequently than others. One communications multiplexer accommodates up to 64 CLT's, and another CLT accommodates one communications line in one direction. CLT requirements for various types of circuits are:

- a. Full duplex—2 CLT's . . .
1 send and 1 receive
- b. Half duplex—2 CLT's . . .
1 send and 1 receive
- c. Simplex—1 CLT . . . (either direction)

The switching speeds of the Multiplexer and the CP offset the possibility of lost characters.

8.2 Basic Circuitry

Each CLT termination "appears" in a set jack corresponding to its associated circuit at the Facilities Control Board. Each CLT is operated directly by the DC loop and associated keying.

A patch cord inserted into the CLT jack, "physically lifts" the CLT from the circuit; thus, disabling that particular leg of that particular circuit.

All CP associated teletype equipment also "appears" on the Facilities Control Board on jacks and are "hard wired" directly to the CLT's which service them.

8.3 Communications Interruptions

Generally, communications interruptions fall into the following two categories:

- a. In-house troubles
- b. Communications path failures

8.3.1 In-house troubles

Generally, in-house troubles are due to equipment failures. At the GSFC, these equipment failures fall into two categories. These categories are:

- a. CP associated equipment
- b. Teletype and associated equipment

Once it is determined that there is an apparent in-house trouble at GSFC, the Facilities Controller must test to determine the source of trouble. This requires close coordination between the Facilities Controller and the maintenance forces of the Western Union and the Telephone Company, as well as the CP engineer.

Certain equipment patching may be implemented to restore service to CP associated teletype equipment. This is done by patching a spare teletype machine into the CLT which normally services the affected teletype equipment. A defective CLT also may be patched to another currently unused CLT, pending restoration of the affected CLT. However, due to the physical wiring relationship of each circuit to its associated CLT, it is imperative that this change be implemented in both the CPCC/490 area and at the Facilities Control. All other CP associated equipment failures must be handled by the CPCC operator and the Univac engineer.

8.4 Circuit Assurance

Circuit assurance functions are conducted by the CP which are very similar to those functions now performed manually. The CP, however, queries stations on an incessant basis rather than at random.

Each station that is equipped with circuit assurance answer-back equipment is tested and if the resulting reply is incorrect, or if there is no reply at all, the CP continues to test this station until it gets a good reply or until it queries the station five times without a good reply. At the time the circuit is tested five times without receipt of a good reply, the circuit assurance program within the CP considers this circuit to be out of service. The CP then takes the appropriate action to insure that no further traffic is sent to this station and alerts the Facilities Control with a message stating the following:

- a. Time
- b. Station
- c. "CA Report"
- d. Number of "no replies"
- e. Number of "garbled replies"

8.5 Facilities Control Advisory Monitor (VLAQFACS)

8.5.1 Purpose

The Facilities Control is equipped with a Model 28 RO which is utilized by the CP to automatically inform the controller of garbled/open circuit and circuit assurance reports.

8.5.2 Circuit Garbled-Open

If a circuit begins garbling or goes at random, the CP informs Facilities Control on VLAQFACS as follows:

- a. Station
- b. Current time
- c. Print-out of garbling/open

9.0 UTILITY ROUTINES

As the NASA Communications Processing System expands, many recurring services to the CP's may be performed more efficiently by designing the various utility routines needed for their execution. In this manner, responsibility for the correct completion of these tasks is transferred from the console operator to the CP. The need

for many routines can be foreseen at present and several of the applications are listed in the following paragraphs.

9.1 Tape to Drum

This routine loads a fixed format tape into a designated drum address in one of the following methods:

- a. With a compare of present contents
- b. Direct load without compare of present contents

Format xxxxxxxxxxxxyyyyyyyyzzzzzzzzzz

- a. x = drum address
- b. y = present contents of drum at this address*
- c. z = new contents to be entered

The Program Tape to be loaded is mounted on a servo and read into the core and started. The operator then makes a console entry designating a compare or no compare load. If no compare is requested, the tape is loaded to the drum and the routine closed out. If compare is requested, the content of the drum address is compared to the tapes present content and, if equal, enter new contents. If present contents do not compare, the routine sends the content of the drum to the console for a decision on making the change to this address. After all changes on tape are made, the program will close out.

9.2 Tape to Core

The routine is the same as tape to drum except the change affects the core only and is lost by bootstrapping the CP.

9.3 Type-Out Counters

Routine counters are placed on a routine-contained list and their contents printed on console at a designated time, or whenever requested. Typeout on console is in a format similar to the one below.

Format wwww is xxxx was yyyy dif zzzzz

- a. w = counter name, number, or address
- b. x = present contents in octal
- c. y = contents when last requested in octal
- d. z = difference in decimal

9.4 Octal to Spurt

This is a conversion routine for tape stored routine which converts machine instructions to a

*If no compare of present contents is required, second word (y) would be deleted when making tape.

XS-3 coded spurt equivalent and stores in a buffer for output to the 1004 printer.

9.5 De-bugging Aid

This is a routine which interlaces with a designated program that is in a process of being debugged. This routine has the ability to save specific registers and other relative information during the occurrence of a designated set of circumstances within the program being debugged.

10.0 MODULE AND CIRCUIT INVENTORY RECORDS

The growth of the NASA Communications Processing System necessitates many equipment and programming changes and a flexible method of establishing and maintaining an inventory of these changes has been devised.

The variety of records that are required seemed to eliminate the possibility of a fixed word count and format. The Module and Circuit Records are outlined below and any additional records of this type will follow a similar format.

10.1 Module Record

All permanent on-line module information is duplicated on magnetic tape. This information, preceded by certain processing program requirements, is the basic module record. One tape block contains one record. The following example is the present format for the variable length module record:

- a. first word: number of words in module; ESI number
- b. first word plus 1: base address; spare
- c. first word plus 2: first data word
- d. first word plus N: last word†

This record serves as an inventory of all modules and their current assignment. Maintenance to these records are frequent and for this reason simplicity in design was stressed.

A Module Record Maintenance Routine reads one tape block, identifies the ESI number, makes the necessary change to the record, and rewrites the new record on a new tape. In event there are no changes to be made, it is written on the new tape "as is" and the next tape block is read into the core.

†First word plus two through word plus N will be variable according to module length.

A Module Record List Program reads a block, converts all information to be printed on the 1004 card processor printer to XS-3 cord, packs the output buffer in the format required, and gives control to the printer routine.

10.2 Circuit Inventory Record

The Circuit Inventory Record contains all information relative to an individual circuit. The information is stored on magnetic tape; one record per tape block. Some of the requirements are listed below:

- a. Circuit number (or record number)
- b. Routine
- c. Drops
- d. Carrier
- e. Associated project (if any)
- f. Line speed
- g. Date of service
- h. History
- i. Part time or full time
- j. Associated equipment
- k. FDX, HDX, Simplex

The records are maintained for report reference, and as an inventory of stations and equipment associated with each circuit.

The accessing programs read one tape block into a designated core area and identifies the circuit being processed. Report programs retrieve needed information from the record, process it and read the next tape block. Record up-dating programs make necessary changes, if any, and rewrite the block on a new tape. Processing programs do not write on the input tape, for new tapes are made when changing record information.

11.0 METHODS OF TESTING

11.1 Format Line Concept

To conform with NASCOP, errors are selected by format line; that is, SOM, routing line and so forth. Each line is broken into a series of probable errors which may be encountered in a normal teletype message. Although each error is preconceived and placed in logical sequence; it is impractical, if not implausible, to generate an error to simulate all possibilities in advance. Experience in operating the system brings to light new error combinations continually.

Errors, which were selected on a past experience basis as typical, were adopted as a practical

means of evaluation. The most common error being extra "letter functions" which are inserted at the beginning of the message to provide a leader to accommodate the tape transmitter.

11.2 Mixed Combinations

Mixed combinations are designed to achieve error results which are not achieved by single format line errors. These combinations when properly synthesized create special conditions such as "misrouting" or "held for validation."

Message identification is obtained by several means. To eliminate identification by means of serial progression from full identification to no identification is a method to determine the message delivery cut-off point.

11.3 Radar and Telemetry Format

A "JJ" identifier in the SOM of a message denotes telemetry or radar data. An attempt is made to disguise this identifier by adding extra letters or leaving one of the J's out.

Header identification is gradually eliminated until no identification is present. This is to test at what point JJ data would be lost because of garbled headers.

11.4 Service Forward, Message Recall and Set Count

The numbering sequence and other characteristics preclude any preconceived fixed messages; however, the format to be used, with examples, is precut. The teletype tape for each test is cut just prior to the testing cycle.

11.5 Master Tape Log

The master log is a serial listing of the teletype test tapes. It is composed of numbers and letters relating to teletype format lines, that is, format line 1A, 1B, 1C . . . format line 2A, 2B, 2C . . . and so on. Also, each category such as JJ Data and Mixed Combinations is listed as JJ1, JJ2, JJ3 and MC1, MC2, MC3 . . . respectively. Each serial number is referenced to a particular error which has been pre-conceived. This numbering system is open-ended to facilitate additional listings where applicable.

11.6 Test Folders

The test folder has a number corresponding to a serial number in the master log. Each folder

contains a master test tape, a "hard copy" of the message on the tape, and a data sheet to annotate test results. The master test tapes are standard five level, clear hole, mylar tapes. A usable test tape may be cut from the master, always maintaining the master as a permanent record. An envelope is affixed to each folder which contains the master tape.

The text of the "hard copy" explains the type of error which is encountered when utilizing the enclosed test tape associated with the message.

The test results are in the form of notations regarding significant information about the characteristics observed during the testing period. These results may be a basis for change in the future operating techniques or format.

11.7 Progress Reports

Progress reports are submitted weekly summarizing the various stages of completion. The intent of each report is to disseminate information among cognizant personnel, concerning the testing and documentation phase.

11.8 Test Run Results

Folders are made that contain the input/output information (receive side and send side) with any intercept or advisory messages which may have been generated. The tests are listed by run numbers relative to each format line.

11.9 Adapting Tapes to Program Debugging

The beginning conditions as set forth in the Univac Systems Manual are programmed to achieve the original requirements. Each tape, when related to program evaluation, is used to check a specific function. Program errors may be detected when a tape fails to be processed correctly.

11.10 Program Analysis

In some cases tape errors have generated conditions which were unusual or which were not postulated. These errors provide the impetus for program changes, and new programming concepts may be formed as a result of these simple or compound errors.

11.11 Adapting Tapes to Operational Debugging

The system input and output is simulated with several ASR-28's. One teletype is chosen as a

"send side" and another is chosen as the "receive side." As each tape is sent, the receive side "hard copy" is evaluated along with any service messages which are generated by the CP and is compared with the send side. When an entire group of error tapes clear the CP and are correctly forwarded, the operational conditions are satisfied.

12.0 490/7094 INTERFACE

The 490 is a replacement switch for the 83B2 (AT&T) switch which was installed to direct radar teletype to the 7094's, and to direct text teletype from the 7094's to the operating sites in support of the Mercury Manned Space Flight Missions. A total of 14 incoming and three outgoing teletype circuits are involved. These circuits are a mixture of 60 wpm and 100 wpm.

During non-mission periods, the circuits are used for network checking and routine messages.

The old interface was a manual "patch cord." A one-for-one situation which passed all traffic on the circuits to the 7094's. To be able to use these circuits to support other missions not involving the 7094's, a means of directing traffic was needed. The 490 provides the directing ability, and the message header provides the direction of the message.

Interfacing in the new manner brought with it several problems, as follows.

12.1 Reliability

To provide 100 percent assurance that all radar data is received by the 7094's, all data formats plus any message which is in such a distorted state that it cannot be recognized as a valid text message is routed to the 7094's.

The 490 programs do not examine the text of these messages, but the 7094 programs do perform checks to insure proper data. Therefore, text messages will be rejected by the 7094's.

12.2 Compatibility with existing equipment

The existing inputs to the 7094's is through an IBM 7261 S/P converter, 7281 channel multiplexer and into the 7094 core buffers. Both 60 wpm and 100 wpm channels are installed in the 7281. To debug the interface programs, we use the same speed output from the 490. Otherwise a lengthy realignment would be necessary each time we tried to debug.

12.3 Formats-Incoming

Two routing indicators are assigned to the 7094's for text type traffic. The original radar routing indicator, consisting of two J's (JJ) is retained.

12.4 Formats-Outgoing

The 7094's and any other originator of messages in that area were given an instruction manual which details the formats to be used in the NASCOM network.

12.5 Message accounting

The 490's are programmed to maintain a message accounting for all messages entering the network. This requires the originator to generate a sequential number for every message. The matter of the sequential number posed a serious problem to the programmers of the 7094's. There are three 7094 computers involved and anyone of them can transmit at any time. Therefore, to generate a sequential number for each message would involve up-dating all computers each time a message is sent. There are several 7094 programs involved also which would necessarily be initialized to the present message count.

It should be noted that a message received out of sequence will not restrict if from being transmitted to the receiving station, but only causes an advisory message to be generated by the 490 explaining the situation, plus, intercepting the complete message for remedial action. In the case of the 7094's, all traffic would be intercepted from the time of the first non-sequential number received.

To offset this, the 490 is programmed to accept a "set count" message. Upon receipt of this message the 490 initializes the message count of the originator to zero. Therefore, each program used in the 7094's need only generate a set count message and start all messages at zero count, which solves the problem of synchronizing a counter in three computers.

12.6 Cadfiss

This is a network checking program which has been in existence for some time. It sends a message on each circuit and waits for the answering message, analyzes the results and forms a norm for traffic reliability for the network.

As there are no programming checks made on the returned messages to the originator, a one-

for-one circuit network is used. When a message enters a particular channel of the 7094 it is assumed to be from that station. This posed a message accounting problem again. To insure the proper return to the 7094's each 490 incoming circuit was assigned a corresponding outgoing circuit to the 7094's. In this way, the 7094 programs are assured that they are receiving the correct returning traffic.

As stated before, the present programs are directed towards our slow speed teletype traffic. While very important, these circuits are relatively easy to handle when compared to the high speed (1200-4800 baud) data. As we pick up the higher speed circuits, new programs must be developed particular to the needs of the circuit involved and unique to the type of data being transmitted. We believe that these new teletype circuits will be relatively easy to incorporate, for we have considered most factors within the program concept. However, new types of data such as the Digital Command System (DCS) with its partial recall criteria requires a complete new concept of the recall program. It will also be a part of the main operating programs and as stated before, any change to these programs requires the most serious, and comprehensive planning and debugging possible.

We also have several requirements for inventory type programs which furnish us a print-out of any or all of the network at any time. We are presently working on the means for up-dating these programs as new circuits, equipment, and so forth, are added to the network.

Fortunately, the matter of circuit inventory up-dating is automatic with the 490. In the past, we have had no centralized system for listing all of our circuits and an inventory required extensive research and must time. With the 490's, however, every circuit must be connected to a CLT and programmed for in advance. We, therefore, have all of the required information to maintain the complete up-to-date inventory at hand. It requires only an up-dating program to maintain the inventory. There are many other uses for the 490, in this type of "off-line" work and which will present themselves in the future. This is a continuing project and as other uses are found and implemented, this report will be up-dated and the information will be distributed.

THE DRIFT OF A 24-HOUR EQUATORIAL SATELLITE DUE TO AN EARTH GRAVITY FIELD THROUGH 4th ORDER*

C. A. WAGNER

This report extends previous investigations of 24-hour near equatorial earth satellites by considering the motion of such satellites in an earth gravity field through the 4th order. The three coupled second order linear differential equations of initial drift from a 24-hour equatorial circular reference orbit are presented. This linear system is analyzed for "stable regions" in the field within which drift motion is self limiting. A somewhat simplified form of the linear equations is integrated to give a general solution which defines the mean daily motion after injection to a high degree of accuracy.

INTRODUCTION

Much analytic work has been done recently on the motions of a near synchronous satellite of the earth with a triaxial (2nd order) gravity field (References 1 and 2). The oscillatory movement of such a satellite about the minor axis of the earth's elliptical equator has been sufficiently well described. This investigation into higher order gravity effects on 24-hour satellites was prompted by a recent refinement of knowledge about these higher order anomalies (Reference 3). The perturbation forces arising from the higher order tesseral harmonics of the earth's gravity field are small, but they are in resonance on a 24-hour satellite. The major conclusion of the investiga-

tion is that, while no absolutely stationary geographic points exist for 24-hour satellites in an earth field to 4th order, the overall features of the regime of motion in the dominant triaxial field (see Reference 1 and page 1024) still hold.

DERIVATION OF THE INTEGRALS OF PERTURBATIONAL MOTION FOR A 24-HOUR EARTH SATELLITE

The force field \mathbf{F} of the earth on a mass point m at earth centered r, ϕ, θ in inertial space (Figure 1) can be written:

$$\mathbf{F} = m \left\{ \hat{r} \frac{\partial V_E}{\partial r} + \frac{\hat{\phi}}{r} \frac{\partial V_E}{\partial \phi} + \frac{\hat{\theta}}{r \cos \phi} \frac{\partial V_E}{\partial \theta} \right\}, \quad (1)$$

where

$$V_E = \frac{\mu_E}{r} \sum_{n=2}^4 \sum_{m=0}^n \left[1 - \left(\frac{R_0}{r} \right)^n P_n^m(\sin \phi) J_{nm} \cos m(\theta - \theta_{nm}) \right] \quad (2)$$

(see Appendixes A and B). The XY plane is the earth's equatorial plane; θ_{nm} is the inertial longitude of the principal axis of symmetry of the earth's mass distribution accounted for by the nm harmonic of the geopotential V_E .

The acceleration of m in inertial space r, ϕ, θ is

$$\begin{aligned} \vec{a} = & \hat{r} (\ddot{r} - r\dot{\theta}^2 \cos^2 \phi - r\dot{\phi}^2) + \hat{\phi} \left[\left(\frac{1}{r} \right) \left(\frac{d}{dt} [r^2 \dot{\phi}] \right) + r\dot{\theta}^2 \cos \phi \sin \phi \right] \\ & + \hat{\theta} \left[\frac{1}{r \cos \phi} \frac{d}{dt} (r^2 \dot{\theta} \cos^2 \phi) \right]. \end{aligned} \quad (3)$$

*Published as NASA Technical Note D-2103, February 1964.

Writing $\mathbf{F}=\mathbf{ma}$, implies the following three scalar equations in the r, ϕ, θ components of force and acceleration:

$$\ddot{r} - r\dot{\theta}^2 \cos^2 \phi - r\dot{\phi}^2 = G_r (r, \phi, \theta - \theta_{nm}), \tag{4}$$

$$\frac{1}{r} \frac{d}{dt} (r^2 \dot{\phi}) + r\dot{\theta}^2 \cos \phi \sin \phi = G_\phi (r, \phi, \theta - \theta_{nm}), \tag{5}$$

$$\frac{1}{r \cos \phi} \frac{d}{dt} (r^2 \dot{\theta} \cos^2 \phi) = G_\theta (r, \phi, \theta - \theta_{nm}), \tag{6}$$

where the G 's are gravitational force components per unit mass (Appendix B).

Consider the XY equatorial plane of the earth, with the earth's equator reflecting the mass distribution due to the nm harmonic of the geopotential (Figure 2). λ is the geographic longitude of m ; λ_{nm} is the geographic longitude of the principal nm axis of earth symmetry. Thus, it is clear from Figure 2 that $\lambda - \lambda_{nm} = \theta - \theta_{nm}$. The potential in Appendix B is thus consistent with that in Reference 4.

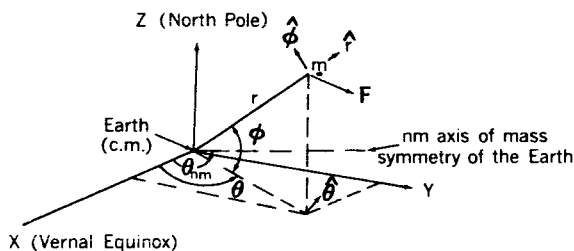


FIGURE 1.—Coordinate system referencing the motion of a 24-hour earth satellite.

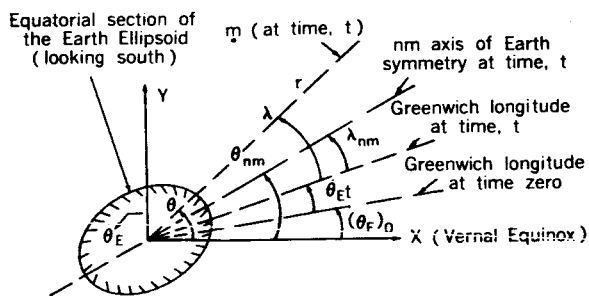


FIGURE 2.—Section of the earth ellipsoid's equator showing the relationship of the various longitude references.

The reference orbit for the synchronous satellite is a circle in the equatorial plane of radius r_s , traversed at the earth's rotation rate. Therefore, we assume a perturbation solution to Equations

4, 5 and 6 of the following form:

$$\begin{aligned} \theta = & [\text{Initial Inertial Longitude}] \\ & + [(\text{Earth Rate}) \cdot (\text{Time})] \\ & + [\text{Geographic Longitude Perturbation}] \\ = & [(\theta_E)_0 + \lambda_0] + [\dot{\theta}_E t] + [\Delta \lambda]; \end{aligned} \tag{7}$$

$$\phi = \Delta \phi; \tag{8}$$

$$r = r_s + \Delta r, \tag{9}$$

where $\dot{\theta}_E$ is the time constant earth rotation rate, λ_0 is a constant equal to the initial geographic longitude position of m , and $\Delta \lambda$ is the change in geographic longitude with time. Small r_s is a time-constant initial radius (the "synchronous radius") of m , to be determined later, Δr the change in radius of m with time, and $(\theta_E)_0$ the initial inertial longitude of Greenwich. The perturbation (Equation 7) replaces $\theta - \theta_{nm}$ in the force field of Appendix B, by $(\theta_E)_0 + \lambda_0 + \dot{\theta}_E t + \Delta \lambda - \theta_{nm}$. But $\theta_{nm} = (\theta_E)_0 + \dot{\theta}_E t + \lambda_{nm}$, from Figure 2. Therefore, for the perturbation solutions, the force field longitude arguments are

$$\lambda_0 - \lambda_{nm} + \Delta \lambda.$$

Let

$$\lambda_0 - \lambda_{nm} = \gamma_{nm} \tag{10}$$

define harmonic constants referred to the initial geographic longitude of m . Thus Equations 7, 8 and 9 are solutions to Equations 4, 5 and 6, if

$$\Delta \ddot{r} - (r_s + \Delta r)(\dot{\theta}_E + \Delta \dot{\lambda})^2 \cos^2 \phi - (r_s + \Delta r)(\Delta \dot{\phi})^2 = G_r(r_s + \Delta r, \Delta \phi, \Delta \lambda + \gamma_{nm}), \quad (11)$$

$$\frac{1}{(r_s + \Delta r)} \frac{d}{dt} (r_s + \Delta r)^2 (\Delta \dot{\phi}) + (r_s + \Delta r) (\dot{\theta}_E + \Delta \dot{\lambda})^2 \cos \Delta \phi \sin \Delta \phi = G_\phi(r_s + \Delta r, \Delta \phi, \Delta \lambda + \gamma_{nm}), \quad (12)$$

$$\frac{1}{(r_s + \Delta r) \cos \phi} \frac{d}{dt} (r_s + \Delta r)^2 (\dot{\theta}_E + \Delta \dot{\lambda}) \cos^2 \Delta \phi = G_\theta(r_s + \Delta r, \Delta \phi, \Delta \lambda + \gamma_{nm}). \quad (13)$$

Let $w = \dot{\theta}_E$, and by performing the indicated differentiations and substitutions, and ignoring all products of perturbations and/or perturbation rates, Equations 11-13 become

$$\ddot{r}_1 + ar_1 + b\dot{\lambda}_1 + c\lambda_1 + d\phi_1 = e \quad (14)$$

$$\ddot{\phi}_1 + f\phi_1 + gr_1 + h\lambda_1 = i \quad (15)$$

$$\ddot{\lambda}_1 + j\lambda_1 + k\dot{r}_1 + \ell r_1 + m\phi_1 = n. \quad (16)$$

where $r_1 = \Delta r/r_s$, $\lambda_1 = \Delta \lambda$, $\phi_1 = \Delta \phi$ are all dimensionless variables. The dots in Equations 14-16 and what follows refer to derivatives with respect to a dimensionless time $T = tw$. Thus

$$\frac{d}{dT} = \frac{1}{w} \frac{d}{dt}; \quad \frac{d^2}{dT^2} = \frac{1}{w^2} \frac{d^2}{dt^2}.$$

The constants in Equations 14-16 are:

$$\begin{aligned} a = & -1 - \frac{\mu_E}{w^2 r_s^3} \left\{ 2 + 6J_{20} \left(\frac{R_0}{r_s} \right)^2 - 36J_{22} \left(\frac{R_0}{r_s} \right)^2 \cos 2\gamma_{22} + 30 \left(\frac{R_0}{r_s} \right)^3 J_{31} \cos \gamma_{31} \right. \\ & - 300J_{33} \left(\frac{R_0}{r_s} \right)^2 \cos 3\gamma_{33} - 11.25J_{40} \left(\frac{R_0}{r_s} \right)^4 + 225J_{42} \left(\frac{R_0}{r_s} \right)^4 \cos 2\gamma_{42} \\ & \left. - 3150J_{44} \left(\frac{R_0}{r_s} \right)^4 \cos 4\gamma_{44} \right\}; \end{aligned} \quad (17)$$

$$b = -2; \quad (18)$$

$$\begin{aligned} c = & \frac{\mu_E}{w^2 r_s^3} \left\{ 18J_{22} \left(\frac{R_0}{r_s} \right)^2 \sin 2\gamma_{22} - 6J_{31} \left(\frac{R_0}{r_s} \right)^3 \sin \gamma_{31} + 180J_{33} \left(\frac{R_0}{r_s} \right)^3 \sin 3\gamma_{33} \right. \\ & \left. - 75J_{42} \left(\frac{R_0}{r_s} \right)^4 \sin 2\gamma_{42} + 2100J_{44} \left(\frac{R_0}{r_s} \right)^4 \sin 4\gamma_{44} \right\}; \end{aligned} \quad (19)$$

$$d = \frac{\mu_E}{w^2 r_s^3} \left\{ 6J_{30} \left(\frac{R_0}{r_s} \right)^3 - 60J_{32} \left(\frac{R_0}{r_s} \right)^3 \cos 2\gamma_{32} + \frac{75}{2} J_{41} \left(\frac{R_0}{r_s} \right)^4 \cos \gamma_{41} \right. \\ \left. - 525J_{43} \left(\frac{R_0}{r_s} \right)^4 \cos 3\gamma_{43} \right\}; \quad (20)$$

$$e = 1 + \frac{\mu_E}{w^2 r_s^3} \left\{ -1 - \frac{3}{2} J_{20} \left(\frac{R_0}{r_s} \right)^2 + 9J_{22} \left(\frac{R_0}{r_s} \right)^2 \cos 2\gamma_{22} - 6J_{31} \left(\frac{R_0}{r_s} \right)^3 \cos \gamma_{31} \right. \\ \left. + 60J_{33} \left(\frac{R_0}{r_s} \right)^3 \cos 3\gamma_{33} + \frac{15}{8} J_{40} \left(\frac{R_0}{r_s} \right)^4 - \frac{75}{2} J_{42} \left(\frac{R_0}{r_s} \right)^4 \cos 2\gamma_{42} \right. \\ \left. + 525J_{44} \left(\frac{R_0}{r_s} \right)^4 \cos 4\gamma_{44} \right\}; \quad (21)$$

$$f = 1 - \frac{\mu_E}{w^2 r_s^3} \left\{ -3J_{20} \left(\frac{R_0}{r_s} \right)^2 + 6J_{22} \left(\frac{R_0}{r_s} \right)^2 \cos 2\gamma_{22} - \frac{32}{2} J_{31} \left(\frac{R_0}{r_s} \right)^3 \cos \gamma_{31} \right. \\ \left. + 45J_{33} \left(\frac{R_0}{r_s} \right)^3 \cos 3\gamma_{33} + \frac{15}{2} J_{40} \left(\frac{R_0}{r_s} \right)^4 - 120J_{42} \left(\frac{R_0}{r_s} \right)^4 \cos 2\gamma_{42} \right. \\ \left. + 420J_{44} \left(\frac{R_0}{r_s} \right)^4 \cos 4\gamma_{44} \right\}; \quad (22)$$

$$g = -\frac{\mu_E}{w^2 r_s^3} \left\{ -9J_{30} \left(\frac{R_0}{r_s} \right)^3 + 90J_{32} \left(\frac{R_0}{r_s} \right)^3 \cos 2\gamma_{32} - \frac{105}{2} J_{41} \left(\frac{R_0}{r_s} \right)^4 \cos \gamma_{41} \right. \\ \left. + 735J_{43} \left(\frac{R_0}{r_s} \right)^4 \cos 3\gamma_{43} \right\}; \quad (23)$$

$$h = -\frac{\mu_E}{w^2 r_s^3} \left\{ 30J_{32} \left(\frac{R_0}{r_s} \right)^3 \sin 2\gamma_{32} - \frac{15}{2} J_{41} \left(\frac{R_0}{r_s} \right)^4 \sin \gamma_{41} \right. \\ \left. + 315J_{43} \left(\frac{R_0}{r_s} \right)^4 \sin 3\gamma_{43} \right\}; \quad (24)$$

$$i = \frac{\mu_E}{w^2 r_s^3} \left\{ \frac{15}{2} J_{30} \left(\frac{R_0}{r_s} \right)^3 - 15 J_{32} \left(\frac{R_0}{r_s} \right)^3 \cos 2\gamma_{32} + \frac{15}{2} J_{41} \left(\frac{R_0}{r_s} \right)^4 \cos \gamma_{41} - 105 J_{43} \left(\frac{R_0}{r_s} \right)^4 \cos 3\gamma_{43} \right\}; \quad (25)$$

$$j = - \frac{\mu_E}{w^2 r_s^3} \left\{ 12 J_{22} \left(\frac{R_0}{r_s} \right)^2 \cos 2\gamma_{22} - \frac{3}{2} J_{31} \left(\frac{R_0}{r_s} \right)^3 \cos \gamma_{31} + 135 J_{33} \left(\frac{R_0}{r_s} \right)^3 \cos 3\gamma_{33} - 30 J_{42} \left(\frac{R_0}{r_s} \right)^4 \cos 2\gamma_{42} + 1680 J_{44} \left(\frac{R_0}{r_s} \right)^4 \cos 4\gamma_{44} \right\}; \quad (26)$$

$$k = 2. \quad (27)$$

$$l = - \frac{\mu_E}{w^2 r_s^3} \left\{ - 30 J_{22} \left(\frac{R_0}{r_s} \right)^2 \sin 2\gamma_{22} + 9 J_{31} \left(\frac{R_0}{r_s} \right)^3 \sin \gamma_{31} - 270 J_{33} \left(\frac{R_0}{r_s} \right)^3 \sin 3\gamma_{33} + 105 J_{42} \left(\frac{R_0}{r_s} \right)^4 \sin 2\gamma_{42} - 2940 J_{44} \left(\frac{R_0}{r_s} \right)^4 \sin 4\gamma_{44} \right\}; \quad (28)$$

$$m = - \frac{\mu_E}{w^2 r_s^3} \left\{ 30 J_{32} \left(\frac{R_0}{r_s} \right)^3 \sin 2\gamma_{32} - \frac{15}{2} J_{41} \left(\frac{R_0}{r_s} \right)^4 \sin \gamma_{41} + 315 J_{43} \left(\frac{R_0}{r_s} \right)^4 \sin 3\gamma_{43} \right\}; \quad (29)$$

$$n = \frac{\mu_E}{w^2 r_s^3} \left\{ 6 J_{22} \left(\frac{R_0}{r_s} \right)^2 \sin 2\gamma_{22} - \frac{15}{2} J_{31} \left(\frac{R_0}{r_s} \right)^3 \sin \gamma_{31} + 45 J_{33} \left(\frac{R_0}{r_s} \right)^3 \sin 3\gamma_{33} - 15 J_{42} \left(\frac{R_0}{r_s} \right)^4 \sin 2\gamma_{42} + 420 J_{44} \left(\frac{R_0}{r_s} \right)^4 \sin 4\gamma_{44} \right\}. \quad (30)$$

Writing Equations 14-16 in operator notation ($s^1=d/dT$, $s^2=d^2/dT^2$, $s=s^1$, $s^0=1$, etc.), we have

$$(s^2 + a) r_1 + (bs + c) \lambda_1 + (d) \phi_1 = e, \tag{31}$$

$$(g) r_1 + (h) \lambda_1 + (s^2 + f) \phi_1 = i, \tag{32}$$

$$(ks + l) r_1 + (s^2 + j) \lambda_1 + (m) \phi_1 = n. \tag{33}$$

Solving Equations 31-33 by Cramer's rule, we obtain

$$r_1 = \frac{\begin{vmatrix} e & bs + c & d \\ i & h & s^2 + f \\ n & s^2 + j & m \end{vmatrix}}{\begin{vmatrix} s^2 + a & bs + c & d \\ g & h & s^2 + f \\ ks + l & s^2 + j & m \end{vmatrix}}$$

$$= \left\{ e [hm - s^4 - jf - s^2(f + j)] - i [m(bs + c) - d(s^2 + j)] + n (bs^3 + cs^2 + bfs + cf - h) \right\} / \left\{ [s^2 + a] [mh - s^4 - s^2(j + f) - jf] - g [m(bs + c) - d(s^2 + j)] + (ks + l) (bs^3 + cs^2 + bfs + cf - hd) \right\}. \tag{34}$$

Or

$$- s^6 r_1 + (A) s^4 r_1 + (B) s^3 r_1 + (C) s^2 r_1 + (D) s^1 r_1 + (E) s^0 r_1 = ehm + ejf - imc + idj + ncf - nhd. \tag{35}$$

This result follows from the evaluation of the determinants because $s(a, b, c, \dots) = 0$, since a, b, c, \dots are all constants. In summation, the three uncoupled linear drift equations of sixth order are:

$$\begin{aligned} [-s^6 + (A)s^4 + (B)s^3 + (C)s^2 + (D)s^1 + (E)s^0] r_1 &= A_1, \\ [-s^6 + (A)s^4 + (B)s^3 + (C)s^2 + (D)s^1 + (E)s^0] \lambda_1 &= A_2, \\ [-s^6 + (A)s^4 + (B)s^3 + (C)s^2 + (D)s^1 + (E)s^0] \phi_1 &= A_3. \end{aligned} \quad (36)$$

The constants are;

$$A = -a + kb - j - f, \quad (37)$$

$$B = kc + lb, \quad (38)$$

$$C = mh - aj - af - jf + gd + lc + bfk, \quad (39)$$

$$D = -gmb + kcf + lbf - khd, \quad (40)$$

$$E = amh - ajf - gmc + dgj + lcf - lhd, \quad (41)$$

$$A_1 = ehm + ejf - imc + idj + ncf - nhd, \quad (42)$$

$$A_2 = aim - anf - gem + gnd + lef - lid, \quad (43)$$

$$A_3 = ahn - aij - gcn + gej + lci - leh. \quad (44)$$

Examination of Equations 31-33 shows that for zero initial conditions ($r_1 = \dot{r}_1 = \phi_1 = \dot{\phi}_1 = \lambda_1 = \dot{\lambda}_1 = 0$, at $T=0$); $\ddot{r}_1 (T=0) = e$, $\ddot{\lambda}_1 (T=0) = n$ and $\ddot{\phi}_1 (T=0) = i$. The necessary and sufficient conditions for the drift to be zero for all time are, then, for the mass m to be placed with zero initial conditions into an orbit for which

$$e = n = i = 0. \quad (45)$$

That this is so may be shown by successive differentiation of Equation 31-33 for the higher derivatives. They will all be zero providing only that the initial perturbation, perturbation rates and perturbation accelerations are zero. But Equations 45 are three transcendental equations in the two unknowns λ_0 and r_s (the initial longitude and radius of the satellite). Therefore, there will be, in general, no simultaneous solution except by coincidence of the constants of those

equations. However, from what is known at present about the earth's gravitational field* (see Appendix C), the perturbation forces due to the latitude antisymmetry of the field (included in the i constant) are small compared to those in the radial and longitudinal directions at near synchronous altitudes. The latitude perturbations, then, may be neglected in considering the conditions for a near zero solution to Equations 31-33. A plot of these perturbation forces with λ_0 at an r_s determined from $e=0$, is found in Appendix C. It is postulated then, and proved later, that essentially stable regions of the gravity field in geographic-geocentric coordinates exist in the neighborhood of one or more points on the equator for which

$$e = n = 0. \quad (46)$$

*Kozai, Y., Private Communication, November 1963.

In the development which follows, the earth gravity field of Kozai is used together with an earth rotation rate from Reference 3. These earth constants are:*

$$\begin{aligned}
 \omega &= .7292115 \times 10^{-4} \text{ rad./sec.} && \text{(Reference 3)} \\
 \mu_E &= 3.986032 \times 10^{20} \text{ cm.}^3/\text{sec.}^2 \\
 R_0 &= 6.378165 \times 10^8 \text{ cm.} \\
 J_{20} &= 1082.48 \times 10^{-6} \\
 J_{30} &= -2.56 \times 10^{-6} \\
 J_{40} &= -1.84 \times 10^{-6} \\
 J_{21} &= 0. \\
 J_{22} &= -1.2 \times 10^{-6} \\
 J_{31} &= -1.9 \times 10^{-6} \\
 J_{32} &= -.14 \times 10^{-6} \\
 J_{33} &= -.10 \times 10^{-6} \\
 J_{41} &= -.52 \times 10^{-6} \\
 J_{42} &= -.062 \times 10^{-6} \\
 J_{43} &= -.035 \times 10^{-6} \\
 J_{44} &= -.031 \times 10^{-6} \\
 \lambda_{22} &= -26.4 \text{ degrees} \\
 \lambda_{31} &= 4.6 \text{ degrees} \\
 \lambda_{32} &= -16.8 \text{ degrees} \\
 \lambda_{33} &= 42.6 \text{ degrees} \\
 \lambda_{41} &= 237.5 \text{ degrees} \\
 \lambda_{42} &= 65.2 \text{ degrees} \\
 \lambda_{43} &= 0.5 \text{ degrees} \\
 \lambda_{44} &= 14.9 \text{ degrees}
 \end{aligned}
 \tag{47}$$

*The constants are taken from Kozai, Y., Private Communication, November 1962, unless otherwise noted.

The two transcendental equations in λ_0 and r_s arising from Equation 46 are:

$$\begin{aligned} \frac{w^2 r_s^3}{\mu_E} = & 1 + \frac{3}{2} J_{20} \left(\frac{R_0}{r_s} \right)^2 + \cos \lambda_0 \left\{ -6J_{31} \left(\frac{R_0}{r_s} \right)^3 \cos \lambda_{31} \right\} + \cos 2\lambda_0 \left\{ -9J_{22} \left(\frac{R_0}{r_s} \right)^2 \cos 2\lambda_{22} \right. \\ & + \left. \frac{75}{2} J_{42} \left(\frac{R_0}{r_s} \right)^4 \right\} + \cos 3\lambda_0 \left\{ -60J_{33} \left(\frac{R_0}{r_s} \right)^3 \cos 3\lambda_{33} \right\} + \cos 4\lambda_0 \\ & \cdot \left\{ -525J_{44} \left(\frac{R_0}{r_s} \right)^4 \cos 4\lambda_{44} \right\} - \frac{15}{8} J_{40} \left(\frac{R_0}{r_s} \right)^4 + \sin \lambda_0 \left\{ -6J_{31} \left(\frac{R_0}{r_s} \right)^3 \sin \lambda_{31} \right\} \\ & + \sin 2\lambda_0 \left\{ -9J_{22} \left(\frac{R_0}{r_s} \right)^2 \sin 2\lambda_{22} + \frac{75}{2} J_{42} \left(\frac{R_0}{r_s} \right)^4 \sin 2\lambda_{42} \right\} \\ & + \sin 3\lambda_0 \left\{ -60J_{33} \left(\frac{R_0}{r_s} \right)^3 \sin 3\lambda_{33} \right\} + \sin 4\lambda_0 \left\{ -525J_{44} \left(\frac{R_0}{r_s} \right)^4 \sin 4\lambda_{44} \right\} ; \quad (48) \end{aligned}$$

and

$$\begin{aligned} 0 = & \sin \lambda_0 \left[-\frac{3}{2} J_{31} \left(\frac{R_0}{r_s} \right) \cos \lambda_{31} \right] + \sin 2\lambda_0 \left[6J_{22} \cos 2\lambda_{22} - 15J_{42} \left(\frac{R_0}{r_s} \right)^2 \right] \\ & + \sin 3\lambda_0 \left[45J_{33} \left(\frac{R_0}{r_s} \right) \cos 3\lambda_{33} \right] + \sin 4\lambda_0 \left[420J_{44} \left(\frac{R_0}{r_s} \right)^2 \cos 4\lambda_{44} \right] \\ & + \cos \lambda_0 \left[\frac{3}{2} J_{31} \left(\frac{R_0}{r_s} \right) \sin \lambda_{31} \right] + \cos 2\lambda_0 \left[-6J_{22} \sin 2\lambda_{22} + 15J_{42} \left(\frac{R_0}{r_s} \right)^2 \right. \\ & \left. \sin 2\lambda_{42} \right] + \cos 3\lambda_0 \left[-45J_{33} \left(\frac{R_0}{r_s} \right) \sin 3\lambda_{33} \right] + \cos 4\lambda_0 \left[-420J_{44} \left(\frac{R_0}{r_s} \right)^2 \sin 4\lambda_{44} \right] . \quad (49) \end{aligned}$$

In general, for every finite non zero r_s , Equation 49 changes sign a minimum of two times and a maximum of eight times over the equator. The earth constants are such that for $r_s > R_0$, the right hand side of Equation 48 is very close to 1 for all points on the equator. Essentially then, Equations 48 and 49 decouple. Equation 48 may be solved separately for a near synchronous radius independent of longitude, leaving a small longitude dependent residual. With the near synchronous radius

so determined, the zero's of Equation 49 establish (to high accuracy) a minimum of two and a maximum of 8 potentially drift free points on the equator. The longitude dependent residual of equation 48 may then be solved to establish (to high accuracy) the radius to each potentially drift free point. The perturbation constants of the earth's gravity field are so much less than 1 that the iteration need not be carried further than outlined above (see Appendix D). The

results of this iteration of Equation 48 and 49 with the earth constants of Equation 47 are: The spherical earth "synchronous radius" is

$$r_s \text{ (Spherical)} = 138333942.5 \text{ ft. (26199.6103 statute miles);} \quad (50)$$

The "oblate earth" (including J_{20} and J_{40} potential terms) "synchronous radius" is

$$r_s \text{ (Oblate)} = 138335648.5 \text{ ft. (26199.9334 statute miles).} \quad (51)$$

With the aforementioned oblate earth synchronous radius, Equation 49 becomes

$$\begin{aligned} 0 = & .430 \sin \lambda_0 - 4.37 \sin 2\lambda_0 + .4175 \sin 3\lambda_0 - .151 \sin 4\lambda_0 - .0345 \cos \lambda_0 \\ & - 5.76 \cos 2\lambda_0 + .538 \cos 3\lambda_0 + .257 \cos 4\lambda_0. \end{aligned} \quad (52)$$

The zero's of Equation 52, which are the potentially drift free longitudes around the equator, are at

$$\lambda_0 = 64.2^\circ, 155.8^\circ, 242.9^\circ \text{ and } 331.3^\circ. \quad (53)$$

It is interesting to compare these potentially stable longitudes with those which would be present if all the tesseral and sectorial harmonics except J_{22} are ignored. In the simpler field (the

so-called "triaxial" gravity field), the zero's of Equation 48 with the oblate earth synchronous radius of Equation 51 are at

$$\lambda_0 = 63.6^\circ, 153.6^\circ, 243.6^\circ \text{ and } 333.6^\circ. \quad (54)$$

In no case (with a full earth potential) do the "stable" longitudes differ by more than 2.3° from those which exist in the simpler "triaxial" field.

The potentially drift free radii to the longitudes of Equation 53 are:

$$r_s(\lambda_0 = 64.2^\circ) = 138335637.5 \text{ ft. (26199.9313 statute miles)} \quad (55a)$$

$$r_s(\lambda_0 = 155.8^\circ) = 138335660.2 \text{ ft. (26199.9356 statute miles)} \quad (55b)$$

$$r_s(\lambda_0 = 242.9^\circ) = 138335635.2 \text{ ft. (26199.9309 statute miles)} \quad (55c)$$

$$r_s(\lambda_0 = 331.3^\circ) = 138335660.3 \text{ ft. (26199.9356 statute miles).} \quad (55d)$$

The uncoupled linearized drift Equations 36 may be simplified and integrated directly by ignoring all terms in the differential coefficients A, B, C, D and E which are much less than 1. The driving terms A_1, A_2, A_3 must retain at least one order of smallness less than 1 so that the resulting solution is sufficiently sensitive to drift acceleration. The initial radius may be chosen for convenience as the mean of those in Equation 55. But to insure the longest possible validity for the resulting solution, it is probably best to solve $e=0$ for r_s at the λ_0 from which the perturbation is desired. In any

case, $r_s = 138335647.7 \pm 12.5$ ft. for near zero solutions to Equation 36 with zero initial conditions. The simplified uncoupled drift equations then become:

$$\begin{aligned} (s^6 + 2s^4 + s^2) r_1 &= -A_1 \\ (s^6 + 2s^4 + s^2) \lambda_1 &= -A_2 \\ (s^6 + 2s^4 + s^2) \phi_1 &= -A_3. \end{aligned} \quad (56)$$

It may be verified that the complete solution to Equation 56 is:

$$\Delta_q = C_{1q} + C_{2q}T + (C_{3q} + C_{4q}T) \sin T + (C_{5q} + C_{6q}T) \cos T - \frac{A_q T^2}{2} \quad (57)$$

where

$$\Delta_q = r_1 \text{ when } q = 1,$$

$$\Delta_q = \lambda_1 \text{ when } q = 2,$$

$$\Delta_q = \phi_1 \text{ when } q = 3.$$

In any dynamics problem utilizing the perturbation solutions (Equation 57), 18 conditions on the

perturbations must be specified.

The coefficients C_{iq} of the approximate perturbation solution (Equation 57) for the synchronous equatorial satellite have been solved in Appendix A of Reference 5 for the general case where any small initial perturbations and perturbation rates may be given. Ignoring terms of second and lower order smallness, the linearized drift solution for the 24-hour equatorial satellite is approximately

$$r_1 = 4s_0^0 r_1 + 2s_0^1 \lambda_1 + e + 2nT + [s_0^1 r_1 - 2n - (125s_0^0 r_1)T] \sin T + [-3s_0^0 r_1 - 2s_0^1 \lambda_1 - e] \cos T; \quad (57a)$$

$$\lambda_1 = -2s_0^1 r_1 + s_0^0 \lambda_1 + 4n + [18s_0^0 r_1 - 3s_0^1 \lambda_1 - 2e]T + (-24s_0^0 r_1 + 4s_0^1 \lambda_1 + 2e) \sin T + [2s_0^1 r_1 - 4n + (6s_0^0 r_1)T] \cos T - \frac{3}{2}nT^2; \quad (57b)$$

$$\phi_1 = i + [s_0^1 \dot{\phi}_1] \sin T + [s_0^0 \phi_1 - i] \cos T \quad (57c)$$

The $s_0^0 \Delta_q$ are initial dimensionless perturbations; the $s_0^1 \Delta_q$ are initial dimensionless perturbation rates.

DRIFT OF A 24-HOUR EQUATORIAL SATELLITE WITH LOW INITIAL RATES

The general character of the drift following near perfect injection at radii for which e is equal to or close to zero, may be found by examining the approximate integrals (Equations 57a, 57b, 57c). There is a coupled daily harmonic oscillation in the drift motion due to the eccentricity introduced principally by the resonant longitudinal perturbation force (when present). This daily harmonic oscillation is also partly due to coupling from the latitude and radial perturbations. The long term drift in radius is controlled by the $2nT$ term for perfect injection ($s_0^1 = s_0^0 = 0$). The long term drift in longitude is controlled by the $-3n/2T^2$ term for perfect injection. The sign of n changes four times around the equator. Table 1 shows the dominant drift effect following a near perfect "synchronous" injection. Thus the long term

Table 1
The Dominant Drift Effect Following a Near Perfect "Synchronous" Injection

Longitude Range (degrees)	Sign (n) (dominant drift in radius)	Sign (-n) (dominant drift in longitude)
62.4 < λ_0 < 155.8:	+	-
155.8 < λ_0 < 242.9:	-	+
242.9 < λ_0 < 331.3:	+	-
331.3 < λ_0 < 62.4:	-	+

radial and longitudinal drift following near perfect injection at near synchronous radii ($e \approx 0$) around the equator has the character (not to scale) shown in Figure 3. The long term motion in longitude and radius is thus highly suggestive of the coupled long term circulation about the minor axis of the earth ellipsoid's equator predicted* and confirmed in the computer studies in Reference 1; both sides made with a triaxial gravitational field. All the results of this study showing the dominance of the J_{22} (triaxial) term in the longitudinal perturbations, point strongly to the conclusion that

*Frick and Garber, Private Communication, 1962.

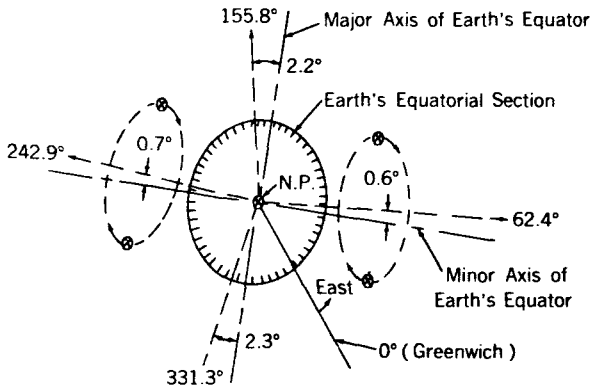


FIGURE 3.—Long term radial and longitudinal drift following near perfect injection at near synchronous radii around the equator.

such a circulation about two dynamically “stable” regions of zero longitude and radial force perturba-

tion, exists for the full earth potential as well.

The latitude drift following a near perfect “synchronous” injection, predicted in (57c) is initially a daily harmonic oscillation of small amplitude, but predominantly on one side of the equator or the other depending on the sign of i which is proportional to the latitude perturbing force at the equator. Since J_{30} (the so-called “pear shaped” harmonic of the earth’s field) dominates i and it is negative, the perturbing force is directed south and the initial daily oscillation is southerly with an amplitude of $2i$.

The drift of a near synchronous earth satellite injected with zero drift rates at -71.4° longitude from Greenwich (45° from the minor axis of the earth ellipsoid’s equator, and thus having close to maximum longitude perturbations) as predicted by Equations 57a, 57b, 57c, is

$$\Delta r = 259t - 41.25 \sin(2\pi t) \text{ ft. (t, in sidereal days);} \tag{57d}$$

$$\Delta \lambda = 34.2 \times 10^{-6} (1 - \cos 2\pi t) - 505 \times 10^{-6} t^2 \text{ degrees, (t, in sidereal days);} \tag{57e}$$

$$\Delta \phi = -10.7 \times 10^{-7} (1 - \cos 2\pi t) \text{ degrees, (t, in sidereal days).} \tag{57f}$$

Numerical integration (Cowell step by step method) of the equations of motion have been carried out on an IBM 7090 computer and the comparison with the predictions of Equations 57d, 57e, 57f is illustrated in Figures 4 and 5 for the longitudinal and radial drifts to 30 days following injection. The full results indicate that the linear theory of Equation 57 will predict the perturbations of a near synchronous, near equatorial satellite due to a “full” earth potential through 4th order, to within 2 percent in the mean daily longitude drift, and 1 percent in the mean daily radial drift for up to 180 days following a near perfect injection. The numerically computed latitude drift of the near synchronous satellite is of the order of $\pm 10^{-6}$ degrees, maximum, over 180 sidereal days, which agrees with (57f) in order of magnitude. Numerically integrated drifts for the above example in a “triaxial” (J_{20} and J_{22} harmonics only) earth field show errors of about 10 percent from the drift in a full potential field in both radius and longitude to 180 days following injection. This “error” was reduced to about 5 percent with the

inclusion of the J_{31} harmonic in the programmed potential function. Inclusion of the J_{33} harmonic

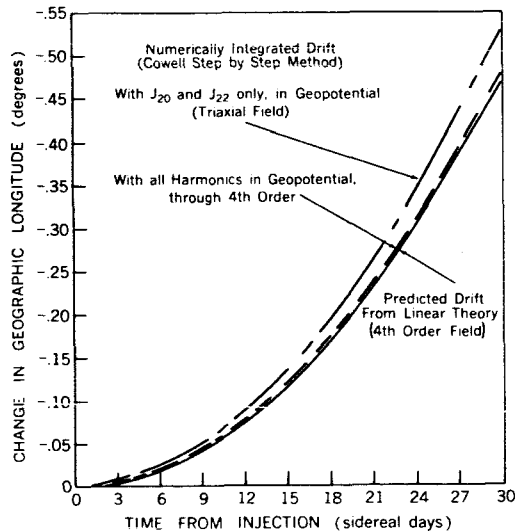


FIGURE 4.—Comparison of numerically integrated and predicted longitude drift of a near 24-hour satellite injected at -71.4° (45° east of the earth ellipsoid’s minor equatorial axis) at a radius of 138,335,650.2 feet, with zero initial perturbation rates.

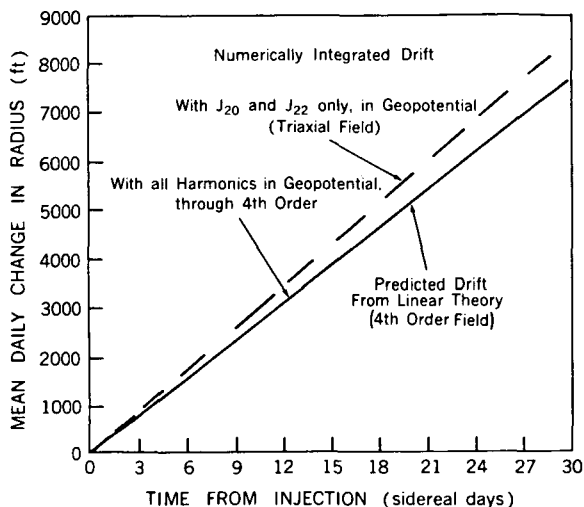


FIGURE 5.—Comparison of numerically integrated and predicted mean daily radial drift of a near 24-hour satellite injected at -71.4° at a radius of 138,335,650.2 feet, with zero initial perturbation rates.

reduces the "error" to about 1 percent over 180 days.

It should be noted that the linear theory presented in this report does *not* predict the change in eccentricity with time of the spiralling orbit of the resonant near synchronous satellite. However, the results of the numerical integration on

the above example show that the *initial* eccentricity of the spiral orbit is well predicted. The radius in the numerically integrated orbit has a daily oscillation of about $\pm(34+5t)$ feet, (t , in days) for a period of 5 days following injection; which is in excellent agreement with (74a) considering the inherent machine error in the numerical integration and the simplifications theory.

STABILITY OF MOTIONS OF THE 24-HOUR SATELLITE NEAR POINTS OF ZERO LONGITUDINAL AND RADIAL PERTURBATION FORCES

If the longitudinal and radial perturbation forces are zero, it may be verified that $e=n=0$ (see Appendixes B and C) and the linearized drift Equations 31-33 become

$$\left. \begin{aligned} (s^2 + a)r_1 + (bs + c)\lambda_1 + (d)\phi_1 &= 0 ; \\ (g)r_1 + (h)\lambda_1 + (s^2 + f)\phi_1 &= i ; \\ (ks + l)r_1 + (s^2 + j)\lambda_1 + (m)\phi_1 &= 0 . \end{aligned} \right\} \quad (58)$$

The characteristic equation of the system (Equations 58) is (evaluating the differential constants of Equation 36 for $e=n=0$)

$$\begin{aligned} q^6(-1) + q^4(-a + kb - j - f) + q^3(kc + lb) + q^2(mh - aj - af - jf + gd + lc + bfk) \\ + q(-gmb + kcf + lbf - khd) + amh + dgj + lcf - lhd = 0 . \end{aligned} \quad (59)$$

The particular solutions to Equation 58 will be constants:

$$\left. \begin{aligned} r_1 (\text{particular}) &= \frac{E}{A_1} ; \\ \lambda_1 (\text{particular}) &= \frac{E}{A_2} ; \\ \phi_1 (\text{particular}) &= \frac{E}{A_3} . \end{aligned} \right\} \quad (60)$$

To simplify the calculation of the roots of the characteristic Equation 59 without losing anything essential in the characteristics of the motion, we will ignore the lower order perturbation constants

and/or products of perturbation constants in each coefficient of Equation 59. With this simplification and with the large constants evaluated, Equation 59 becomes

$$q^6(-1) + q^4(-2) + q^3(2c - 2l) + q^2(-1) + q(2c - 2l) + 3j = 0. \tag{61}$$

It is noted that the characteristic values of the motion from Equation 61 are independent of the latitude perturbations of the earth's gravity field. Consider $q_{1,2}$ as solutions to Equation 61, where $|q_{1,2}| \ll 1$. For these solutions, ignore orders of q smaller than q^2 . Equation 61 in $q_{1,2}$ then becomes;

The solutions to Equation 62 are

$$q_{1,2} = -(l - c) \pm [(l - c)^2 + 3j]^{\frac{1}{2}}. \tag{63}$$

Since $|l - c|, |3j| \ll 1$, the approximation to Equation 62 for these roots is valid.

There are also solutions to Equation 61: $|q| \approx (-1)^{\frac{1}{2}}$. Therefore, let

$$q^2 + 2(l - c)q - 3j = 0. \tag{62} \qquad q_{3,4} = +(-1)^{\frac{1}{2}} + \epsilon_{3,4}. \tag{64}$$

where $|\epsilon_{3,4}| \ll 1$. With Equation 64 substituted into Equation 61, Equation 61 becomes

$$\begin{aligned} & [(-1)^{\frac{1}{2}} + \epsilon_{3,4}]^6 + 2[(-1)^{\frac{1}{2}} + \epsilon_{3,4}]^4 + 2(l - c) [(-1)^{\frac{1}{2}} + \epsilon_{3,4}]^3 \\ & + [(-1)^{\frac{1}{2}} + \epsilon_{3,4}]^2 + [(-1)^{\frac{1}{2}} + \epsilon_{3,4}] [2(l - c)] - 3j = 0. \end{aligned} \tag{65}$$

Since $|\epsilon_{3,4}| \ll 1$, the expansion of Equation 65 ignoring terms in ϵ_3 and higher, is

$$\begin{aligned} & 6\epsilon [(-1)^{\frac{1}{2}}]^5 + 8\epsilon [(-1)^{\frac{1}{2}}]^3 + 6(l - c)\epsilon [(-1)^{\frac{1}{2}}]^2 + 2\epsilon(-1)^{\frac{1}{2}} - 2(l - c)(-1)^{\frac{1}{2}} \\ & + 2\epsilon(l - c)(-1)^{\frac{1}{2}} + 2(l - c)(-1)^{\frac{1}{2}} + 15\epsilon^2 - 12\epsilon^2 + 6(l - c)\epsilon^2(-1)^{\frac{1}{2}} \\ & + \epsilon^2 = 0. \end{aligned}$$

This last equation reduces to

$$\begin{aligned} & \epsilon^2 [15 - 12 + 6(l - c)(-1)^{\frac{1}{2}} + 1] + \epsilon [6(-1)^{\frac{1}{2}} - 8(-1)^{\frac{1}{2}} - 6(l - c) + 2(-1)^{\frac{1}{2}} \\ & + 2(l - c)(-1)^{\frac{1}{2}}] = 0. \end{aligned}$$

Thus, the approximate roots to Equation 65 are

$$\epsilon_3 = 0, \qquad \epsilon_4 = \frac{-2(l - c) [-3 + (-1)^{\frac{1}{2}}]}{4}. \tag{66}$$

Similarly, letting

$$q_{5,6} = -(-1)^{\frac{1}{2}} + \epsilon_{5,6}. \tag{67}$$

The expansion of Equation 61, with the same order of approximation as for $\epsilon_{3,4}$, reduces to

$$6\epsilon \left[-(-1)^{\frac{1}{2}} \right]^5 + 8\epsilon \left[-(-1)^{\frac{1}{2}} \right]^3 + 6(l-c)\epsilon \left[-(-1)^{\frac{1}{2}} \right]^2 + 2\epsilon \left[-(-1)^{\frac{1}{2}} \right] + 2(l-c)(-1)^{\frac{1}{2}} - 2\epsilon(l-c)(-1)^{\frac{1}{2}} - 2(l-c)(-1)^{\frac{1}{2}} + 15\epsilon^2 - 12\epsilon^2 - 6(l-c)\epsilon^2(-1)^{\frac{1}{2}} + \epsilon^2 = 0.$$

The roots to the above equation are

$$\epsilon_5 = 0, \quad \epsilon_6 = \frac{-2(l-c) \left[-(-1)^{\frac{1}{2}} - 3 \right]}{4}. \quad (68)$$

Thus Equations 66 and 68 substituted into Equations 64 and 67 determine the other four roots to Equation 61 as approximately

$$q_{3,5} = \pm(-1)^{\frac{1}{2}}, \quad q_{4,6} = \pm(-1)^{\frac{1}{2}} \left[1 - \frac{1}{2}(l-c) \right] + \frac{3}{2}(l-c). \quad (69)$$

Stability in the Triaxial Geopotential Field

In this case $l=c=0$ at the points of zero longitudinal and radial perturbation forces. The four characteristic solutions are

$$q_{1,2} \text{ (triaxial)} = \pm(3j)^{\frac{1}{2}}, \quad (70a)$$

$$q_{3,4} \text{ (triaxial)} = \pm(-1)^{\frac{1}{2}}. \quad (70b)$$

While it is true that there are two more independent complementary solutions to Equation 58 that may be found for the triaxial case, their constants will be found to be zero when a natural set of initial conditions are specified. For the triaxial case, $g=h=d=i=m=0$. Therefore Equation 58 uncouples in the latitude variation and there are only four characteristic solutions to the set in the longitude and radial perturbations. It may be shown that (70a) and (70b) are just these four. The redundant $q_{5,6} = \pm(-1)^{\frac{1}{2}}$ are the two characteristic solutions to the latitude variation in the triaxial case.

On or Near the Major Axis in the Triaxial Field

$$3j (\lambda_0 = 153.6^\circ, -26.4^\circ) = +0.99 \times 10^{-6}$$

The motion thus has a slowly divergent component.

On or Near the Minor Axis in the Triaxial Field

$$3j (\lambda_0 = 63.6^\circ, -116.4^\circ) = -0.99 \times 10^{-6}$$

The motion consists of two non-damped self-limiting oscillations. One has a period in the

neighborhood of one day (from $q_{3,4}$). The other has a frequency of $(.99)^{\frac{1}{2}} \times 10^{-3} = 0.995 \times 10^{-3}$ (dimensionless) $= .995 \times 10^{-3} \text{ w}$ (dimensions of time^{-1}). This long period oscillation has a period in the neighborhood of

$$\frac{2\pi(\text{rad./cycle})}{.995 \times 10^{-3} \times 2\pi(\text{rad./day})} = 1005 \text{ days} = 2.76 \text{ years}.$$

Stability in the Geopotential Field Through 4th Order

For motion in the vicinity of the 4 points of zero longitudinal and radial perturbation forces:

1. There will be 4 characteristic solutions giving damped oscillations with periods near one [day $q_{3,4,5,6}$ from Equation 69]
2. There will be 2 characteristic solutions giving either a weak-negatively damped oscillation with a period near 1000 days, or two exponentials, one of which is slowly divergent in character. [$q_{1,2}$ from Equation 63]

For Motion in the Vicinity of the Zero Perturbations at $\lambda_0 = 64.2^\circ$

$$l-c = -0.36 \times 10^{-9},$$

$$3j = -1.026 \times 10^{-6}$$

Perfect injection in the vicinity of this quasi-stable point is thus followed by two self-limiting oscillations with periods in the

vicinity of one day which are weak-positively damped ($q_{3, 4, 5, 6}$). There is also a self-limiting oscillation with a long period of $1/(1.026)^{\frac{1}{2}} \times 10^{-3} = 987$ Days = 2.71 Years, which is weak-negatively damped ($q_{1, 2}$).

For Motion in the Vicinity of the Zero Perturbations at $\lambda_0 = 155.8^\circ$

$3j$ is controlled by the J_{22} term, is greater than zero, and $\pm(3j)^{\frac{1}{2}}$ is of the order of $\pm 10^{-3}$. $|l-c|$ is of the order of 10^{-9} (as before), so that one of the characteristic solutions will be a slowly divergent exponential.

For Motion in the Vicinity of the Zero Perturbations at $\lambda_0 = 242.9^\circ (-117.1^\circ)$

$$l - c = + 5.3 \times 10^{-9} .$$

$$3j = - 1.123 \times 10^{-6} .$$

Perfect injection in the vicinity of this quasi-stable point is thus followed by two self-limiting coupled oscillations with periods in the vicinity of one day which are weak-negatively damped ($q_{3, 4, 5, 6}$). There is also a coupled self-limiting oscillation with a long period near $1/(1.123)^{\frac{1}{2}} \times 10^{-3} = 944$ days = 2.58 year, which is weak-positively damped ($q_{1, 2}$).

For Motion in the Vicinity of the Zero Perturbations at $\lambda_0 = 331.3^\circ (-28.7^\circ)$

$3j$, controlled by the J_{22} term, is greater than zero. $\pm(3j)^{\frac{1}{2}}$ is of the order of $\pm 10^{-3}$.

$|l-c|$ is of the order of 10^{-9} as before. Therefore, the resultant motion is slowly divergent in character as one of the $q_{1, 2}$ solutions will be positive of the order of $+10^{-3}$.

In conclusion, for the geopotential gravity field through 4th order, two regions on the equator have been found within which small initial perturbations in geographic longitude, latitude, and radius are self limiting in the sense that the resulting motion of an earth satellite in these regions is essentially harmonic in character with very weak damping. These regions are in the neighborhood of 64.2° East of Greenwich and $26,199.9314$ statute miles and 117.1° West of Greenwich and $26,199.9309$ statute miles from the center of mass of the earth. They are both within $2\frac{1}{2}$ degrees of the minor axis of the earth ellipsoid's equator. The damping is of the order of 10^{-9} wt, or of the order of $10^{-8} t$ (days). Thus, initial amplitudes of the damped harmonic perturbations in these two regions suffer a twofold change in magnitude in the order of $t = \ln(2)/10^{-8} = 6.93 \times 10^7$ days = 1.9×10^5 years. These regions then, can be considered to be essentially stable.

While the regions around 64.2° and 242.9° on the equator have an inherent dynamic stability in the sense shown above, the regions around the zero perturbation points at 155.8° and 331.3° are only very weakly dynamically unstable. At the $n=e=0$ point where $\lambda_0 = 331.3^\circ$ and $r_s = 138,335,660.3$ ft., by assuming zero initial perturbations and rates, the resultant motion from a complete evaluation of the C_{iq} coefficients of Equation 57 in Appendix A of Reference 5, is

$$r_1 \doteq - 1.63 \times 10^{-16} + (- 1.76 \times 10^{-16}T) + (+ 2.64 \times 10^{-16} - .815 \times 10^{-16}T) \sin T + (1.63 \times 10^{-16} - .88 \times 10^{-16}T) \cos T - 21.3 \times 10^{-24}T^2 , \tag{71a}$$

$$\lambda_1 \doteq - 6.15 \times 10^{-16} + (3.26 \times 10^{-16}T) + (- 4.89 \times 10^{-16} + 1.76 \times 10^{-16}T) \sin T + (+ 6.15 \times 10^{-16} + 1.63 \times 10^{-16}T) \cos T + 13.2 \times 10^{-17}T^2 , \tag{71b}$$

$$\phi_1 \doteq - 6.4 \times 10^{-9} + (- 23.7 \times 10^{-14}T) \sin T + (6.4 \times 10^{-9} \cos T) + 24.5 \times 10^{-16}T^2 . \tag{71c}$$

Equation 71a predicts a change in injection radius of only $-.0003$ feet in 5 years. Equation 71b predicts a change in injection geographic longitude of only $+1.00 \times 10^{-6}$ degrees in 5 years. Equation 71c predicts a change in injection latitude of only $+1.85 \times 10^{-5}$ degrees in 5 years. These are all mean daily drifts. The amplitudes of the daily oscillations are even smaller quantities. The conclusion is that station keeping requirements for near synchronous satellites placed with low initial rates near the major axis of the earth ellipsoid's equator, will be virtually unaffected by the noncentral character of the earth's gravitational field.

CONCLUSIONS

The major conclusion of this paper may be summarized as follows:

There are four longitudes, located within $2\frac{1}{2}$ degrees of the axes fixed by the earth's elliptical equator, into which a 24-hour satellite may be placed and maintained with negligible drift for extended periods of time.

At intermediate longitudes (about 45 degrees from these "stable points") such a satellite will, even if injected "perfectly", experience a minimum of about 4 degrees of drift in 3 months following injection.

Maximum drift of a perfectly injected 24-hour satellite in an earth gravity field to 4th order differs by about 10 percent from the drift experienced in a triaxial field.

Except in the immediate vicinity of the "stable points," the initial mean daily drift in radius following a perfect injection into a 24-hour orbit is proportional to time, and the initial mean daily drift in geographic longitude is essentially proportional to the square of time.

The maximum longitude perturbational force on a 24-hour satellite in an Earth potential field to 4th order is approximately 12 percent greater than the maximum perturbation experienced in a triaxial field.

REFERENCES

1. BLITZER, L., BOUGHTON, E. M., KANG, G., and PAGE, R. M., "Effects of Ellipticity of the Equator on 24-Hour Nearly Circular Satellite Orbits," *J. Geophys. Res.* 67(1):329-335, January 1962.
2. BARRETT, C. C., "The Perturbations of a Synchronous Satellite Resulting from the Gravitational Field of a Triaxial Earth," Mission Analysis Report-MAR 2, GSFC Document No. X-623-62-160, September 10, 1962.
3. ISLEY, W. C., "A Summary of Constants Associated with Orbital Analysis of Earth Satellites Including the Influence of Their Uncertainties Upon Gravitational Measurements for Synchronous Satellites," Mission Analysis Report-MAR 1, GSFC Document No. X-623-62-169, September 1962.
4. WAGNER, C. A., "The Gravitational Potential of a Triaxial Earth," Mission Analysis Report-MAR-3, GSFC Document No. X-623-62-206, October 31, 1962.
5. WAGNER, C. A., "The Perturbations of a 24 Hour Near Equatorial Satellite Due to an Earth Gravity Potential Through 4th Order," Mission Analysis Report-MAR 6, GSFC Document No. X-623-63-62, February 11, 1963.

APPENDIX A

LIST OF SYMBOLS

- A, B, C, D, E, A₁, A₂, A₃ Perturbation constants of the uncoupled perturbation equations of motion of the near synchronous satellite of the earth
- a, b, c, d, e, f, g, h, i, j, k, l, m, n Perturbation constants of the coupled linearized perturbation equations of motion with respect to the synchronous circular orbit of radius r_s
- d The total differential operator
- \vec{F} The earth's gravity field
- g_s The acceleration of earth gravity at the near synchronous radius
 $r_s \cdot g_s \approx 32.15 R_0/r_s)^2 \approx 32.15 \times .02288 = .7355 \text{ ft/sec}^2$
- i In addition to a perturbation constant, used as an index to the coefficients of the drift Equations 57
- m A point mass
- n, m When used in the earth potential function; indicates the harmonic of order n and power m
- q An index for the coefficients of the drift Equations 57. Also, a characteristic solution of the uncoupled motion Equations 36
- R₀ or R_e The mean equatorial radius of the earth ellipsoid
- r, λ, φ Spherical coordinates of the near synchronous satellite; geocentric radius; geographic longitude with respect to the greenwich meridian, and geocentric latitude from the earth's equator
- r₁, λ₁, φ₁ Dimensionless perturbation coordinates

$$r_1 = \frac{\Delta r}{r_s}, \quad \lambda_1 = \Delta \lambda \text{ (radians)}, \quad \phi_1 = \Delta \phi \text{ (radians)}$$

- Δr, Δλ, Δφ Perturbation coordinates in radius, geographic longitude and latitude from a reference equatorial orbit which is circular, having the near synchronous radius r_s and a period of exactly one sidereal day. These coordinates measure the drift of a near synchronous earth satellite from a point in space at a fixed r_s from the c.m. of the earth, moving along the equator at the earth's rotation rate so as to maintain a fixed geographical longitude at all times
- $\hat{r}, \hat{\theta}, \hat{\phi}$ Unit Vectors for the spherical coordinate system: r, θ, φ
- r, θ, φ Spherical coordinates of geocentric radius, inertial longitude from the vernal equinox, and geocentric latitude from the earth's equator, locating the near synchronous satellite m in inertial space
- r_s A nominal or calculated orbit radius at injection for a near synchronous earth satellite
- s A differential operator:

$$s^0 () = (); \quad s^1 () = \frac{d()}{dt}, \quad s^2 () = \frac{d^2 ()}{dt^2}, \text{ etc.};$$

and, s₀⁰ () = () at time T=0, etc.

- T Dimensionless time variable: $T = wt$, where w = the earth's sidereal rotation rate, and t is real time
- t Real time, from a zero at the point of injection of the satellite into its near synchronous orbit
- ($\dot{}$), ($\ddot{}$) Differentiation with respect to real time t , prior to Equation 14, and differentiation with respect to dimensionless time T in and after Equation 14
- V_E The earth's gravity potential field
- $\gamma_{nm}, J_{nm}, \lambda_{nm}, \theta_{nm}$ Constants of the earth's gravity potential:
- ϵ A small parameter
- $(\theta_E)_0$ The right ascension of Greenwich (GHA) at $t=0$
- $\dot{\theta}_E, w$ The earth's "constant" sidereal rotation rate
- λ_0 or λ_o The initial geographic longitude of the near synchronous satellite (i.e. the geographic longitude at injection)
- μ_E The earth's gaussian gravitational constant

APPENDIX B

THE EARTH GRAVITY POTENTIAL AND GRAVITATIONAL FORCE FIELD USED

The Gravity Potential used is the exterior potential derived in Reference 4 for geocentric spherical coordinates referenced to the earth's spin axis and it's center of mass (see Figure 1). The Harmonic series is truncated after J_{44} .

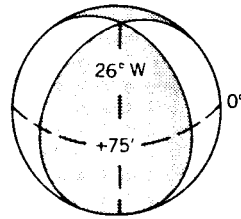
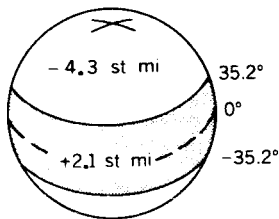
The inertial gravity constants θ_{nm} are related to the usual geographical gravity constants by:

$$\theta_{nm} = (\theta_E)_0 + \theta_E t + \lambda_{nm}$$

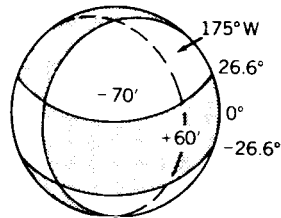
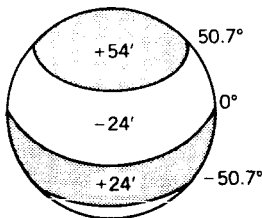
so that

$$\theta - \theta_{nm} = \lambda - \lambda_{nm} \quad (B1)$$

(see Figure 2). The longitudes and minimax deviations from an average earth sphere for each harmonic are taken from the potential of Kozai.* The J_{nm} and λ_{nm} of this potential are presented in Equations 47. The earth's gravity potential may be illustrated as follows:

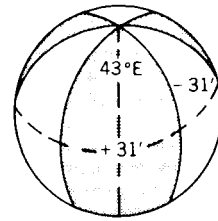
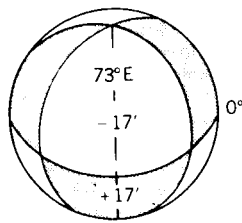


$$V_E = \frac{\mu_E}{r} \left\{ 1 - \frac{J_{20} R_o^2}{2r^2} (3 \sin^2 \phi - 1) - 3J_{22} \frac{R_o^2}{r^2} \cos^2 \phi \cos 2(\theta - \theta_{22}) \right.$$

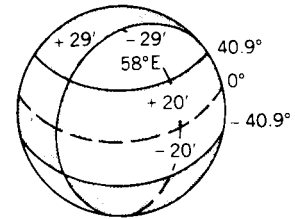
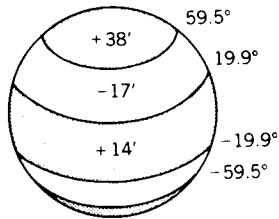


$$- \frac{J_{30} R_o^3}{2r^3} (5 \sin^3 \phi - 3 \sin \phi) - \frac{J_{31} R_o^3}{2r^3} \cos \phi (15 \sin^2 \phi - 3) \cos (\theta - \theta_{31})$$

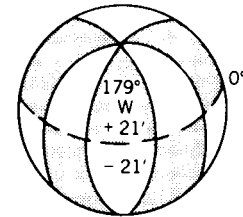
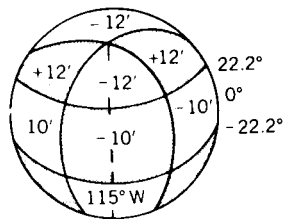
*Wagner, C. A., "The Gravitational Potential of a Triaxial Earth," Goddard Space Flight Center Document Number X-623-62-206, 1962.



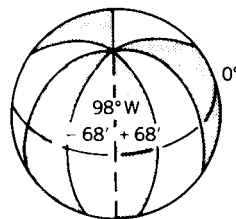
$$- 15J_{32} \frac{R_o^3}{r^3} \cos^2 \phi \sin \phi \cos 2(\theta - \theta_{32}) - 15J_{33} \frac{R_o^3}{r^3} \cos^3 \phi \cos 3(\theta - \theta_{33})$$



$$- \frac{J_{40} R_o^4}{8r^4} (35 \sin^4 \phi - 30 \sin^2 \phi + 3) - \frac{J_{41} R_o^4}{8r^4} [140 \sin^3 \phi - 60 \sin \phi \cos \phi \cos(\theta - \theta_{41})]$$



$$- \frac{J_{42} R_o^4}{8r^4} [420 \sin^2 \phi - 60] \cos^2 \phi \cos 2(\theta - \theta_{42}) - \frac{J_{43} R_o^4}{8r^4} [840 \sin \phi] \cos^3 \phi \cos 3(\theta - \theta_{43})$$



$$- \frac{J_{44} R_o^4}{8r^4} [840 \cos^4 \phi \cos 4(\theta - \theta_{44})]$$

$$\vec{F} = \hat{r} F_r + \hat{\theta} F_\theta + \hat{\phi} F_\phi. \tag{B2}$$

$$F_r = m G_r = m \frac{\partial V_E}{\partial r}.$$

where

$$\begin{aligned} G_r = \frac{\mu_E}{r^2} \left\{ -1 + \left(\frac{R_o}{r}\right)^2 \left[\frac{3}{2} J_{20} (3 \sin^2 \phi - 1) + 9 J_{22} \cos^2 \phi \cos 2(\theta - \theta_{22}) \right. \right. \\ + 2 \left(\frac{R_o}{r}\right) J_{30} (5 \sin^2 \phi - 3) (\sin \phi) + 6 \left(\frac{R_o}{r}\right) J_{31} (5 \sin^2 \phi - 1) \cos \phi \cos(\theta - \theta_{31}) \\ + 60 \left(\frac{R_o}{r}\right) J_{32} \cos^2 \phi \sin \phi \cos 2(\theta - \theta_{32}) + 60 \left(\frac{R_o}{r}\right) J_{33} \cos^3 \phi \cos 3(\theta - \theta_{33}) \\ + \frac{5}{8} \left(\frac{R_o}{r}\right)^2 J_{40} (35 \sin^4 \phi - 30 \sin^2 \phi + 3) \\ + \frac{25}{2} \left(\frac{R_o}{r}\right)^2 J_{41} (7 \sin^2 \phi - 3) \cos \phi \sin \phi \cos(\theta - \theta_{41}) \\ + \frac{75}{2} \left(\frac{R_o}{r}\right)^2 J_{42} (7 \sin^2 \phi - 1) \cos^2 \phi \cos 2(\theta - \theta_{42}) \\ \left. + 525 \left(\frac{R_o}{r}\right)^2 J_{43} \cos^3 \phi \sin \phi \cos 3(\theta - \theta_{43}) + 525 \left(\frac{R_o}{r}\right)^2 J_{44} \cos^4 \phi \cos 4(\theta - \theta_{44}) \right] \Big\} \tag{B3} \end{aligned}$$

$$F_\theta = m G_\theta = \left[\frac{m}{r \cos \phi} \right] \frac{\partial V_E}{\partial \theta}.$$

where

$$\begin{aligned} G_\theta = \frac{\mu_E}{r^2} \left(\frac{R_o}{r}\right)^2 \left\{ 6 J_{22} \cos \phi \sin 2(\theta - \theta_{22}) + \frac{3}{2} \left(\frac{R_o}{r}\right) J_{31} [5 \sin^2 \phi - 1] \sin(\theta - \theta_{31}) \right. \\ + 30 \left(\frac{R_o}{r}\right) J_{32} \cos \phi \sin \phi \sin 2(\theta - \theta_{32}) + 45 \left(\frac{R_o}{r}\right) J_{33} \cos^2 \phi \sin 3(\theta - \theta_{33}) \\ + \frac{5}{2} \left(\frac{R_o}{r}\right)^2 J_{41} (7 \sin^2 \phi - 3) \sin \phi \sin(\theta - \theta_{41}) + 15 \left(\frac{R_o}{r}\right)^2 (7 \sin^2 \phi - 1) \\ \cdot J_{42} \cos \phi \sin 2(\theta - \theta_{42}) + 315 \left(\frac{R_o}{r}\right)^2 J_{43} \cos^2 \phi \sin \phi \sin 3(\theta - \theta_{43}) \\ \left. + 420 \left(\frac{R_o}{r}\right)^2 J_{44} \cos^3 \phi \sin 4(\theta - \theta_{44}) \right\}. \tag{B4} \end{aligned}$$

$$F_{\phi} = \dot{m} G_{\phi} = \frac{\dot{m}}{r} \frac{\partial V_E}{\partial \phi},$$

where,

$$\begin{aligned}
 G_{\phi} = \frac{\mu_E}{r^2} \left(\frac{R_o}{r} \right)^2 & \left\{ -3J_{20} \sin \varphi \cos \varphi + 6J_{22} \cos \varphi \sin \phi \cos 2(\theta - \theta_{22}) \right. \\
 & - \frac{3}{2} \left(\frac{R_o}{r} \right) J_{30} (5 \sin^2 \varphi - 1) \cos \phi + \frac{3}{2} \left(\frac{R_o}{r} \right) J_{31} (15 \sin^2 \varphi - 11) \sin \phi \cos(\theta - \theta_{31}) \\
 & + 15 \left(\frac{R_o}{r} \right) J_{32} (3 \sin^2 \varphi - 1) \cos \phi \cos 2(\theta - \theta_{32}) \\
 & + 45 \left(\frac{R_o}{r} \right) J_{33} \cos^2 \phi \sin \phi \cos 3(\theta - \theta_{33}) - \frac{5}{2} \left(\frac{R_o}{r} \right)^2 J_{40} (7 \sin^2 \varphi - 3) \\
 & \cdot \sin \phi \cos \phi + \frac{5}{2} \left(\frac{R_o}{r} \right)^2 J_{41} (28 \sin^4 \varphi - 27 \sin^2 \varphi + 3) \cos(\theta - \theta_{41}) \\
 & + 30 \left(\frac{R_o}{r} \right)^2 J_{42} (7 \sin^2 \varphi - 4) \cos \phi \sin \phi \cos 2(\theta - \theta_{42}) \\
 & + 105 \left(\frac{R_o}{r} \right)^2 J_{43} (4 \sin^2 \varphi - 1) \cos^2 \phi \cos 3(\theta - \theta_{43}) \\
 & \left. + 420 \left(\frac{R_o}{r} \right)^2 J_{44} \cos^3 \phi \sin \phi \cos 4(\theta - \theta_{44}) \right\}. \tag{B5}
 \end{aligned}$$

The radial perturbation of the gravitational field referred to in this report is the residual of the sum of the gravitational and centrifugal forces on \dot{m} at the moment of injection at the radius r_s . Since \dot{m} is injected at an angular rate w for conditions where the initial perturbation rates with respect to the reference synchronous orbit are

zero, the centrifugal force on \dot{m} at the moment of this injection at r_s is $+\hat{r}(w^2 r_s \dot{m})$. Thus, the radial perturbation force, as defined above, at the moment of injection is: radial perturbation force $= \dot{m} w^2 r_s + \dot{m} G_{r_s} = \dot{m} w^2 r_s e$. Thus, the condition for the radial perturbation force (for a "perfect" injection, at injection) to be zero is: $e=0$.

APPENDIX C

MAGNITUDES OF THE LONGITUDE AND LATITUDE PERTURBATION FORCES PER UNIT ON A NEAR SYNCHRONOUS EQUATORIAL SATELLITE DUE TO GRAVITY

It may be verified from Appendix B, that:

1. The longitude perturbation force per unit mass on the near synchronous satellite is given by

$$(G_\phi)_{r_s} \doteq (g_s)(a) \quad (C1)$$

2. The latitude perturbation force per unit mass on the near synchronous satellite is given by

$$(G_\phi)_{r_s} \doteq (g_s)(i) \quad (C2)$$

where g_s is the radial acceleration of gravity in a $1/r^2$ earth field at the synchronous radius:

$$g_s = \frac{\mu_E}{r_s^2} = \frac{\mu_E}{(R_o)^2 \left(\frac{R_o}{r_s}\right)^2} = 32.15 \times .02288 = .735 \text{ ft./sec.}^2 \quad (C3)$$

It is noted that at the synchronous radius

$$\frac{\mu_E}{w^2 r_s^3} \approx 1 \quad (C4)$$

In Figure C1, the perturbation forces per unit mass are plotted for two earth gravity fields. One is for a field through 4th order due to Kozai,* and the other is for a simpler triaxial field which includes only the second order harmonic potential

constants from the same source. It is seen that, while the triaxial components clearly dominate the longitudinal perturbation force function, the maximum force in the "full" field is 12 percent greater than the maximum force in the triaxial (Table C1). The latitude perturbation in the triaxial field is, of course, zero. It is noted that the extreme magnitudes of the latitude perturbation function around the equator are about one

Table C1

Comparison of Maximum Longitudinal Perturbation Forces at Near Synchronous Altitudes Between the Triaxial and Full 4th Order Gravity Fields of Kozai.

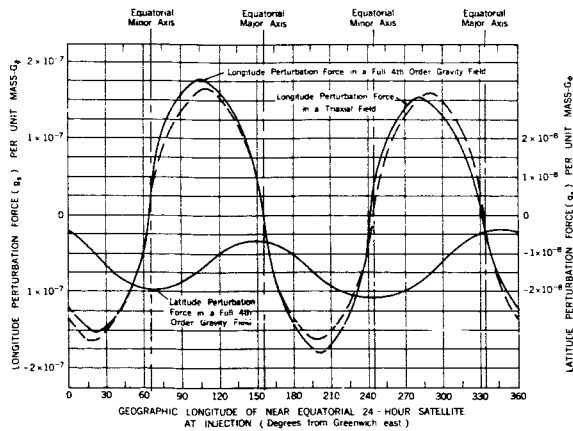


FIGURE C-1.—The longitude and latitude perturbation forces on a near equatorial 24-hour satellite around the equator, due to the geopotential of Kozai.

Field	Minimax $(G_\phi)_{r_s}$ (units of g_s)	λ_o (degrees)	Percent difference of minimax G_ϕ between full and triaxial field (full-triax/ triax) x 100
Triaxial Full	-1.65×10^{-7} -1.51×10^{-7}	18.6 21.7	+8.5
Triaxial Full	$+1.65 \times 10^{-7}$ $+1.78 \times 10^{-7}$	108.6 109.7	+7.9
Triaxial Full	-1.65×10^{-7} -1.85×10^{-7}	198.9 200.0	-12.0
Triaxial Full	$+1.65 \times 10^{-7}$ $+1.53 \times 10^{-7}$	288.9 282.3	-7.3

*Kozai, Y., Private Communication, November 1962.

order of magnitude less than the extremes of the longitude perturbation function. This fact gives justification to the assumption that the regions of stability for the near synchronous satellite may be considered to be fixed by the longitude perturbation alone (i.e., by the zeros of n). It is also interesting that the minimum of the latitude

perturbation in the full field at synchronous radius occurs close to the dynamically unstable region near the earth equator's major axis while the maximum latitude perturbation occurs near the dynamically stable region close to the minor axis. This coincidence tends to minimize the drift in the neighborhood of the major axis.

APPENDIX D

PROCEDURE FOR DETERMINING THE INJECTION RADIUS AND LONGITUDE INTO A NEAR SYNCHRONOUS ORBIT WITH MINIMAL INITIAL PERTURBATION ACCELERATIONS

The two conditions for a minimal drift, near synchronous earth orbit, are assumed to be

$$e = 0 \quad (D1)$$

$$n = 0 \quad (D2)$$

These two conditions give a set of four injection radii and longitudes for minimum drift in a near synchronous equatorial earth orbit, with the potential of Kozai.* It can be assumed that injection at the earth rate into an inclined orbit whose nodes are near these geographic longitudes at near synchronous radii, will give orbits with similarly small nodal drift if the inclination is not excessive.

Let;

$$r_s = r_o + \Delta r_s \quad (D3)$$

where $r_s \ll r_o$ is assumed and where r_o is the solution to:

$$\frac{w^2 r_o^3}{\mu_E} = 1 + \frac{3}{2} J_{20} \left(\frac{R_o}{r_o}\right)^2 - \frac{15}{8} J_{40} \left(\frac{R_o}{r_o}\right)^4 \quad (D4)$$

r_o may be thought of as the synchronous radius for the "oblate" earth. With Equations D3 and D4 into Equations D1 and D2, by ignoring all powers of $\Delta r_s/r_o$ greater than one and all terms in $J_{nm}\Delta r_s/r_o$; the synchronous conditions become

$$\begin{aligned} \frac{\Delta r_s}{r_o} = & \left\{ (K_{31})_1 \cos \lambda_o + (K_{31})_2 \sin \lambda_o + [(K_{22})_1 + (K_{42})_1] \cos 2\lambda_o + [(K_{22})_2 \right. \\ & + (K_{42})_2] \sin 2\lambda_o + (K_{33})_1 \cos 3\lambda_o + (K_{33})_2 \sin 3\lambda_o + (K_{44})_1 \cos 4\lambda_o \\ & \left. + (K_{44})_2 \sin 4\lambda_o \right\} / \frac{3w^2 r_o^3}{\mu_E} \quad (D5) \end{aligned}$$

and

$$\begin{aligned} 0 = & (K_{31})_3 \sin \lambda_o + [(K_{22})_3 + (K_{42})_3] \sin 2\lambda_o + (K_{33})_3 \sin 3\lambda_o + (K_{44})_3 \sin 4\lambda_o \\ & + (K_{31})_4 \cos \lambda_o + [(K_{22})_4 + (K_{42})_4] \cos 2\lambda_o + (K_{33})_4 \cos 3\lambda_o \\ & + (K_{44})_4 \sin 4\lambda_o \quad (D6) \end{aligned}$$

*Kozai, Y., Private Communication, November 1962

In Equations D5 and D6, the gravitational constants $(K_{nm})_i$ are

$$\begin{aligned} (K_{31})_1 &= -6J_{31} \left(\frac{R_o}{r_o}\right)^3 \cos \lambda_{31} , \\ (K_{22})_1 &= -9J_{22} \left(\frac{R_o}{r_o}\right)^2 \cos 2\lambda_{22} , \\ (K_{42})_1 &= \frac{75}{2} J_{42} \left(\frac{R_o}{r_o}\right)^4 \cos 2\lambda_{42} , \\ (K_{33})_1 &= -60J_{33} \left(\frac{R_o}{r_o}\right)^3 \cos 3\lambda_{33} , \\ (K_{44})_1 &= -525J_{44} \left(\frac{R_o}{r_o}\right)^4 \cos 4\lambda_{44} , \quad (D7) \\ (K_{31})_3 &= -\frac{3}{2} J_{31} \left(\frac{R_o}{r_o}\right) \cos \lambda_{31} , \\ (K_{22})_3 &= 6J_{22} \cos 2\lambda_{22} , \\ (K_{42})_3 &= -15J_{42} \left(\frac{R_o}{r_o}\right)^2 \cos 2\lambda_{42} , \\ (K_{33})_3 &= 45J_{33} \left(\frac{R_o}{r_o}\right) \cos 3\lambda_{33} , \\ (K_{44})_3 &= 420J_{44} \left(\frac{R_o}{r_o}\right)^2 \cos 4\lambda_{44} . \end{aligned}$$

where

$$\begin{aligned} (K_{nm})_2 &= (K_{nm})_1 \tan(n\lambda_{nm}) , \\ \text{and: } (K_{nm})_4 &= -(K_{nm})_3 \tan(n\lambda_{nm}) . \end{aligned}$$

To solve Equation D4 for the oblate earth synchronous radius r_o , let

$$r_o = r_{o1} + \Delta r_o , \quad (D8)$$

where r_{o1} is the "spherical earth" synchronous radius satisfying the expression

$$\frac{w^2 r_{o1}^3}{\mu_E} = 1 . \quad (D9)$$

If $\Delta r_o \ll r_{o1}$ is assumed, then with Equations D9 and D8 into Equation D4, by ignoring the $J_{nm}\Delta r_o/r_{o1}$ as of second order, the oblate synchronous radius is determined from

$$\frac{\Delta r_o}{r_{o1}} = \left(\frac{R_o}{r_{o1}}\right)^2 \left[\frac{3}{2} J_{20} - \frac{15}{8} J_{40} \left(\frac{R_o}{r_{o1}}\right)^2 \right] / 3 . \quad (D10)$$

The procedure to determine the elements of the "stable" synchronous orbit is:

1. Solve for the spherical synchronous radius r_{o1} from Equation D9.
2. Solve for the oblate synchronous radius r_o from the results of the solution for Δr_o from Equation D10 into Equation D8.
3. Solve for the constants K in Equation D7.
4. Solve for the "stable" synchronous injection or nodal longitudes λ_o from Equation D6.
5. Solve for the "stable" synchronous radii r_s corresponding to these longitudes from the results of the solution for Δr_s from Equation D5 in Equation D3.

For the earth potential used, the magnitude assumptions in Equations D1 and D2 are valid and the aforementioned uncoupled procedure establishes the "stable" synchronous elements to high accuracy.

IX. GENERAL

RESONANCE SCATTERING AT LYMAN-ALPHA BY AN ATOMIC HYDROGEN CELL*

J. E. BLAMONT†, P. DELACHE† AND A. K. STOBER

This paper describes the experimental conditions for direct photoelectric observation of the optical resonance and polarization of the hydrogen Lyman-alpha line (1216A).

INTRODUCTION

Rocket and satellite studies of the Lyman-alpha (1216A) line emitted in the upper atmosphere have been made using detectors which have large bandwidths, such as ion chambers with various window materials and filling gases (References 1 and 2), or photomultipliers with cesium iodide or copper iodide photocathodes (Reference 3). High-resolution photographic measurements have been made with a grating spectrograph (Reference 4) and photographic plates; the system requires long exposures and a recovery system to obtain data.

This paper is concerned chiefly with the experimental results of a hydrogen (filter) cell and ion chamber for obtaining photoelectric data.

Hydrogen Cells

Lyman-alpha is the resonance line of the hydrogen atom. If a cell containing a large number of hydrogen atoms is placed in front of the detector and is excited by an incident beam of light, the hydrogen atoms will absorb selectively the central frequency (core) of the Lyman-alpha emission from the existing beam with a width that is a function of the temperature of the atoms and their number density (optical thickness). This experiment has been performed by the Naval Research Laboratory (Reference 5). The number density of the hydrogen atoms is not very critical when the exciting spectrum is a line spectrum containing a narrow Lyman-alpha emission line. In such a cell there is a lower limit for the density of hydrogen atoms above which the cell will absorb

all the Lyman-alpha contained in the exciting beam. Such a system, a broad bandpass filter with a narrow absorption feature, can be used when the energy of the spectrum of interest is defined by the detector (1050 to 1300A for an ion chamber) and contains two parts of the same order of magnitude: the absorbable part, centered at the Lyman-alpha line with a width of 10^{-2} A; and the remainder of the spectrum from 1050-1300A, nonabsorbable. This system cannot be used to observe the Lyman-alpha radiation coming from the sun, the absorbable part being too small compared with the non-absorbable part (not more than 1 part in 100).

On the contrary, a scattering cell, where the resonance light, scattered by hydrogen atoms, is observed at 90 degrees of the incident beam, is equivalent to a very narrow band filter (about 10^{-2} A). Of course, the "transmission" in the center of the line will be small because the geometry of the cell forbids the detection of all the scattered photons, but the transmission will be exactly zero (except from the effects of stray light) for all the other wavelengths.

DESCRIPTION OF THE EXPERIMENT

The first cell was a rather complicated cell, Figure 1(a), in which the hydrogen atoms were produced by dissociation of hydrogen molecules on a hot tungsten wire. The cell contained two filaments mounted in light horns located out of direct view of the detector. Four LiF (lithium fluoride) windows were epoxy-sealed to the cell. One was the "entrance" window facing the monochromator light source, and one the "exit" window on the opposite side of the cell for observing

*Published as *NASA Technical Note D-2148*, June 1964.

†Centre National de la Recherche Scientifique, France.

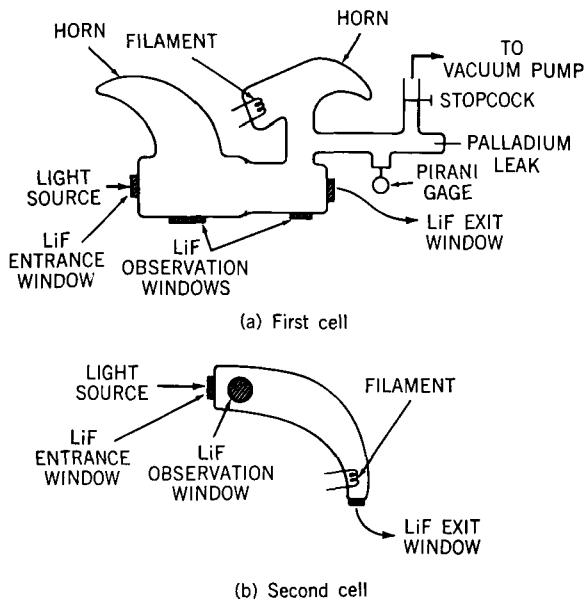


FIGURE 1.—Atomic hydrogen cells.

the direct (absorbed) light. Two other windows were mounted at right angles to the light beam: one near the entrance window, and one near the exit window.

Hydrogen was allowed to enter the cell through a palladium leak. The cell was isolated from the vacuum system by a stopcock, thus permitting evacuation and pressure regulation.

The Light Source

The light source system was a 3-meter high-dispersion vacuum monochromator illuminated by a hydrogen-discharge continuum lamp. The resolution of the monochromator was on the order of $1/20\text{\AA}$ at Lyman-alpha emission.

Experiments with the First Cell

The absorption was first observed. The first problem was the outgassing due to the large diameter of the filaments, high filament current needed, and envelope glass temperature. A good test for the "true" absorption was obtained by observing a band of the hydrogen molecule. When the heating of the filament would not create any absorption at 1234\AA , an absorption measurement would be made at Lyman-alpha.

The Lyman-alpha profile of the light coming out of the system was observed by scanning the

line with the monochromator, with and without the filament heated. This profile was already strongly self-reversed because of the hydrogen atoms at the exit port of the light source. However, a careful study of the profiles showed that this reversal was higher in the presence of hydrogen atoms in the cell. Under these conditions, even with very large slits, it was impossible to see any resonance signal at either of the two windows.

Modification of the Light Source

To have a higher intensity in the center of the Lyman-alpha line, the density of hydrogen inside the lamp was lowered by running the discharge with a mixture of helium and hydrogen. The procedure was the following:

A discharge was started in pure helium. The Lyman-alpha profile was observed with the slits of the monochromator compatible with the detector sensitivity. This may not be the Lyman-alpha line of the hydrogen atoms left in the gas of the helium tank but the H-alpha line of singly ionized helium; its intensity was very small. The hydrogen pressure was slowly increased by using a needle valve. As the intensity of Lyman-alpha increased, the line profile did not self-reverse. It then appeared that conditions changed very suddenly in the discharge lamp; the profile became strongly self-reversed, and the total intensity increased greatly. At the same time, the color of the discharge in the light source changed from orange to red. It was thus possible to determine the best conditions (maximum intensity, minimum self-reversal) from the color of the discharge.

The Second Cell

The design of the second cell, Figure 1(b), took into account the facts mentioned above.

The small rhenium wire heater filament was mounted on an aluminum oxide ring and required 5 watts for white heat. The detectors were attached directly to the windowless cell by O-rings or Apiezon wax seals.

The outgassing problem was solved by eliminating the taper joint system of the first cell and by using a secondary pump.

The stray light was removed by the use of a light horn; however, the detector at the bottom

of the horn "saw" part of the direct (absorbed) light reflected on the curved part of the horn.

The palladium leak system was the same as that used on the first cell. Pressure was measured with a thermocouple gage. Foreign gas could be introduced into the cell through a needle valve. (The experimental apparatus is shown in Figure 2).

Results Observed

With the above-mentioned conditions, both absorbed and scattered light signals were observed.

A typical recording is shown in Figure 3. On the left channel, the scattered light signal is seen and, on the right channel, the absorbed light signal.

Reading from below, we (1) start with Lyman-alpha excitation and atomic hydrogen (filament-heated); (2) close the shutter—that is, suppress the excitation—and both signals fall to zero (the zero of the left recorder is not the zero of the chart paper); (3) return to the former situation, Lyman-alpha+hydrogen; (4) stop heating the filament (no more atomic hydrogen). The "absorbed" signal rises back to the value marked Lyman-alpha, and the resonance signal falls to zero at the same time, indicating no stray light. Then we (5) close the shutter, and the absorbed signal falls again to zero while the resonant signal does not show any variation. This indicates that the filament is not seen by the detector. It can be noted at this point that another proof of the fact that there are hydrogen atoms which absorb and scatter the light, and not any outgassing, is the very short time constant of variation of the signals. We can say that it is less than the electronic time constant (0.2 second).

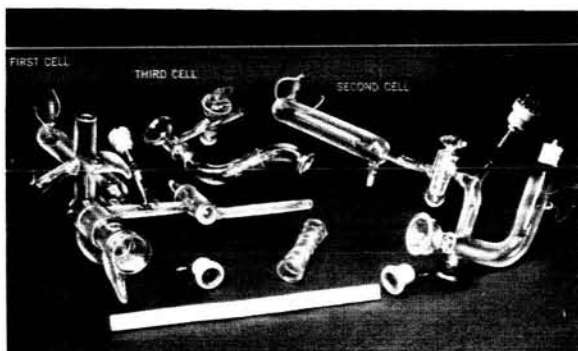


FIGURE 2.—The experimental apparatus.

0 = Zero
 $L\alpha$ = Lyman-alpha
 $L\alpha + H$ = Lyman-alpha + Hydrogen
 H = Hydrogen

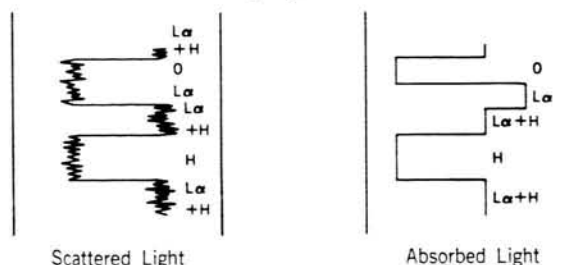


FIGURE 3.—Recording scattered light and absorbed light.

A study of the resonance signal as a function of the hydrogen pressure showed that a maximum of scattered light was obtained at the total pressure of 30 to 40 microns. Introduction of helium makes the signal drop by a factor of 2 or more, whatever the total pressure is.

Evaluation of the Number of Hydrogen Atoms in the Cell

This is a rather difficult point, since we do not know whether the density is uniform throughout the cell. However, the density can be evaluated from two experiments:

1. By constantly increasing the filament heat, the absorption reaches a plateau while the resonance signal is still increasing. This seems to indicate that the optical thickness of the whole system between the entrance window and the detector is rather high (2 or more).

2. The resonance signal increases but does not reach a maximum. However, at higher filament current the light we observe is roughly 1/1000 of the "absorbable light"; the absorbable light is that part of the light effectively absorbed as observed by the bottom detector (50 percent of the increasing light). If an optical depth of 1 is assumed, the scattered light would be 1/500 of the incident absorbable light. Since an optical depth of 1 is 10^{12} atoms/cm² it is believed that the density in hydrogen atoms is on the order of 10^{11} cm⁻³. The dissociation of hydrogen is then on the order of 10^{-4} .

Best Performance Obtained

Power dissipated in the filament for maximum resonance signal: 6 watts

Hydrogen pressure for maximum resonance signal: 30 microns

Signal-to-noise ratio: 20

A Third Cell for Observing the Polarization of the Lyman-Alpha Light

Figure 2 shows the three cells: the first cell at the left, the horn-shaped cell at the right, and a third one in the center. This cell was made to observe the polarization of the light at Lyman-alpha by reflection from a LiF crystal at the Brewster angle. A Malus-type design was adopted. We observed a reflected light strongly polarized (at least 90 percent) with a reflection coefficient of the order of 15 percent. This system will be used for measuring the polarization of the Lyman-alpha airglow.

REFERENCES

1. CHUBB, T. A., and FRIEDMAN, H., "Photon Counters for the Far Ultraviolet," *Rev. Sci. Inst.* 26(5):493-498, May 1955.
2. STOBER, A. K., "Ceramic Vacuum Ultraviolet Ion Chambers," NASA Technical Note D-1180, March 1962.
3. DUNKELMAN, L., FOWLER, W. B., and HENNES, J. P., "Spectrally Selective Photodetectors for the Middle and Vacuum Ultraviolet," *Appl. Optics* 1(6):695-700, November 1962.
4. PURCELL, J. D., and TOUSEY, R., "The Profile of Solar Lyman- α ," *Memoirs Soc. R. Sci. Liege* 4:283-294, 1961.
5. MORTON, D., and PURCELL, J. D., "Observations of the Extreme Ultraviolet Radiation in the Night Sky Using an Atomic Hydrogen Filter," *Planet. Space Sci.* 9:455-458, August 1962.

FREQUENCY AND TIME SYNCHRONIZATION AT NASA TRACKING STATIONS*

ANDREW R. CHI

Time, being an independent variable in physical measurements, is desired to be kept as constant and as synchronous with a primary time standard as possible. Since there are several time scales in use, it may be fruitful to give the definitions again.

TIME STANDARDS:

The primary standard of time is based on the orbital rotational period of the earth about the sun known as Ephemeris Time (ET). The ephemeris second,¹ is the fraction of $1/31, 556, 925.9747$ of the tropical year for 12^h ET of January 0, 1900. The ephemeris time is considered as uniform and has been compared with atomic time since 1958 and has not found any divergence to date. Another time standard, which is widely used for navigation and astronomical observations is known as Universal Time (UT). It is determined by the angular position of the earth about its axis. The unit of time, the mean solar second,² is defined as $1/86, 400$ of a mean solar day. The rate of rotation of the earth is not constant. There is a slow secular change due to tidal friction and irregular seasonal changes which may change the length of the day by as much as four milliseconds.

The observed angular position of the earth with reference to the equatorial positions of the stars is known as uncorrected Universal Time or UT-0. UT-1 incorporates the corrections applied for polar motion to UT-0; and UT-2 incorporates the corrections for seasonal variations to UT-1. These corrections are issued by the Bureau International de l'Heure (B.I.H.), Paris. UT-2 time is, in fact, an artificial kind of time scale; it is an extrapolated time, corrected and confirmed after

the observation; and in practice, is measured with a photographic zenith tube with respect to the stars. Atomic time,³ A.1., is defined as a clock which keeps time by advancing one second in the interval required for 9, 192, 631, 770 oscillations of cesium at zero field. Its epoch started at $0^h 0^m 0^s$ UT-2 on January 1, 1958, which defined the value of A.1 as $0^h 0^m 0^s$. A.1 time has been adopted temporarily as the international unit of time by the 12th General Conference of Weights and Measures in Paris on October 8, 1964.

Figure 1 shows the location of NASA tracking stations known as STADAN, station tracking and data acquisition network. The tracking of artificial satellites on a worldwide basis requires the value of universal time by practice as well as by necessity as a fact of life since the measurements are made relative to the positions of the earth. It further requires that the Universal Time be homogeneous and synchronous. It is for these reasons that Universal Time is used and requirements for precise time generated.

TIME SYNCHRONIZATION:

Time synchronization is a means of introducing an epoch. This can be done by using:

- (a) Time tick or codes in HF and VLF transmissions;
- (b) Portable clock; and
- (c) Satellite clock.

The first method has been in use with the transmission of standard frequencies in the HF band since 1923 and time tick since 1942 from NBS Radio Station WWV, Beltsville, Maryland. Due to the Doppler variations and the instability of the propagation medium for high frequency, the precision of time measurement is limited to one millisecond. New techniques, using multiple

*Presented in Session 2.13 Time and Frequency Measurements using very low frequency techniques of the 19th Annual Instrument Society of America Conference, October 12-15, 1964, New York, New York.

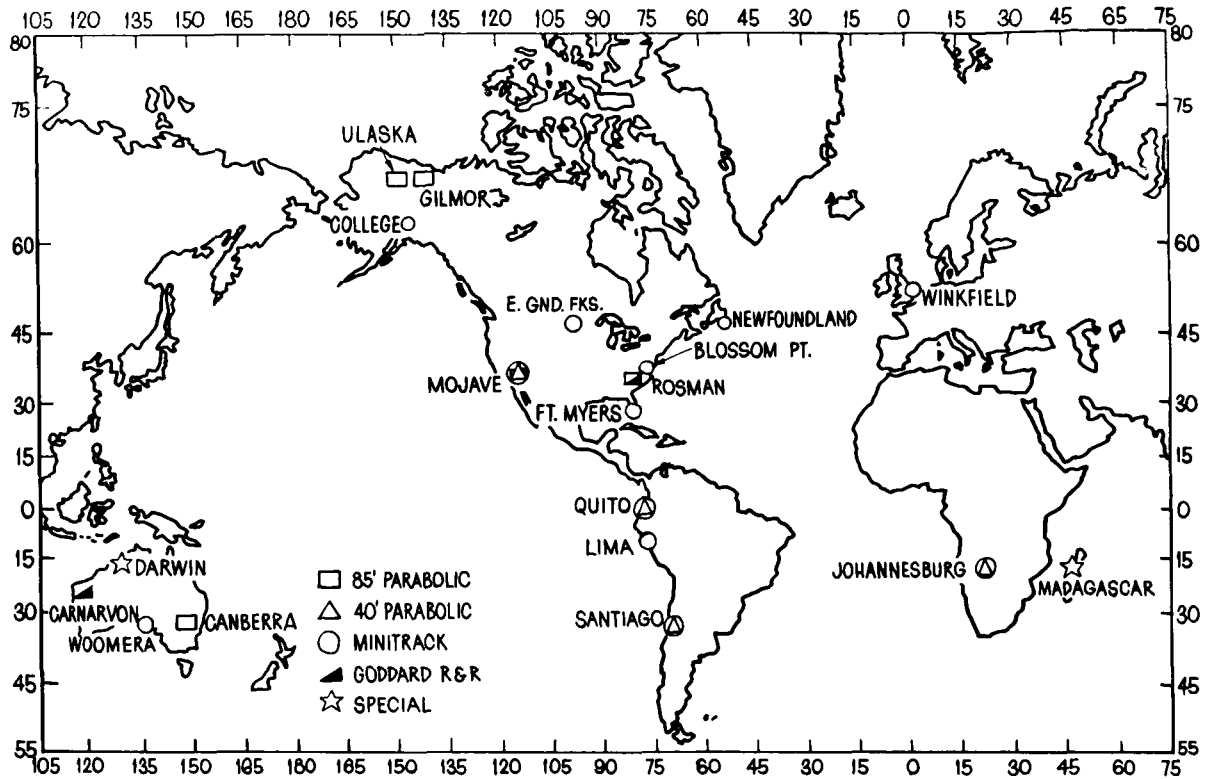


FIGURE 1.—STADAN (Station Tracking and Data Analysis Network) locations.

VLF transmissions and the highly stable characteristics of the propagation medium for VLF, have been suggested and are currently being developed.⁴ The expected capability of such techniques is in the order of microseconds. There are several other suggested techniques with the aim to resolve the cycle ambiguity of the carrier wave of VLF transmission. Such techniques, if developed, will also allow microseconds time synchronization. In addition, one can also use a transponder in a satellite to synchronize station clocks. The precision of such a technique is one microsecond if equipment delays can be calibrated. This technique has been tried with Telestar and is being tried with Syncom III and will be tried again with Relay II.

The use of a portable clock for time synchronization is not new, and is improved with the use of quartz crystal controlled oscillator and atomic frequency standards. As the new techniques for time synchronization are developed, a portable clock is an indispensable means to check the precision of the techniques and to increase accuracy.

It is also a reliable standby for introducing time to a remote station where epoch of time is not established or interrupted.

A satellite clock is another means to establish a single frame of time reference among the tracking stations. Experimental observations in the satellite as well as tracking of the satellite are referred to the satellite clock. The advantage of such a time reference is the increased opportunities for time synchronization of all tracking stations. Thus, the timing errors among the tracking stations are reduced and the overall accuracy is increased. Such problems as propagation effects in VLF transmission are also reduced. It may also become obvious that any invariant time scale may be used in a satellite clock. This technique is not yet fully developed. Under considerations are the calibration of time delay of the communication systems and techniques of coding timing information.

The present timing systems in use in the tracking stations have a capability of clock synchronization of one millisecond. The frequency stan-

dards are the conventional 100 kc/s, 1 mc/s and 5 mc/s types of quartz crystal controlled oscillators with stabilities ranging from 5×10^{-9} to about 5×10^{-10} per day. There are standby power supplies as well as oscillator redundancy in some of the tracking stations. The clocks (in the Mini-track stations) are of Goddard design and are capable of correcting timing errors with a resolution of 10 microseconds. The frequency of the oscillators are kept to 1×10^{-9} . The station clock is checked daily with WWV transmitted signals and is kept to within 1 ms. Operationally, the time correction is made in increments of 100 microseconds.

Improved timing systems are planned for the tracking stations. Prototype models are designed and developed. The new design includes: (a) all solid state circuitry, (b) redundancy with logic circuits to reject failures or deviations exceeding pre-set maxima, (c) phase comparison of standby

frequency standard with the standard in use, and (d) continuous phase tracking of the frequency standard with standard VLF transmissions.

The following six figures which I shall show, rapidly show the planned timing systems. Figure 2 shows the block diagram of the planned time standard. The phase shifter to the right of the 2.5 mc frequency standard is for fine time correction from zero to 10 microseconds continuous adjustment. The frequency divider provides the coarse time adjustment in steps of 10 microseconds up to 10 milliseconds. Figure 3 shows one commercial 2.5 mc/s oscillator and the time comparison unit. Figure 4 shows a digital clock designed at GSFC. Figure 5 shows a time code generator. Figure 6 shows the signal distribution patch panel. Figure 7 shows a station identification data generator. Figure 8 shows the preceding units together with the auxiliary measuring equipment installed in two racks. It should be borne in mind

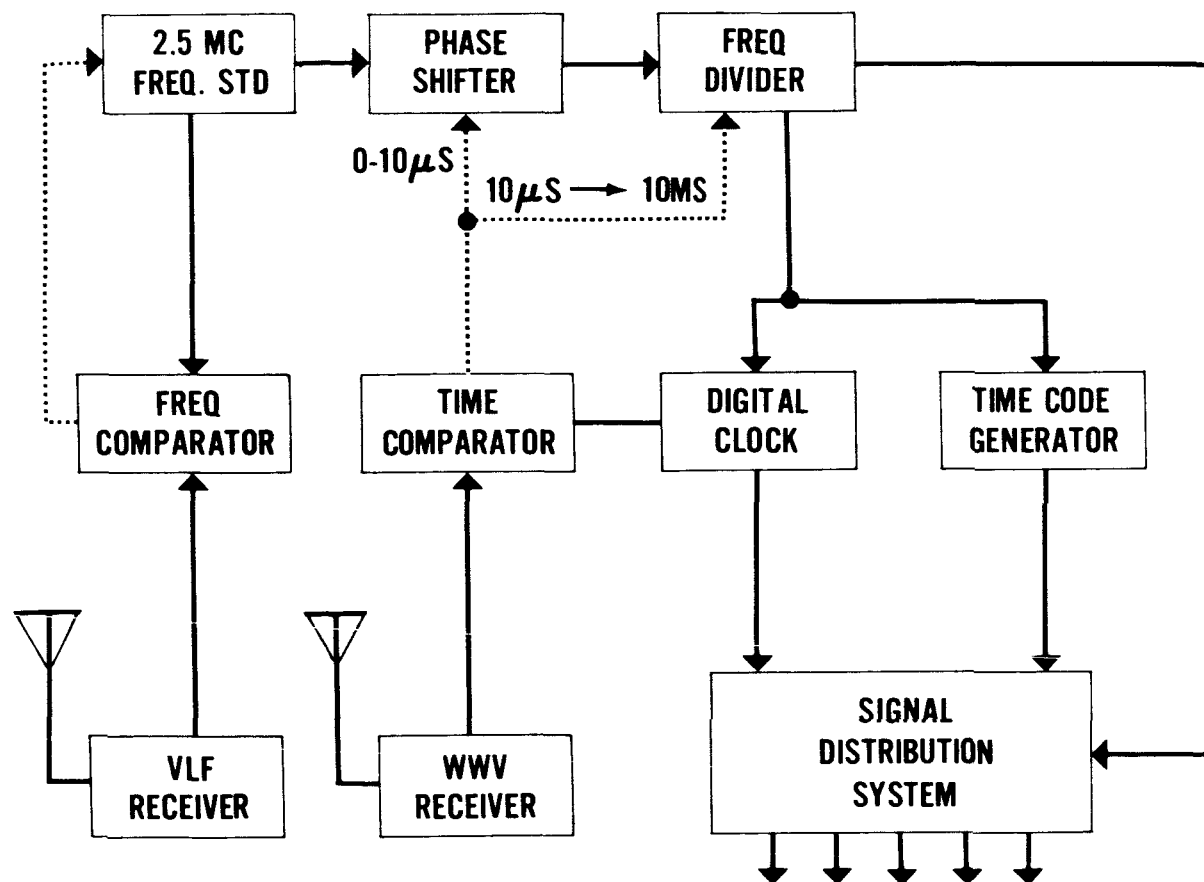


FIGURE 2.—Planned Time Standards.

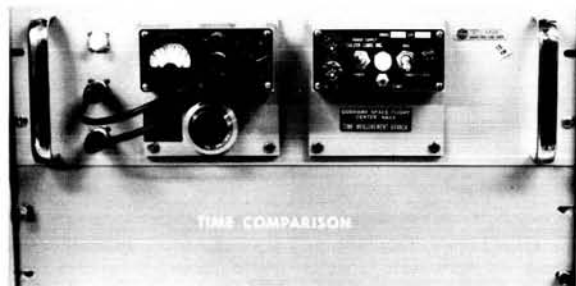


FIGURE 3.—Oscillators, 2.5Mc/s and time comparison chassis.

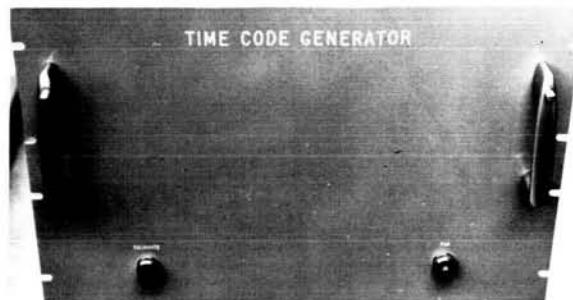


FIGURE 6.—Time code generator.

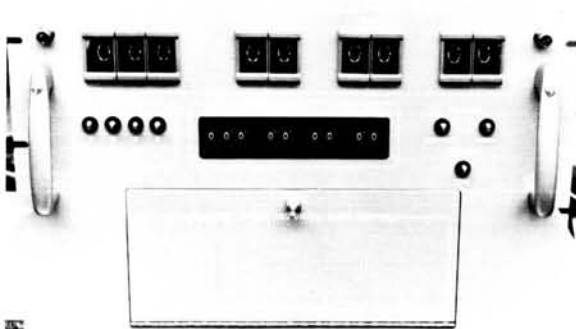


FIGURE 4.—Digital clock, NASA design.



FIGURE 7.—Station identification data generator.

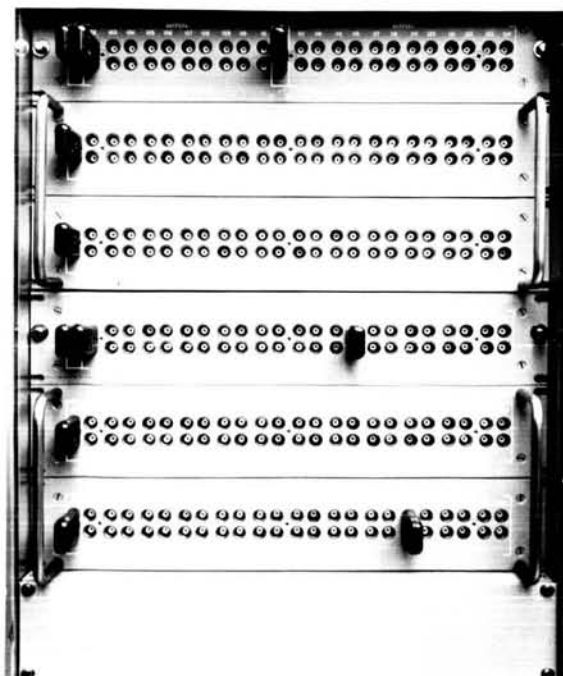


FIGURE 5.—Signal distribution patch panel.

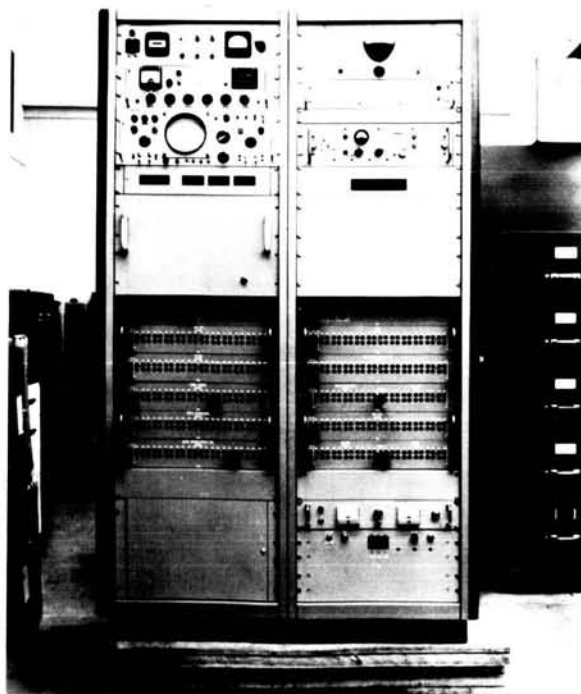


FIGURE 8.—Timing equipment racks.

that these units are results of earlier development and do not necessarily represent the future systems which are now under development.

Once the epoch of a clock is established, it remains to keep it on time by a highly precise quartz crystal controlled oscillator or an atomic frequency standard. The former suffers a frequency change with time known as aging. The latter while suffering no known frequency aging,

is subjected to physical effects such as variations of ambient temperature, humidity, magnetic field, etc., resulting in random frequency changes. Thus, it becomes necessary to synchronize a local frequency standard with a primary frequency standard. One technique which is in increasing use is the VLF transmission. There exist problems in interpretation of the propagation of VLF signals.

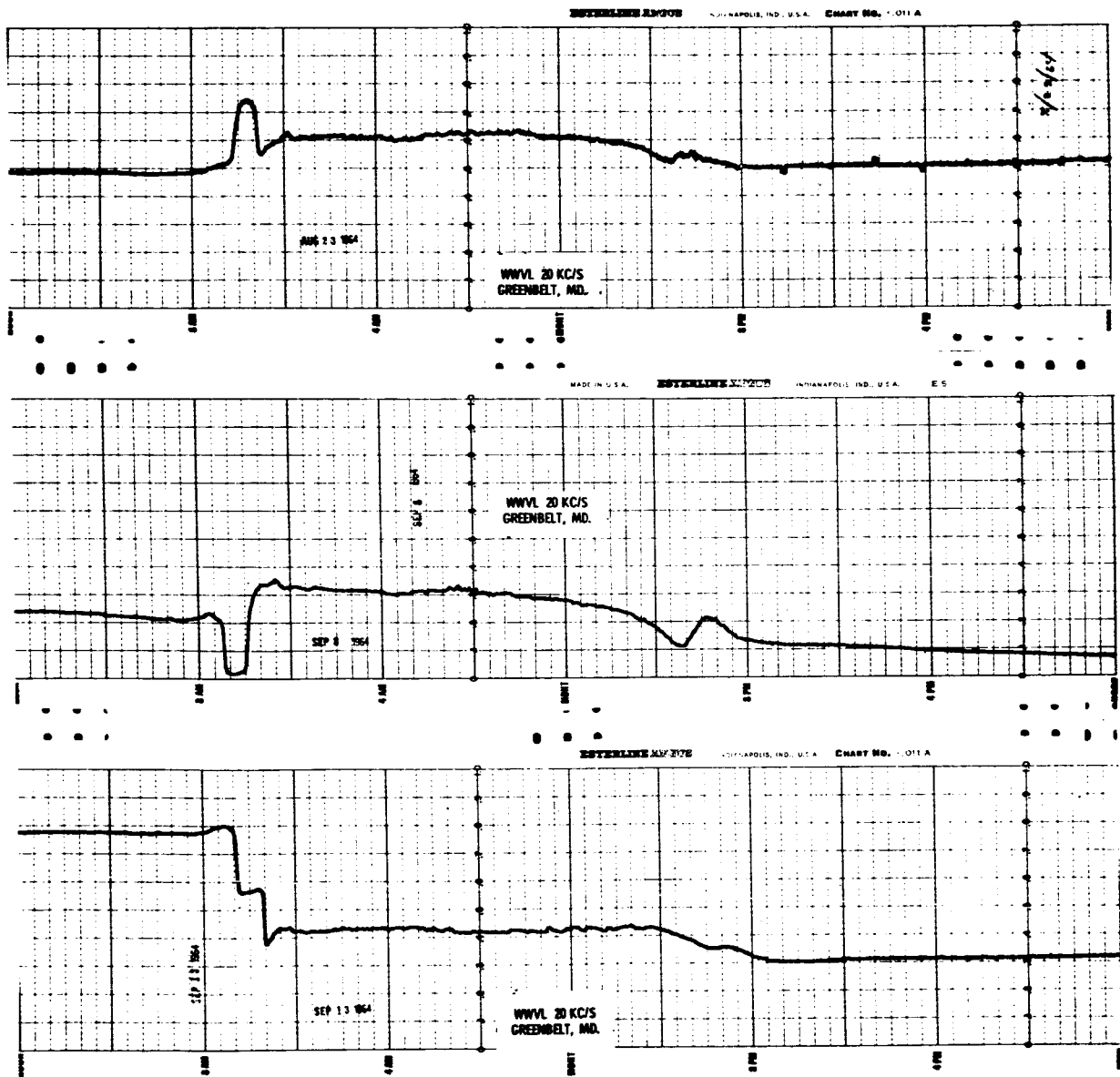


FIGURE 9.—20 Kc WWVL signal as received at NASA/GSFC, Greenbelt, Maryland, 1510 miles or 2430 km from the transmitter.

FREQUENCY SYNCHRONIZATION WITH VLF TRANSMISSIONS:

The frequency stability requirement for generating precise frequencies and time can be met by judicious selection of oscillators and control of the oscillator frequency by using a frequency stabilized VLF transmission. Such technique has been used to achieve an ultimate frequency comparison to a precision of 2×10^{-11} .⁵ Limitations which degrade this precision are the long term sta-

bility of the local oscillator, of the oscillator which controls the VLF transmitter, and the mean height variations of the reflecting medium. The diurnal phase variations also change with the distance from the transmitter, especially, at a range between 1700 and 4300 Km.⁶ These diurnal phase variations contribute errors in phase synchronization of the local frequency standards if care is not exercised.

Figure 9 shows the phase record of 20 kc/s

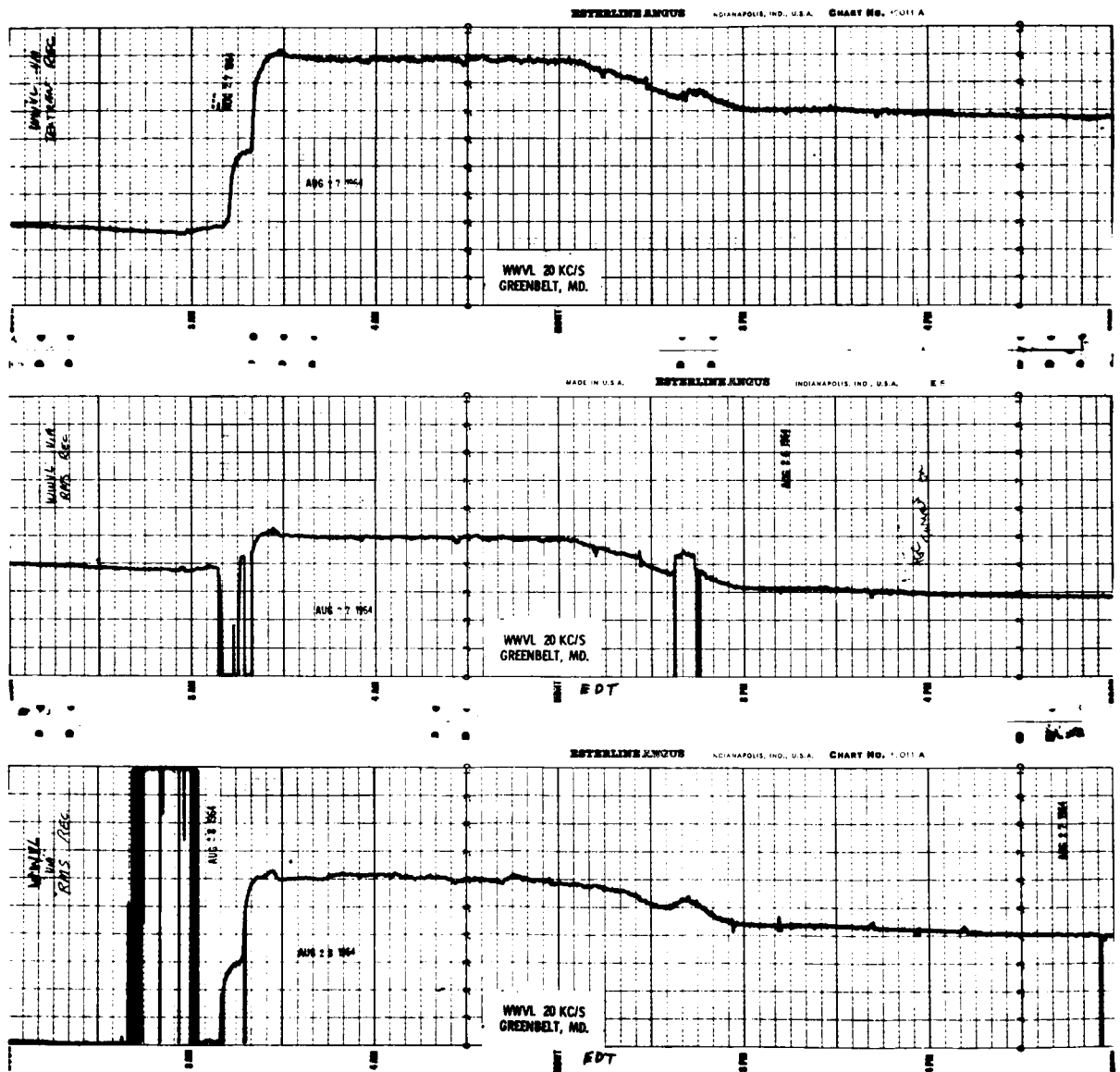


FIGURE 10.—20 Kc WWVL signal as received at NASA/GSFC, Greenbelt, Maryland, by a single mode, of propagation, during sunrise two-hops.

transmission from NBS Radio Station WWVL at Fort Collins, Colorado, as received at NASA/GSFC, Greenbelt, Maryland. In this figure the phase delay increases from the bottom to the top with a full scale of 100 microseconds and the time increases from right to left as marked in Eastern Daylight Time. The top trace is a typical phase record and is the most often seen. The two lower traces are seen occasionally. The difference is the phase change observed during sunrise. The phase change as shown in the top and center traces is tentatively interpreted as the resultant of two signals, the first and the second-hop sky waves, nearly at antiphase. The increase in phase delay due to the increase of the ionospheric height from day to night is in the range of 1.25 to 1.88 radians or 10 to 15 microseconds. The rapid phase change of about 20 microseconds observed during sunrise as shown in the top and center traces is then according to the tentative interpretation as the interference of the second-hop sky wave by the first-hop sky wave resulted from an increase in the signal strength of the latter. It has been observed that at certain times the attenuation for the first-hop sky wave is higher than that for the second-hop. This is especially so during sunrise and results in a phase record showing a step phase-advance as shown by the bottom figure and the next figure, Figure 10. The top

and center traces of the phase records are the outputs from two commercial VLF phase tracking receivers (Textran, now known as Tracor Corp., Model 599-cs and RMS Engineering Inc., Model 1312) tuned to 20 Kc/s. The two receivers were used to remove any possible malfunction of the phase lock-loop in the VLF receiver. In the center trace of this figure, the phase lock-out for the receiver signal level was set too high in the RMS receiver and resulted in a lock-out of the servo motor and zero recorder output when the signal level was low. This was corrected although not completely as shown in the bottom trace (of this slide). The phase change during sunrise shows that a single mode, two-hop sky wave transmission, is definitely predominant. The frequency of occurrence of these observations, single mode, two-hop sky wave propagation, is 6, 5, and 5 during July, August, and September of 1964 respectively or about 20 percent of the time during the three months observed. The phase advance due to a single two-hop sky wave is about 7.54 radians or .60 microseconds. The phase advance is measured from the mean night-time phase to that after sunrise. The net phase change measured before the sunset of the previous day and after the sunrise of the next day is about one cycle of the carrier wave or 50 microseconds.

Comparisons of 20 Kc/s diurnal phase records

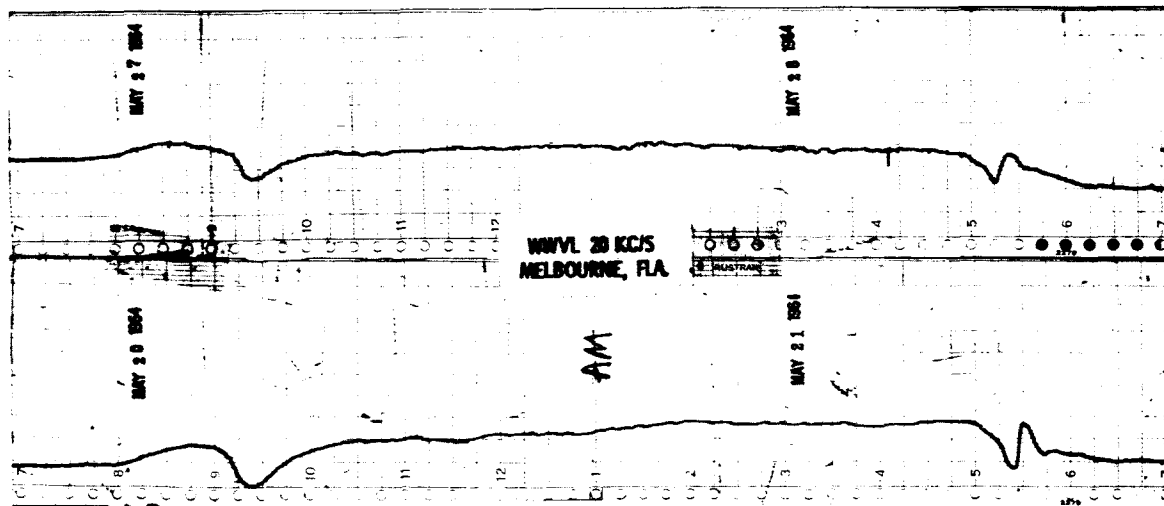


FIGURE 11.—20 Kc WWVL signal as received at Melbourne, Florida, 1550 miles or 2490 km from the transmitter (by courtesy of Mr. Marvin S. Friedland of Geo Space Corps.)

as received at GSFC with those observed at a location about the same distance from the transmitter showed good agreement. Figure 11 shows the phase records of the same 20 Kc/s signal from WWVL as received at Geo Space Corporation, Melbourne, Florida. This record is supplied through the courtesy of Mr. Marvin S. Friedland of the Geo Space Corporation. The time scale in this figure is from left to right and the phase delay is still increasing in the upward direction. Melbourne, Florida, is about 1550 miles or 2490 kilometers from Fort Collins, Colorado, where the

WWVL transmitter is located. Greenbelt, Maryland, is about 40 miles closer to the transmitter than Melbourne. Greenbelt is 1510 miles or 2430 kilometers from the transmitter. Both NASA/GSFC and Geo Space Corporation observed the diurnal phase variations of the 20 Kc/s signal due to mixed modes propagation of the first- and second-hop sky waves.

At a shorter distance than 1500 miles from the transmitter one should then expect to receive the single-hop propagated 20 Kc/s waves. Such is the case as shown in the next figure, Figure 12

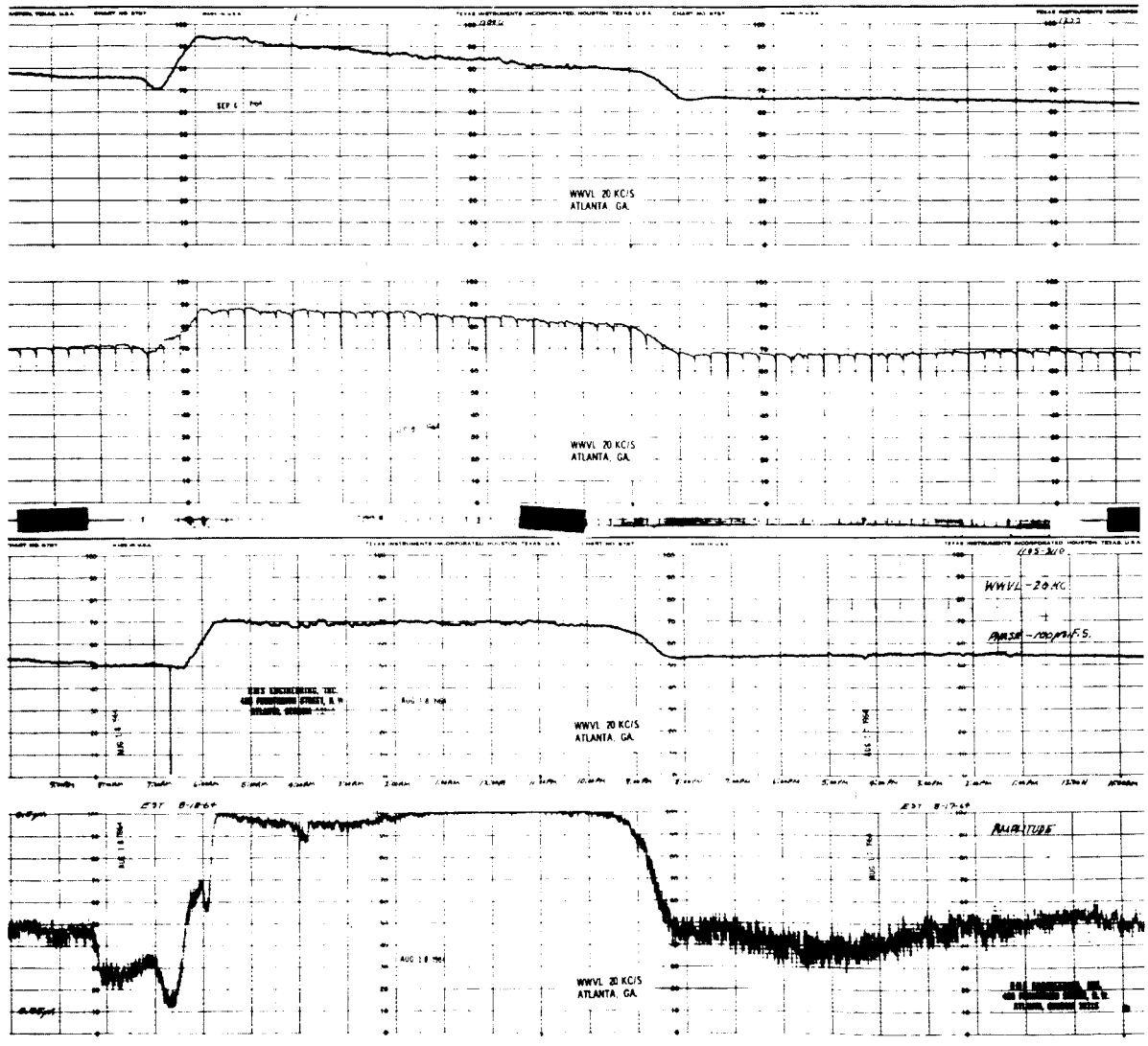


FIGURE 12.—20 Kc WWVL signal as received at Atlanta, Georgia, 1206 miles or 1950 km from the transmitter (by courtesy of Mr. W. Whittier Wright of RMS Engineering, Inc.)

The records in this figure are furnished through the courtesy of Mr. W. Whittier Wright of the RMS Engineering, Inc., Atlanta, Georgia. Atlanta, Georgia, is about 1206 miles or 1940 kilometers from the WWVL transmitter. In this figure, the time increases from right to left again. The increase in phase delay is still upward. In this figure, reading from top to bottom, the first and the third traces are phase records; the second

is, however, one cycle or 50 microseconds for the phase record as shown by the top trace. The bottom trace is the corresponding amplitude record. It can be seen that this phase record agrees with that shown in the previous figure.

The results of these observations are being applied in a continuing program at GSFC in cooperation with the National Bureau of Standards Boulder Laboratories to determine the

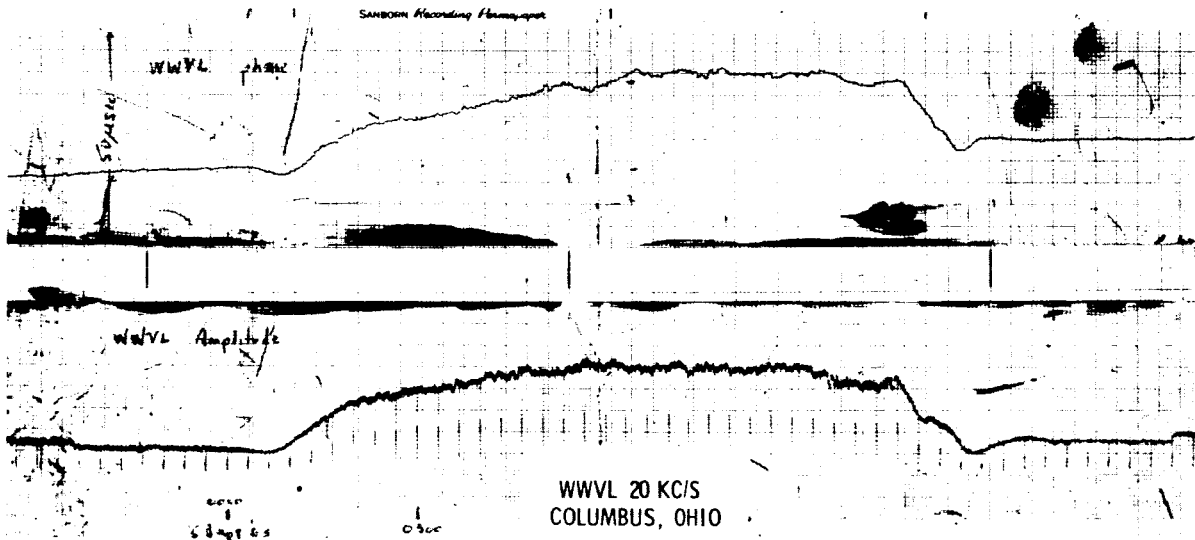


FIGURE 13.—20 Kc WWVL signal as received at Columbus, Ohio, 1170 miles or 1880 km from the transmitter (by courtesy of Dr. William A. Teso of Ohio State University).

and the fourth are the corresponding amplitude records. One can see the typical trapezoid diurnal phase records of a VLF signal in the first and third traces. The trapezoid shaped phase record is considered as the phase of a transmitted wave by a single mode of propagation; in this case, a single-hop transmission. Another phase record of 20 Kc/s signal recorded at Ohio State University, Columbus, Ohio, is shown in the last figure, Figure 13. These records are furnished through the courtesy of Dr. William A. Teso of Ohio State University. Columbus, Ohio, is about 36 miles closer to the WWVL transmitter than Atlanta. Columbus is 1170 miles or 1880 kilometers from the transmitter. In this figure the time increases from left to right and the phase delay is still increasing in the upward direction. The full scale

feasibility of using VLF transmission to synchronize time in a world-wide basis.

REFERENCES

1. The ephemeris second is so defined in 1956 by the International Committee of Weights and Measures.
2. U.S. Naval Observatory Circular No. 49, March 8, 1954.
3. MARKOWITZ, W., HALL, R. G., ESSEN, L., and PARRY, J. V. L. *Phys. Rev. Ltrs.* 1, 105, 1958. Also *TIME Service Notice No. 6* dated January 1, 1959, issued by U.S. Naval Observatory, Washington, D.C., 20390.
4. (a) MORGAN, A. H., "A New Method of Time Signal Modulation and Demodulation of VLF Carriers," NBS Report 7236, National Bureau of Standards Boulder Laboratories, July 20, 1963.
(b) ECHOLS, J. D., and STONE, C. S., "Optimum Utilization of Existing Timing Techniques," presented at ISA Conference, Huntsville, Alabama, April 18, 1963.

- (c) LOONEY, C. H., JR., "VLF Utilization at NASA Satellite Tracking Stations," Journal of Research NBS 68D, 1., January 1964.
- (d) NBSBL Contract with NASA/GSFC Contract No. S-49285G.
5. PIERCE, J. A., WINKLER, G. M. R., CORKE, R. L., Comparison Between Cesium-controlled Oscillators," Nature 187, pp. 914-916, 1960.
6. CHILTON, C. J., CROMBIE, D. D., JEAN, A. G., pp. 265-267 and discussion by W. T. BLACKBAND, pp. 286-290, Chapter 19, "Phase Variations in VLF Propagation," AGAR Dograph 74: "Propagation of Radio Waves at Frequencies Below 300 Kilocycles," Pergamon Press.
- The GBR Experiment "A Trans-Atlantic Frequency

PRELIMINARY INVESTIGATION OF A RADIOISOTOPE FUELED THERMIONIC DIODE POWER SYSTEM FOR THE 0.5 to 0.3 KW POWER RANGE REQUIREMENTS ASSOCIATED WITH ELECTRIC PROPULSION DEVICES*

DENNIS F. HASSON

The use of radioisotopes as thermal sources in a thermionic diode power system has been investigated. The power system was envisioned to meet the requirements of electric propulsion devices in the 0.5 to 3.0 KW range. The diode loss and output, fuel block thermal, and system weight analyses utilized to obtain the system efficiencies and system performances (watts per pound) are presented herein. Also included are the resultant costs and electrical power availability analyses from the above information. The analyses indicate that the best isotope fuel from a technical feasibility viewpoint is plutonium primarily because the other competitive fuels are eliminated due to the heavy shielding weights associated with radiological safety requirements. The costs for plutonium compounds, however, are prohibitive in the 0.5 KW to 3.0 KW range and the present and planned availabilities are hardly equal to a requirement of 0.5 KW. Further, it is indicated that direct radioisotope heating of a propellant in the thruster might be a more practical and feasible approach to propulsion requirements.

ACKNOWLEDGMENTS

The author wishes to express appreciation to Mr. K. I. Duck who assisted with the calculations and Mr. R. Weitzel who provided information on the power conditioning aspects of the diode power system for their assistance in preparation of this paper.

DEFINITION OF TERMS

A = isotope atomic mass, gm

A' = constant in Equation (B-5), $120 \frac{\text{amps}}{\text{cm}^2 - \text{°K}^2}$

A_{eff} = isotope effective cross sectional area of fuel block, cm^2

B = buildup factor, dimensionless

C = constant - dimensionless

D = conversion constant, $7.75 \times 10^5 \frac{\text{watts} - \text{dis} - \text{days}}{\text{mev}}$

D_E = emitter diameter, cm

D_F = fuel block diameter, cm

D_R = dose rate, rem/hr

$D_{R \text{ req.}}$ = required dose rate level at 1 meter, rem/hr

E_{AV} = average energy per disintegration, $\frac{\text{mev}}{\text{dis}}$

h_{cs} = cesium heat conduction coefficient, $\frac{\text{watts}}{\text{cm}^2 - \text{thousands } \text{°K}}$

*Published as *Goddard Space Flight Center Document X-623-64-169*, June 1964.

- J = diode output current density, $\frac{\text{amp}}{\text{cm}^2}$
 K_1 = thermal shield loss constant, $\frac{\text{watts}}{\text{cm}^2}$
 K_2 = thermal shield edge loss constant, $\frac{\text{watts}}{\text{cm}}$
 k_F = molybdenum fuel block heat conduction coefficient $\frac{\text{watt}}{\text{cm}^2-\text{°C}}$
 L_F = fuel block length, cm
 ℓ = shield thickness for 100 watt-thermal, cm
 ℓ_s = shield thickness for present cases, cm
 N_o = Avogadro number, $6.0247 \times 10^{23} \frac{\text{atoms}}{\text{gm atomic weight}}$
 P_D = diode power density, $\frac{\text{watts}-\text{electrical}}{\text{cm}^2}$
 P_{d_i} = isotope power density, $\frac{\text{watts}-\text{thermal}}{\text{gm}}$
 P_{d_c} = isotope compound power density, $\frac{\text{watts}-\text{thermal}}{\text{cm}^3}$
 P'_{d_c} = effective isotope compound power density (See Equation (C-1)), $\frac{\text{watts}-\text{thermal}}{\text{cm}^3}$
 P_{out} = system electrical output power, watts
 P_{TD} = thermionic diode output power, watts
 Q_{cs} = cesium conduction loss, watts
 $Q_{\text{dc-dc}}$ = DC-DC converter loss, watts
 $Q_{E-C_{\text{rad}}}$ = emitter-collector radiation loss, watts
 Q_{ec} = electron cooling loss, watts
 Q_{ss} = structural support and lead loss, watts
 Q_{ts} = thermal shield loss, watts
 S = system specific power, $\frac{\text{watts}}{\text{lb}}$
 T = decay time, days
 $T_{\frac{1}{2}}$ = half-life of radioisotope, days
 T_{bs} = temperature of bottom shield, °K
 T_C = collector temperature, °K
 T_{cs} = cesium reservoir temperature, °C
 T_E = emitter temperature, °K
 T_s = temperature of side shield, °K
 t = time, sec
 V_J = fuel block volume, cm³
 V_{out} = output voltage, volts
 x = distant from isotope source, cm
 W_D = weight of diode, structural support and leads, lbs.
 $W_{\text{dc-dc}}$ = weight of DC-DC converter, lbs.
 W_F = weight of fuel block, lbs.
 ΔT_L = longitudinal temperature difference, °K
 ΔT_R = radial temperature, difference, °K
 ϵ = experimental value of $\frac{1}{\frac{1}{\epsilon_R} + \frac{1}{\epsilon_C} - 1}$, dimensionless

ϵ_E = emitter total emissivity evaluated at T_E , dimensionless

ϵ_C = collector total emissivity evaluated at $\sqrt{T_E - T_C}$, dimensionless

η_s = system efficiency, percent

η_{dc-dc} = converter efficiency, percent

ϕ = neutron flux rate, $\frac{\text{neutron}}{\text{cm}^2\text{-sec}}$

ϕ^* = equivalent emitter work function from Richardson equation at T_E and J , ev

μ = neutron attenuation coefficient, cm^{-1}

ρ_I = isotope density, $\frac{\text{gm}}{\text{cm}^3}$

ρ_F = molybdenum fuel block density, $\frac{\text{gm}}{\text{cm}^3}$

ρ_{Pb} = lead density, $\frac{\text{gm}}{\text{cm}^3}$

σ = Stephan-Boltzmann constant

for Equation (B-1) $5.679 \times 10^{-12} \frac{\text{watts}}{\text{cm}^2 - \text{°K}^4}$

for Equation (B-5) $8.6167 \times 10^{-5} \frac{\text{ev}}{\text{°K}}$

INTRODUCTION

Electric propulsion devices have been proposed to provide the attitude control and station keeping functions of long term (3 years) synchronous satellites. It is generally felt that the electric power requirements for these devices fall in the range of 0.5 to 3.0 KW for the subject mission (See, e.g., Reference 1). Further analysis indicated that a radioisotope-fueled high temperature thermionic diode would offer the highest system efficiencies and best performance, watts per pound of system as the power generator in a power system. The geometry of the opposing emitter-collector surfaces of a thermionic diode can either be planar (two opposing disks) or cylindrical (two opposing cylinders). The selection of the geometry, however, is determined by the power density of the radioisotope heat source. The cylindrical diode configuration will not operate satisfactorily with isotopes of power density of less than about 12 thermal watts per cubic centimeter, and this eliminates the low and medium power density isotopes which have some of the most desirable radiological safety characteristics. Further, the most favorable operating conditions for the cylindrical diode would be met with the short life, very high power density isotopes which would not meet the mission life requirements of the present task. Thus, this

investigation was performed for a thermionic diode with the planar geometry, and the results of the investigation are given in this report. Some comparisons with solar cell and nuclear thermoelectric power systems are also included.

The investigation has been performed in the following manner: first, radioisotope physical properties, cost and availability data have been updated and evaluated; second, the thermal power losses in both the fuel block and thermionic diode of the power generator have been considered in some detail; and last, system weights have been estimated. The results of the latter two parts of the analysis were integrated with the radioisotope data to provide the performance data presented here.

SYSTEM DESCRIPTION

The system under study consists of a radioisotope fueled high temperature thermionic diode power generator connected electrically in series with a DC-DC converter, which in turn is connected electrically directly to the load. The thermionic diode power generator with the diode in the planar geometric configuration consists of a circular emitter opposite a collector of the same geometry separated by a narrow gap into which cesium vapor is injected under pressure. The operation of a cesium filled thermionic diode can

be summarized as follows: first, the temperature of the emitter is raised sufficiently by the radioisotope heat source to the point where electrons are emitted from the surface of the emitter, i.e., the work function of the emitter is exceeded; and secondly, the cesium vapor in the gap reduces the space charge which builds up as more and more electrons are freed, thus allowing the electrons to be set in motion toward the collector; and hence, the thermionic diode acts as a power generator. The DC-DC converter takes the output of the thermionic diode generator (which is inherently a low voltage device) and raises its voltage to some

required load value. The radioisotope heat source is joined thermally to the emitter by a heat conduction path in the form of a molybdenum cylinder. The radioisotope fuel consists of cylindrical pellets inserted into many small diameter holes drilled in the molybdenum block. This combination is called the "fuel block" in this paper. The fuel block is surrounded by thermal shields to reduce heat losses by heat radiation to space. A schematic of the above system is shown in Figure 1.

The efficiency of the system is defined by the following equation:

$$\eta_s = \frac{P_{out}}{\underbrace{Q_{ts}}_{\text{Fuel block losses}} + \underbrace{Q_{E-C_{rad}} + Q_{ss} + Q_{cs} + Q_{ec}}_{\text{Thermionic Diode Losses}} + \underbrace{Q_{dc-dc} + Q_{out}}_{\text{Thermionic Diode Output, } P_{TD}}} \quad (1)$$

The system performance parameter, specific power (watts output to load per pound of system weight), is defined by Equation (2) below.

$$S = \frac{P_{out}}{W_D + W_{dc-dc} + W_F} \quad (2)$$

For the purpose of the present analysis a specific size of power supply (50 watts-electrical output to the load) was selected. This power supply consists of one of a pair of back-to-back diode power generators of a 100 watt electrical output module. This pair of power generators was connected in series to give at least 1.0 volt DC output or greater to the DC-DC converter which was assumed to have an 80 percent conversion efficiency at this voltage. Thus, the power generator used in the analysis had an output of 62.5 watts electrical. The fuel block was designed to be 57.7 percent fuel and 42.3 percent molybdenum by volume. Further, the fuel block was covered laterally and on the side of the diode by multi-layers of thin tantalum sheet, except at the interface of the fuel block for the other half of the back-to-back diode of the 100 watt module.

ANALYSIS

The analysis is divided into several sections to expand on the quantities in Equations (1) and (2), and also to introduce the summary of the radioisotope data and subsequently the method of

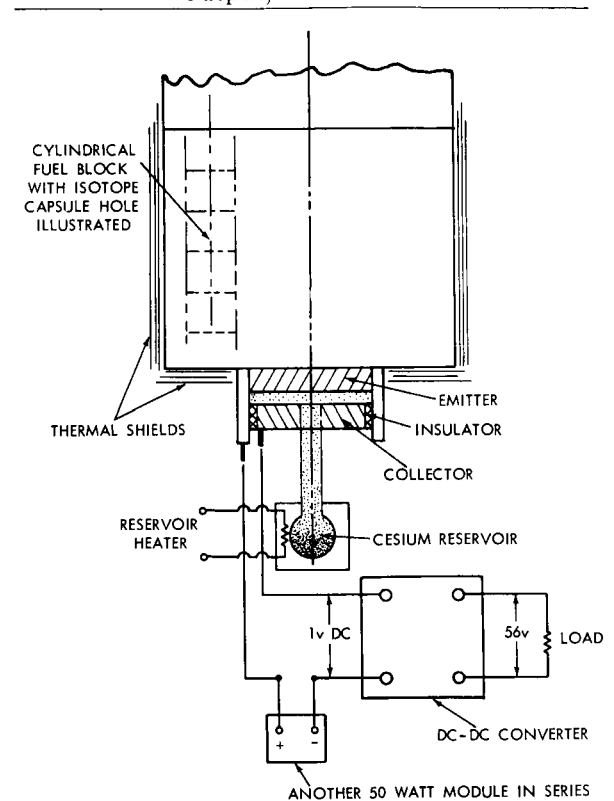


FIGURE 1.—Schematic of isotope fueled thermionic power supply system.

computing the isotope cost and the electrical availability data. A summary of the cases and parameters investigated is given in Table I.

TABLE I.—Summary of Cases and Parameters Studied for Thermionic Power Supply Systems. (All Cases Evaluated for 1 and 3 Years*)

Fuel	$T_E, ^\circ K$		
	1700	1790	1984
SrTiO ₃	X		
CeO ₂	X		
Tm ₂ ¹⁷⁰ O ₃	X		
Tm ₂ ¹⁷¹ O ₃	X		
PuO ₂	X	X	X
Pu-Alloy.....	X	X	X
Cm ₂ ²⁴⁴ O ₃	X		

*Except Pu²³⁸ compounds where all information quoted corresponds to 3 year mission data.

A. Summary of Radioisotope Data

Data on the physical properties, cost, and availability of radioisotopes useable as heat sources are given in Table II. In the present

study, several isotopes were rejected before performing the analysis, these are: Cobalt 60—due to its extreme radiological hazards and low melting point; Cesium 137—due to hazard, low power density and low melting point; Promethium 147—low power density, low melting point and high cost; Actinium 227, Thorium 228 and Uranium 232, due to their low availability and predicted high cost. Also excluded were Polonium 210 and Curium 242 because of short half life. The power density values shown in the table for the isotopes used in the present investigation were calculated by the method indicated in Appendix A.

Although the mission time requirement of this report is three years, the analysis below was performed for mission times of one and three years essentially to illustrate the fall off in performance of the shorter half life isotopes considered in the analysis. The resulting data should be considered as bonus information however not directly applicable to the stated requirements of the present report.

TABLE II.—Data on Radioisotopes Useable as Heat Sources

Isotope	Properties								
	Half-life (yr)	Isotopic purity %	Compound form	Density of compound (gm/cc)	Power Density of compound (at end of 1 year) watts/cc	Decay Mode & Energy (Mev)		Melting point of Compound (°c)	Curies per watt (th)
						Primary	Secondary		
Co ⁶⁰	5.3	10	Metal	8.9	15.5	β(.31)	γ(1.33)	1480	65
Sr ⁹⁰	28	24	SrTiO ₃	4.6	1.14	β(.54)		~1900	149
Cs ¹³⁷	30	35	Glass	3.2	0.215	β(.52)	γ(.662)	~1275	209
Ce ¹⁴⁴	0.78	80	CeO ₂	6	12.6	β(.31)	γ(.034)	~1950	135
Pm ¹⁴⁷	2.7	95	Pm ₂ O ₃	6.6	1.8	β(.23)	γ(.12)	1027	2400
Tm ¹⁷⁰	0.348	100	Metal	9.32	19.95	β(.96)	γ(.084)	1525	500
Tm ¹⁷⁰	0.348	100	Tm ₂ O ₂	8.6	12.5	β(.96)	γ(.084)	~2400	500
Tm ¹⁷¹	1.9	100	Metal	9.32	1.86	β(.097)	γ(.067)	1525	5600
Tm ¹⁷¹	1.9	100	Tm ₂ O ₃	8.6	0.95	β(.097)	γ(.067)	~2400	5600
Po ²¹⁰	0.378		Metal	9.4	1320	α(5.3)	γ(.8)	~170	31
Ac ²²⁷	21.2		AcO	8.15	110	α(4.94)	β(.043)	~1600	34
Th ²²⁸	1.9	95	ThO ₂	9	1270	α(5.42)	γ(.08)	3050	31
U ²³²	74	85	UO ₂	10	33	α(5.32)	γ(.057)	>2000	31
Pu ²³⁸	86.4	80	PuO ₂	11.46	4.4	α(5.49)	γ(.044)	~2000	30
Pu ²³⁸	86.4	80	Metal	17	9.3	α(5.8)	γ(.043)	635	30
Cm ²⁴²	0.444		CmO ₂	8.4	1000	α(6.11)	γ(.044)	2240	28
Cm ²⁴⁴	17.6	90	Cm ₂ O ₃	11.75	27	α(5.8)	γ(.043)	~1500	29

TABLE II.—Data on Radioisotopes Useable as Heat Sources (Continued).

Isotope	Production Capability [kilowatt(th)/yr]			Isotopic Fuel Costs [\$/watt(th)]		
	Present	Planned	Future	Present	Planned	Future
Co ⁶⁰	Large ^a			32		
Sr ⁹⁰	20.13	33.56	67.12	447	149	19.58
Cs ¹³⁷		48.00 ^b	110.00		104	
Ce ¹⁴⁴		27.30	78.00	64	64	0.92
Pm ¹⁴⁷		11.00	25.00		485	
Tm ¹⁷⁰			40.00 ^c	50 ^d		10
Tm ¹⁷⁰			40.00	50		
Tm ¹⁷¹			3.57	1100		400
Tm ¹⁷¹			3.57	1100		400
Po ²¹⁰		14.00	14.00		190	
Ac ²²⁷					1000 ^e	
Th ²²⁸					100	
U ²³²						
Pu ²³⁸	1.16	2.90	11.56	1285	1285	695
Pu ²³⁸	1.77	4.42	17.70	895	895	695
Cm ²⁴²		9.50	9.50		165	
Cm ²⁴⁴	0.46	3.45	Large		4350	1100

NOTES:

- (a) Co⁶⁰, Sr⁹⁰, Ce¹⁴⁴, Pu²³⁸ (dioxide), Pu²³⁸ (metal), and Cm²⁴⁴ production capability and isotopic fuel costs were obtained from preliminary information provided by R. W. Shivers, Production & Materials Branch, Division of Isotopes Development, United States Atomic Energy Commission, 10/17/63.
- (b) Cs¹³⁷, Pm¹⁴⁷, Po²¹⁰ and Cm²⁴² production capability and isotope fuel costs were obtained from *Nucleonics* article "Isotope Coats and Availability," Vol. 21, No. 3, March, 1963 authored by Harold L. Davis.
- (c) Tm¹⁷⁰ and Tm¹⁷¹ production estimates indicated in Report No. TE-2026-120, "Analyses of Costs in the Production of Tm-170 and Tm-171," November 1963, ISO/Serve Inc., Cambridge, Mass. and Thermo Electron Eng. Corp., Waltham, Mass.
- (d) Tm¹⁷⁰ and Tm¹⁷¹ cost estimates verified by W. K. Eister, Production & Materials Branch, Division of Isotopes Development, United States Atomic Energy Commission, in a private communication, 9/16/63.
- (e) Ac²²⁷ and Th²²⁸ cost estimates from T. Slosek, General Electric Company, Vallecitos, via private communication, 8/22/63.

B. Thermionic Diode Power Generator Losses Output

The thermionic diode power generator losses treated include those associated with the emitter-collector radiation, structural support and leads, cesium conduction and electron cooling. The power generator output has been considered to be fixed in all cases at 62.5 watts electrical before entering the DC-DC converter, whose efficiency allowed a 50 watt-electrical system power output.

The sum of all these losses and the diode power

generator output for one of the cases investigated are shown in Figure (2) as a thermal power flow diagram for the system under study. Also shown are the equations which were used to calculate the losses. A more detailed analysis of the diode losses is given in Appendix B. Below is a compilation of the total diode losses for the 62.5 watts thermionic diode and the experimental diode power densities and output voltages which set the emitter sizes in this study. The approximate corresponding emitter-collector spacing is indicated.

T_E , °K	Thermionic Diode Losses plus P_{TD} , Watts	P_D , watts/cm ²	V_{out} , volts	Emitter-Collector Spacing, Mils
1700	674.94	3	0.5	12
1790	533.89	6.9	0.5	4
1984	419.90	10	0.8	2

E. Cost

The cost per 50 electrical watts output deliverable to the load was determined by multiplying the thermal watts required per 50 electrical-watt module by the cost per thermal watt given in Table II. These cost data with the sources noted represent the best available information.

F. Availability

The availability in electrical-kilowatts was calculated, based on the 50 watt system analysis by multiplying the system efficiency by the corresponding isotope thermal kilowatt availability shown in Table II. The availability data with sources noted represent the best available information suitable for the engineering judgments made herein.

G. Other Power Systems

Data were obtained on solar cells and radio-isotope fueled thermoelectric power systems (References 2 and 3, respectively) in order to provide a comparison with the radioisotope fueled thermionic devices treated in this paper. Actual hardware data were obtained on the solar cells for the Eccentric Orbiting Geophysical Observatory, while the data obtained on the thermoelectrics corresponds to predictions based on the most current development data.

PRESENTATION OF RESULTS

The results are presented in the following figures and tables:

	<i>Figure</i>
System Efficiency versus Isotope Power Density	3
Loss, Percent of Total versus Isotope Density	4
(a) Mission time 1 year	
(b) Mission time 3 years	
System Specific Power versus Isotope Power Density	5
(a) Mission time 1 year	
(b) Mission time 3 years	
Available Electric Power, KW, versus Emitter Temperature for Plutonium Compounds	6
	<i>Table</i>
Variation of Cost per 50-watts-electrical output (P_{out}) for the Various Isotopes Listed in Increasing Value of Power Density	III
Variation of Available Electric Power, KW, for the Various Isotopes Listed in Increasing Value of Power Density	IV

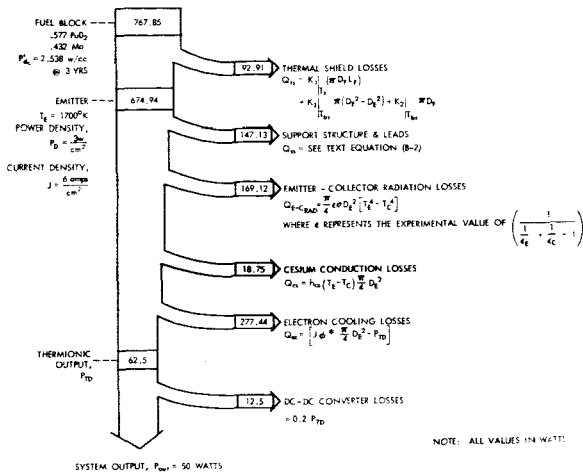


FIGURE 2.—Thermal power flow diagram for PuO₂-fueled-thermionic power supply system (Rhenium emitter-molybdenum collector.)

C. Thermal Shield Losses

The thermal shield losses are computed after the diode losses, since the diode losses and the output are used to initially size the fuel block. After the fuel block is initially sized, thermal shield losses corresponding to this size are determined. The sum of the diode losses, power generator output, and this initial thermal shield loss are then used to determine a new fuel block requirement to supply this amount of thermal power. Further iterations could be performed, but it was found that even the loss associated with the first iteration could be compensated for, if it were considered permissible to increase the percentage of the isotope fuel in the fuel block. The details of this analysis are given in Appendix C.

D. System Weight

The system weight in this analysis includes three components namely; W_D —the diode weight including support structure and leads, W_{dc-dc} —the DC-DC converter weight; and W_F —the fuel block weight. The various items of system weight are described in more detail in Appendix D. Not included in the system weight calculated therein is the weight of radiological safety shielding, but for the fuels requiring it estimates have been made, considering only the radiation from the lateral direction of the cylindrical fuel block. The equations and methods used are given in Appendix E.

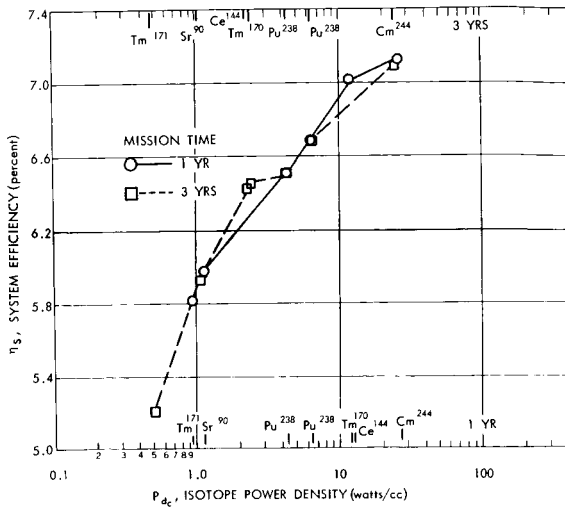


FIGURE 3.—System efficiency versus isotope power density.

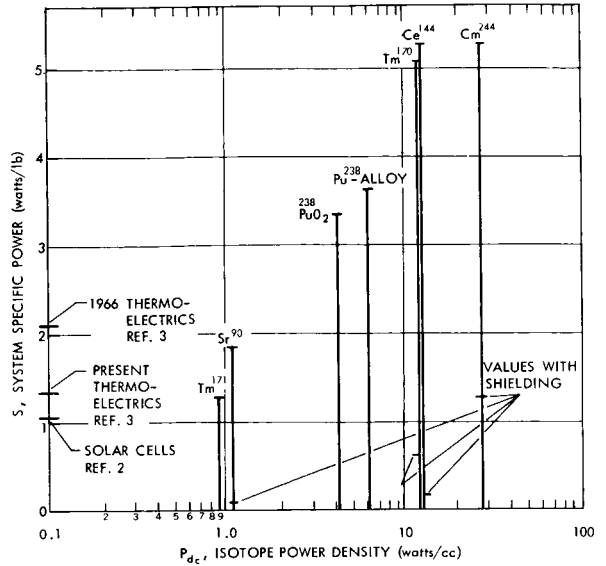


FIGURE 5a.—Mission time 1 year system specific power versus isotope power density.

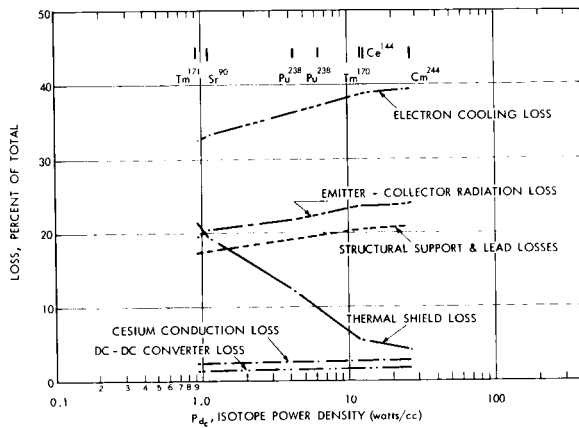


FIGURE 4a.—Mission time 1 year loss, percent of total versus isotope power density.

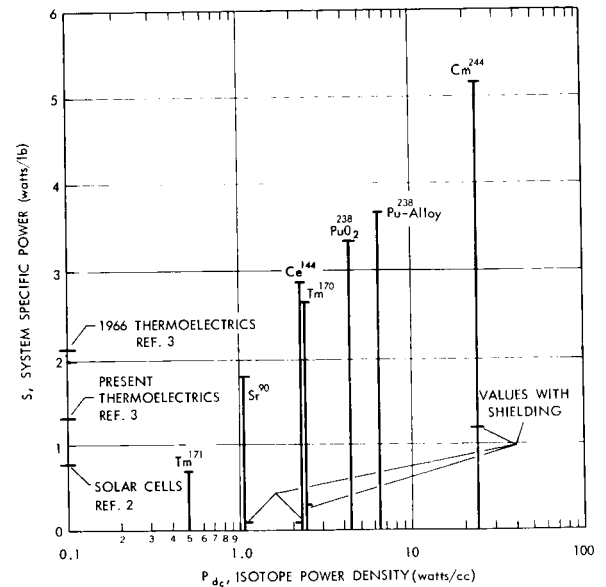


FIGURE 5b.—Mission time 3 years system specific power versus isotope power density.

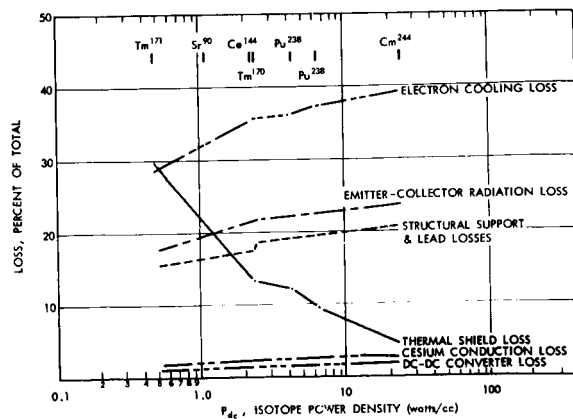


FIGURE 4b.—Mission time 3 years loss, percent of total versus isotope power density.

DISCUSSION OF RESULTS

The discussion follows the figures presented above. It should be pointed out that all the figures correspond to an emitter temperature of 1700°K unless otherwise noted. This emitter temperature was chosen since it is representative of state-of-the-art materials capabilities. Fur-

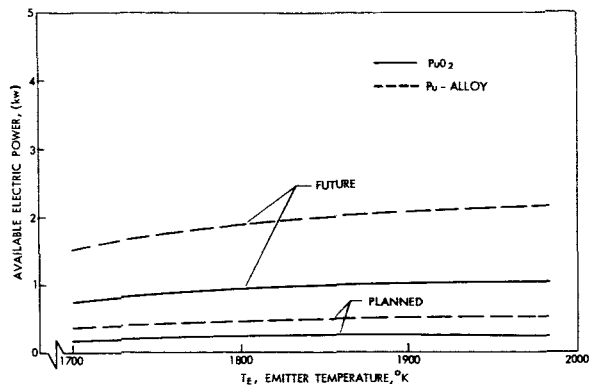


FIGURE 6.—Available electric power, KW versus emitter temperature for plutonium compounds.

ther, it should be noted that the remarks below are in general applicable to the three year mission time except as noted.

System Efficiencies

The system efficiencies for the various isotopes (See Figure 3) show a variation from about 5 to 7 percent with the higher power density isotopes yielding the higher efficiencies. This result is attributable to reduced thermal shielding requirements, because the higher power density fuel required a smaller volume, hence smaller surface

area fuel block. The efficiencies appear low at first glance for high temperature thermionic devices. These efficiencies include the fuel block and converter efficiencies which explains why a diode of efficiency about 10 percent (corresponding to our case for $T_E=1700^\circ\text{K}$) is reduced to the values given above, with the variation in the final result mainly a function of the fuel block efficiency. The percentage contribution of the various losses is shown in Figure 4 to illustrate the effect of isotope density. The previous remark is illustrated by noting that the percent loss for thermal shielding is smaller for the higher power densities.

System Performance

In Figure 5, the system performance parameter, watts per pound, is shown for the various isotopes considered for one and three year missions, respectively. The three highest power density isotopes, at first glance namely, thulium 170, cerium 144 and curium 244 appear to offer the highest performance for the 1 and 3 year missions. If, however, radiological hazards associated with these isotopes are considered (if the requirement to attenuate the dose rate to 10 mrem/hr at one meter must be fulfilled) drastic reductions in the performance occur as shown in the figure. To investigate the effect of relaxing this criterion, a

TABLE III.—Variation of Cost per 50 Watts-Electrical Output (P_{out}) for the Various Isotopes Listed in Increasing Value of Power Density

Isotope	Present		Planned		Future	
	Mission Time		Mission Time		Mission Time	
	1 yr.	3 yr.	1 yr.	3 yr.	1 yr.	3 yr.
Tm ¹⁷¹946	1.058	-----	-----	.344	.385
Sr ⁹⁰375	.377	.125	.126	.016	.017
Pu ²³⁸ (Oxide).....	.986	.986	.986	.986	.768	.768
Pu ²³⁸ (Alloy).....	.669	.669	.669	.669	.520	.520
Tm ¹⁷⁰036	.007	-----	-----	.039	.008
Ce ¹⁴⁴046	.050	.046	.050	.001	.001
Cm ²⁴⁴	-----	-----	3.059	3.071	.703	.706

NOTES:
 (a) Emitter temperature for all cases is 1700°K.
 (b) Units in millions of dollars.

TABLE IV.—Variation of Available Electric Power, KW for the Various Isotopes Listed in Increasing Value of Power Density

Isotope	Present		Planned		Future	
	Mission Time		Mission Time		Mission Time	
	1 yr.	3 yr.	1 yr.	3 yr.	1 yr.	3 yr.
Tm ¹⁷¹ -----					.21	.20
Sr ⁹⁰ -----	1.20	1.19	2.00	1.19	4.01	3.97
Pu ²³⁸ (Oxide)-----	.08	.08	.19	.19	.75	.75
Pu ²³⁸ (Alloy)-----	.15	.15	.38	.38	1.52	1.52
Tm ¹⁷⁰ -----					2.80	2.58
Ce ¹⁴⁴ -----			1.91	1.76	5.47	5.02
Cm ²⁴⁴ -----	.03	.03	.25	.24	*	*

NOTES:

(a) Emitter temperature for all cases is 1700°K.

* Predicted to be large.

level was set at 100 millirem/hr at one meter for the Thulium 170 case (1 year mission). This computation showed an increase in the shielded watts per pound value from 0.61 to 0.83. While this is a gain, it falls short of the original value of 5.1 and makes the system not competitive, even with the solar cells. The above results essentially leave the plutonium compounds as the best from a watts per pound standpoint and inherently (even though not stated previously), from a hazard viewpoint. It should be noted that the present approach shows some improvement over solar cells and radioisotope fueled (fuel is plutonium 238 for data shown in this report) thermoelectrics, but the improvement is not as large as previously anticipated.

Cost

The cost results (Table III) show that some of the more hazardous isotopes are less costly than plutonium. The smaller cost, however, may not be realizable because of other cost resulting from safety handling requirements. In general, (excepting Strontium-90) the cost would be high compared to solar cells (approx. $\$0.015 \times 10^6$ per 50 watts elec.). The cost for the plutonium compounds, even at the future price for the highest power density fuel, would be excessive for

0.5 KW-electrical (Approximately $\$5.2 \times 10^6$ compared to $\$0.15 \times 10^6$ for solar cells).

Available Electrical Power

The available electric power, Table IV, shows some rather disappointing results; namely, that for almost all the isotopes, the present and planned availability will not satisfy the minimum requirements of 0.5 KW. Further, the future availability for plutonium just barely satisfies this requirement and will never satisfy the upper limit of the requirement, 3 KW electrical. The power availability values indicate fairly clear that radioisotope fueled thermionics will never have a wide application as a basic power unit for a multitude of space missions. A comment might be made about the fact that if the hazard problem can be accepted, a way to satisfy the power availability requirement is to use either strontium or cerium for the long and short duration missions, respectively.

General

Since the plutonium offers the best useable performance and inherent lack of hazard, it is of interest to investigate methods to alleviate the problems of cost and availability. One approach to increase the availability and reduce the cost of

plutonium is to increase the emitter temperature. The variation in availability with emitter temperature for the plutonium compounds is given in Figure 6. Although there is some improvement in the availability with increasing emitter temperature, the planned production just meets the minimum requirement of 0.5 KW and the future availability is still about 30 percent short of the maximum requirement. Thus, the capability of isotope fueled thermionics in general to satisfy the 0.5 to 3.0 KW requirement at a reasonable cost and at a performance level (watts per pound) far in excess of other power systems such as solar cells and isotope fueled thermoelectrics appear to be somewhat short of the goal. However, for specific missions such as one of a spacecraft traveling in a direction of reduced solar constant or one in which the spacecraft is in an orbit in the most intense region of the radiation belt, radioisotope fueled thermionics might offer a solution to the power system requirement if the power requirement is in the range of 0.5 KW. The answer to the initial problem, however, of providing a power system for the requirements of electrical propulsion devices for attitude control and station keeping does not exist in the radioisotope fueled thermionics at present, and more investigative effort into possible systems to meet this requirement should go on. An answer to the problem, however, could be provided presently if the electric thrust system could be replaced by a competitive one. Such a system is a direct radioisotopically fueled heater in the thruster to avoid the inherent losses in the thermal-electrical-thermal conversion of the present system. This latter method would also appear to be more compatible with the availability and cost picture for radioisotopes at present.

CONCLUSIONS

The conclusions pertinent to the various parameters considered herein are given as follows:

(a) System Efficiencies

The system efficiencies vary from 5 to 7 percent for an emitter temperature of 1700°K. These values reflect the reduction in the usual quoted thermionic efficiency by the DC-DC converter and fuel block thermal shield losses with the

variance in the efficiency due to isotope power density.

(b) System Performance

The specific power data, watts/pound, show significant gains over such systems as solar cells and thermoelectrics, but this apparent advantage in reality only exists for the plutonium fuels because the others require very heavy shielding for radiological safety considerations. Thus, the plutonium fuels show the only real potential for isotope fueled thermionics.

(c) Cost

The isotopes with radiological safety hazards are less costly than plutonium, where the cost for a 0.5 KW electrical plutonium fueled unit would be $\$5.2 \times 10^6$. This appears prohibitively high except for missions where the cost might be written off against requirements unattainable by any other means.

(d) Available Electric Power

The available electric power shows even a more drastic picture when plutonium is considered the only suitable fuel. Planned availability, even when higher emitter temperatures are considered, is barely sufficient for providing 0.5 KW of electrical power. Future availability would provide at best about 2 KW and the upper bound of the task, 3 KW, could never be met.

(e) General

Other than for specific missions such as traveling away from the sun or traveling in an orbit in the radiation belts, radioisotope fueled thermionic power systems do not appear practical from a cost viewpoint or feasible from an isotope availability standpoint for any known 0.5 to 3.0 KW electrical power supply requirement. One scheme to stay within isotope availability and at reasonable cost by reducing some of the conversion losses is to use direct radioisotope heating of the propellant in the thruster.

REFERENCES

1. ISLEY, W., "Investigation of Attitude Stabilized Spacecraft in the 24 Hour Synchronous Orbit." Technical Report TR-2, NASA-Goddard Space Flight Center Internal Report, June 11, 1962.
2. HARDIS, L., GSFC, "EGO Electrical Power System and Power System Weights," private communication, 1964.

3. EPSTEIN, J., GSFC, "Radioisotope Fueled Thermoelectrics," private communication, 1964.
4. BOWES, W. M., and BURKHALTER, P. G., "Energy Conversion Systems Reference Handbook," Volume XI, Radioisotope System Design, p. B-13, WADD Technical Report 60-699, September 1960.
5. JAKOB, M., and HAWKINS, G. A., "Elements of Heat Transfer," p. 229, Third Edition, John Wiley and Sons, 1957.
6. DUNLAY, J., Thermo Electron Engineering Corporation, private communication, 1963.
7. First Quarterly Report for the Thermionic Emitter Materials Research Program (TEE-4015-1) for the period, 1 July 1961—30 September 1961, Thermo Electron Engineering Corporation.
8. KITRILAKIS, S. S., MEEKER, M. E., and RASOR, N. S., "Annual Technical Summary Report for the Thermionic Emitter Materials Research Program," 1 July 1961—30 June 1962, Contract NONR-3563 (00) (TEE 4015-3), Report No. 2-63, Thermo Electron Engineering Corporation.
9. "Metals Handbook," Volume 1, 8th Edition, p. 1216, American Society for Metals.
10. STROMINGER, D., HOLLANDER, J. M., and SEABORG, G. T., "Tables of Isotopes," Review of Modern Physics, p. 829, April 1958.
11. EROZOLIMSKII, B. G., et. al. "Physics of Nuclear Fission," Chap. 6, p. 78, Pergamon Press, 1958.
12. ROCKWELL, T. III, "Reactor Shielding Design Manual," D. Van Nostrand Company, Inc., 1956.
13. "500 Watt Generator Development Program," TEE-P-1006, Thermo Electron Engineering Corporation, 1962.

APPENDIX A

ISOTOPIC COMPOUND POWER DENSITY CALCULATION

The power density, P_{dc} , of the various isotopic compounds was calculated in the following steps:

Step 1—Isotope Powder Density, P_{di} :

The isotope power density is given by the following equation (see reference 4):

$$P_{di} = \left(\frac{DE_{AV}}{AT_{1/2}} \right) \exp \left[-0.693 \left(\frac{T}{T_{1/2}} \right) \right] \quad (A-1)$$

The average energy, E_{AV} , for the isotopes considered in this report are given below:

Isotope	Mode of Decay	E_{AV} , Mev
Sr ⁹⁰ -----	β	2.237
Tm ¹⁷⁰ -----	β	0.335
Tm ¹⁷¹ -----	β	0.030
Ce ¹⁴⁴ -----	β	0.414
Cm ²⁴⁴ -----	α	5.880
Pu ²³⁸ -----	α	5.486

It should be noted that for the β emitters an integration of the continuous emission spectra must be performed to obtain E_{AV} .

Step 2—Isotopic Purity:

The isotope power densities from the first step are multiplied by a constant factor to reflect the fact that the fuel after irradiation and processing is not always 100 percent pure radioactive. A list of this multiplication factor is given below:

Isotope	Constant
Sr ⁹⁰ -----	0.24
Tm ¹⁷⁰ -----	1.00
Tm ¹⁷¹ -----	1.00
Ce ¹⁴⁴ -----	0.80
Cm ²⁴⁴ -----	0.80
Pu ²³⁸ -----	0.80

Step 3—Isotope Compound Power Density, P_{dc} :

The value of power density from Step 2 is multiplied again by the ratio of the radioisotope mass to the compound mass to account for chemical dilution in the compound form. A list of these constants is given below. After these multiplications, the resultant values are multiplied by the corresponding densities in Table II to yield the isotope compound power density, P_{dc} in watts per cubic centimeter. These final power density values are also given below.

Isotope Compound	Constant	P_{dc} , Power Density, Watts/cm ³	
		after 1 year	after 3 years
SrTiO ₃ -----	.484	1.14	1.07
Tm ₂ ¹⁷⁰ O ₃ -----	.876	12.54	2.40
Tm ₂ ¹⁷¹ O ₃ -----	.877	.95	.49
CeO ₂ -----	.818	12.61	2.37
Cm ₂ O ₃ -----	.910	27.60	24.93
PuO ₂ -----	.881	4.52	4.48
Pu-Alloy-----	-----	6.63	6.50

It should be noted that the Pu-Alloy data were estimated by taking 0.7 of the value for the pure metallic nondiluted radioisotope.

APPENDIX B

DIODE LOSSES AND OUTPUT ANALYSIS

A discussion of the various diode losses and the output along with the equations and data utilized to determine them where applicable follows.

1. Emitter-Collector Radiation Losses

This loss is the radiation heat transfer between the emitter and collector and the governing equation corresponding to the condition of two materials of different emissivities is given from Reference 5.

$$Q_{E-c_{rad}} = \frac{\pi}{4} \epsilon \sigma D_E^2 [T_E^4 - T_C^4] \quad (B-1)$$

where ϵ represents an experimental value of $(1/\epsilon_E + 1/\epsilon_c - 1)^{-1}$ for the particular emitter-collector materials (Rhenium-Molybdenum) operating at temperature of about 1700°K. (If materials handbook values for the emitter and collector material emissivities were used in this quantity, the resultant value would be about one-half of the experimental value.) It is noted that use of the handbook values would not correspond to actual conditions because the emissivities are quoted for rather ideal conditions such as polished surfaces, for example. The value for ϵ used in this report (0.18) was obtained from Reference 6.

2. Structural Support and Lead Losses

Although these losses can be fairly accurately determined when the detail design is known, the present analysis had to resort to a rule of thumb. The rule used was to take 87 percent of the emitter-collector radiation losses, because this

percentage was representative of other lysosana available to the author. Thus

$$Q_{ss} = 0.87 Q_{E-c_{rad}} \quad (B-2)$$

3. Cesium Conduction Losses

This is the loss of heat from emitter by the mechanism of conduction through the cesium vapor. The equation for this loss is given simply by

$$Q_{cs} = h_{cs} (T_E - T_C) \frac{\pi D_E^2}{4} \quad (B-3)$$

The quantity h_{cs} , the heat conduction coefficient, was obtained from Reference 7, (Figure 3-10), for cesium reservoir temperature corresponding to the various emitter temperatures investigated. Below is a short list of these values:

$T_E, \text{ }^\circ\text{K}$	$T_{cs}, \text{ }^\circ\text{C}$	$h_{cs}, \left[\frac{\text{watts}}{\text{cm}^2 - \text{thousands } ^\circ\text{K}} \right]$
1700.....	277	1
1790.....	302	2.5
1984.....	332	3.5

4. Electron Cooling Losses

This quantity represents an irreversible process in which a portion of the energy carried from the emitter to the collector by the electrons is converted to heat. That is, a portion of the energy carried by the electrons is recovered as the output and the remainder is converted to heat which then

is rejected at the collector temperature to the thermodynamic heat sink. The relation for the electron cooling losses is defined to be

$$Q_{ec} = J \phi^* \pi \frac{D_E^2}{4} - P_{TD} \quad (\text{B-4})$$

Values of J and ϕ^* were obtained from Reference 8, (page 46, Figure 15 and page 27, Figure 7, respectively) corresponding to the emitter temperature and cesium reservoir conditions stated in the previous paragraph. An estimate of ϕ^* was made utilizing the experimental current density in the Richardson equation at the emitter temperature for the ignited mode of the diode, to provide a check on the experimental value. The following version of the Richardson equation (See Reference 7) was used:

$$\phi^* = \sigma T_E \ln \left(\frac{A' T_E^2}{J} \right) \quad (\text{B-5})$$

A compilation of the experimental and computed values is given below:

T_E , °K	T_E/T_{cs}	J	ϕ^* (Ref. 8), ev	ϕ^* (Calc.), ev
1700-----	3.09	6	2.72	2.62
1790-----	3.11	13.8	2.75	2.64
1984-----	3.28	12.5	2.85	2.98

Since the correlation between the predicted and experimental values was good, experimental values were used in the present analysis. The above four items comprise the diode losses.

5. Output and DC-DC Converter Loss

The power generator output as mentioned in the body of the report was fixed at 62.5 watts and a DC-DC converter efficiency of 80 percent was assumed with a resultant 50 watt system output and attendant 12.5 watt DC-DC converter loss.

APPENDIX C

THERMAL SHIELD LOSS ANALYSIS

The thermal shield loss is directly dependent on the thermionic diode analysis, and cannot be performed until the thermionic diode losses and the power generator output are known. The method used in this analysis consists of initially determining a fuel block size and, hence, thermal shield loss first by using just the thermionic diode losses and the output. Next, this estimate is iterated to determine the fuel block size to provide the corresponding amount of thermal power. The iterated thermal shield losses are then determined. Further iterations were not performed, because it was found that the incremental difference in the thermal shield losses for the first iteration actually could have been taken care of by changing the proportional amount of isotope fuel in the fuel block. The steps to perform the thermal shield loss calculations are detailed below.

Step 1: Initial Fuel Block Size

First, the volume of the fuel block is obtained from the following equation—

$$V_F = \frac{(\text{Diode losses} + P_{TD})XC}{P'_{d_c}} \quad (\text{C-1})$$

where

- (a) $C = 1.577$ for α emitting fuels to account for void volume required for helium buildup and is equal to 1, obviously, for β emitters.
- (b) $P'_{d_c} = P_{d_c} X (.577)$.

The fuel block diameter is then defined as—

$$D_F = 3 \sqrt{\frac{4V_F}{\left(\frac{L_F}{D_F}\right)\pi}} \quad (\text{C-2})$$

The ratio of (L_F/D_F) was 0.8 for all cases investigated with the exception of the curium 244 cases where it was 0.64 and 0.69 for the 1 and 3 year mission time, respectively. The fuel block length can be solved for directly from these ratios.

Step 2: Initial Estimate of Longitudinal Temperature Gradient

The one dimensional axial temperature difference for uniform heat generation is given by

$$\Delta T_L = \frac{(\text{Diode losses} + P_{TD})XL_F}{2k_F A_{\text{eff}}} \quad (\text{C-3})$$

where

- (a) k_F (shown versus temperature in Figure 7) obtained from data in Reference 9 evaluated at T_E .
- (b) A_{eff} was assumed to be 0.328 of the cross sectional area of the fuel block.

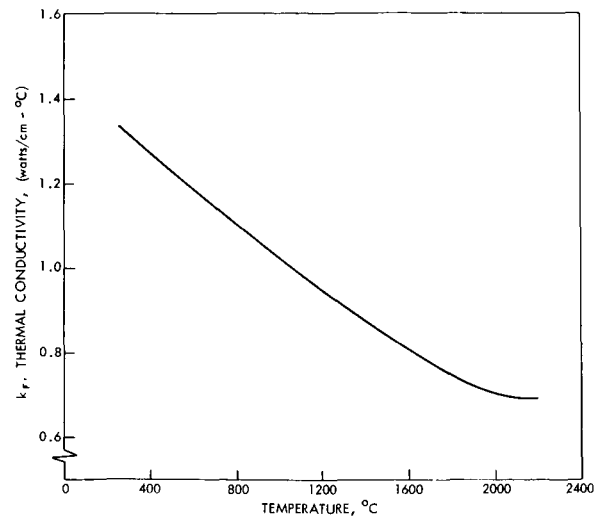


FIGURE 7.—Thermal conductivity of molybdenum versus temperature.

Step 3: Initial Estimate of Radial Temperature Gradient

The one dimensional radial temperature gradient for uniform heat production in a solid cylinder is given by—

$$\Delta T_R = \frac{P_{d_c} D_F^2}{16k_F} \quad (\text{C-4})$$

Step 4: Initial Fuel Block Surface Temperatures

It was assumed that the side shield radiates at a temperature of

$$T_S = T_E + \Delta T_L - \Delta T_R \quad (C-5)$$

Since the heat path is from top to bottom and

radially inward to the emitter, the bottom shield was assumed to be at a temperature given by

$$T_{bs} = T_E + \Delta T_R \quad (C-6)$$

and the edge of the bottom shield also to be at this temperature.

Step 5: Thermal Shield Loss Calculation

$$Q_{is\ initial} = K_1 \Big|_{T_s} (\pi D_F L_F) + K_1 \Big|_{T_{bs}} \pi (D_F^2 - D_E^2) + K_2 \Big|_{T_{bs}} \pi D_F \quad (C-7)$$

where K_1 and K_2 are experimental heat loss constants (see Figures 8 and 9, respectively) evaluated at the conditions indicated.

Step 6: Iteration of Thermal Shield Loss Calculation

An iteration of the above result to obtain a slightly more accurate and representative evaluation of this loss is accomplished as follows:

The initial thermal shield loss determined by equation (C-7) in Step 5 above is added to the

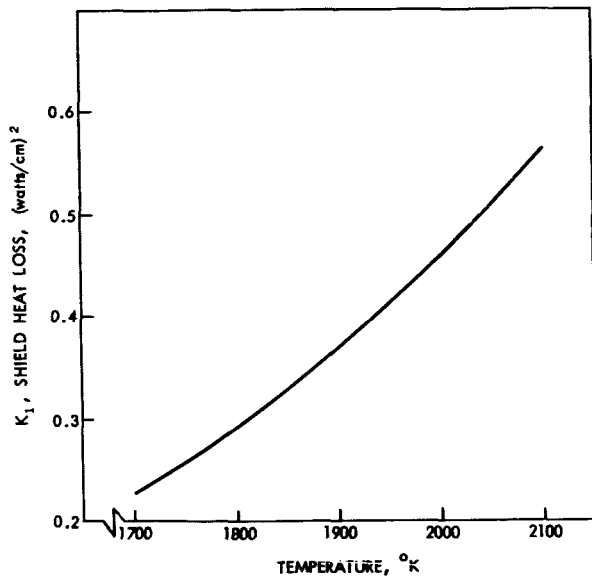


FIGURE 8.—Thermal shield loss constant versus temperature.

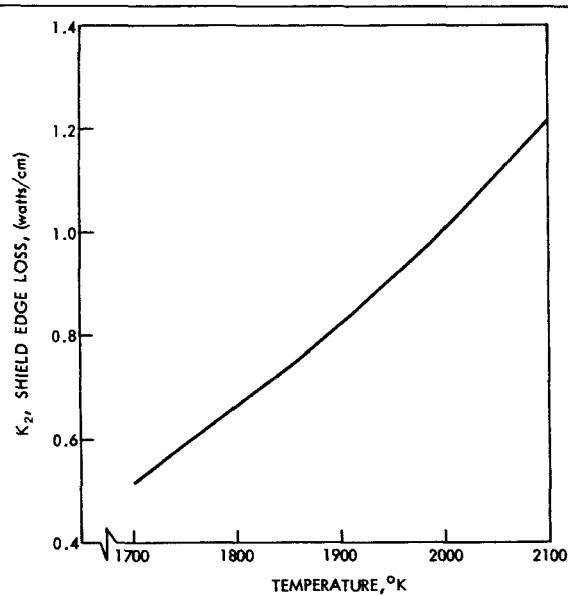


FIGURE 9.—Thermal shield edge loss constant versus temperature.

diode losses and power generator output to determine a new volume as per Step 1. The fuel block size is redetermined and the analysis Steps 2 to 5 repeated to reflect the effect of the increased size. The only difference between the initial and first iteration is that k_F is now evaluated at the calculated value of T_s and T_{bs} for the equation (C-3) and (C-4) computations, respectively, of this iteration.

APPENDIX D

SYSTEM WEIGHTS

The diode weight used, including support and leads, was based on values obtainable on a 50 watt diode preliminary design. The weight was 5.4 pounds of which 2.3 pounds was for internal diode structure, 2.4 pounds for collector radiator casing, 0.2 pounds for leads and 0.5 pounds for structural support.

The DC-DC converter weight was estimated for the 50 watt system to be 2.4 pounds based on a state-of-art converted electrical power to converter weight ratio of 20.5 watts/pound at a 500 electrical watt power level.

The fuel block weight consists of the fuel plus the cylindrical molybdenum block into which it is inserted. The fuel block weight is defined by the following equation:

$$W_F = 2.2 \times 10^{-3} \frac{V_F}{C} [0.577\rho_I + 0.423\rho_F] \quad (\text{D-1})$$

The isotope densities used in the above equation are given in Table II, and the value used for molybdenum was 10.2 gm/cm³.

APPENDIX E

ESTIMATE OF SHIELDING WEIGHTS

The shielding weight calculation consists of first determining the dose rate associated with the particular isotope, and then the determination of the shield thickness to attenuate this dose rate to some prescribed dose rate level at a specified fixed distance. An example of a dose rate determination is given below and the equations used to determine the shield thicknesses for the isotopes requiring shielding and consequent weight are also given.

(1) Dose Rate Determination

The isotope whose dose rate is to be calculated is Curium 244. Although there is slight gamma radiation resultant from the radioisotope decay process, the dose rate from Curium 244 is primarily associated with the neutrons which are yielded in the spontaneous fissioning of the isotope. Thus, the calculations below will be for the fission neutron contribution to the dose rate only.

Step A—Neutron Flux Rate

First, the number of neutrons per second per gm must be determined as follows:

The number of fissions per second per gm is defined by the following equation—

$$\frac{\text{Fissions}}{\text{sec-gm}} = \frac{0.693N_0}{AT_{1/2}} \quad (\text{E-1})$$

where $T_{1/2}$ (half life of fission neutron) equals 1.4×10^7 years (Reference 10). The value of (E-1) is then 3.87×10^6 fissions per second per gm. The number of fission neutrons yielded in the spontaneous fission process is given in Reference 11 as 2.6 ± 0.15 neutrons per fission, and for the present calculations, a value of 2.75 was assumed for conservatism. When this value is combined with that obtained from equation (E-1) the following result is obtained—

$$\frac{\text{Neutrons}}{\text{sec-gm}} = 1.06 \times 10^{17} \quad (\text{E-2})$$

In order to perform the shield calculation and set a reference value for the correspondence between the required thermal watts of isotope and the dose rate, 100 thermal watts was chosen as this reference. Thus, to convert the neutron flux in (E-2) to this reference, the following relation was used:

$$\frac{\text{Neutrons}}{\text{sec-100 watt}_{th}} = (\text{E-2}) \times \frac{100}{P_{dt}|_{t=0}} \quad (\text{E-3})$$

When the indicated operation is performed, the result is 0.373×10^9 neutrons per second per 100 watts-thermal. The final operation in determining the neutron flux rate is to consider the geometry and fix the distance at which you want to determine this rate. This equation for a point source radiating radially is—

$$\phi = \frac{(\text{E-3})}{4\pi x^2} \quad (\text{E-4})$$

For this example at one meter from the source, the value is 0.297×10^4 neutrons per second per 100 watts-thermal per square centimeter.

Step B—Dose Rate

First it was assumed that the fission neutrons had an average energy of 2 Mev, because this energy level is representative of the probable energy of a fast neutron. Using this energy value, the correlation between neutron flux rate and dose rate was found in Reference 12 (p. 22) and this relation is as follows:

$$\frac{1 \text{ neutron}}{\text{sec-cm}^2} = \frac{0.12 \text{ mrem}}{\text{hr}} \quad (\text{E-5})$$

The dose rate for the 100 watt-thermal source at one meter is then determined by multiplying the result of equation (E-4) by the (E-5) relation to yield.

$$D_R = 6.356 \frac{\text{rem}}{\text{hr}} \text{ at 1 meter} \quad (\text{E-6})$$

(2) Determination of Lead Shield Thicknesses

The basic equation for the shield thickness to change the dose rate level D_R to some required value is given as follows—

$$\ell = -\frac{1}{\mu} \ln \left(\frac{D_{R_{req}}}{BD_R} \right) \quad (E-7)$$

In the present investigation, this level was fixed at 10 mrem at one meter. Since the dose rate calculation is somewhat long and involved, lead shield thickness values were obtained from the literature for the reference dose rate of 10 mrem per hr at 1 meter and these are as follows—

Isotope	Thermal Source Watts	ℓ , cm	Reference
Sr ⁹⁰ -----	10	15.24	3
Ce ¹⁴⁴ -----	10	19.05	3
Tm ¹⁷⁰ -----	0.5	5.4	13
Cm ²⁴⁴ -----	100	6.5	Above calculation

The above values were then used in the following equation to calculate the shield thickness in a lateral direction as this equation is

$$\ell_s = \ell - \ln \left(\frac{\text{Reference thermal source}}{\text{Required thermal source}} \right) \quad (E-8)$$

Shielding was assumed to be only required in the lateral direction, since the diode structure and perhaps the spacecraft itself could provide attenuation in the axial direction. Thus, the shield weight was then calculated by the following equation—

$$W_s = \rho_{Pb} \frac{\pi}{4} (\ell_s^2 - D_F^2) L_F \quad (E-9)$$

The above approach to the shielding was taken, because if the source were covered spherically to shield the radiation in all directions, the shield weight would be even more horrendous than determined from the assumption above.

FACILITY DESCRIPTION. GODDARD ANTENNA TEST RANGE*

PAUL A. LANTZ

INTRODUCTION

The Goddard Antenna Test Range is located in Prince George's County, Maryland, on Beaverdam Road, 0.8 miles east of the intersection of this road with Soil Conservation Service Road. The range is one mile northeast of the NASA Goddard Space Flight Center (see Figure 1). The property was made available to GSFC through a use agreement with the U.S. Department of Agriculture Beltsville Research Center.

The range is available for use by all antenna research personnel at Goddard Space Flight Center. The purpose of this Facility Description is to describe the range, to discuss its capabilities, and to establish its limitations for prospective users.

DESCRIPTION OF RANGE

General Description

The range contains two 75-foot, vertical steel towers (one for transmitting and one for receiving) located 1000 feet apart on a 49-acre plot of ground graded to ± 6 inches from a plane surface. In addition, the range contains a unique vertical RF anechoic chamber. This chamber is the terminal end of a rectangular chamber, erected vertically and covered with a radome. A field-generating (or receiving) antenna is suspended above the radome. Finally, a renovated farmhouse on the range serves as a laboratory, an office, and a small machine shop. Figure 2, an illustration of the over-all range, shows the location of a proposed additional receiving tower, in line with the present towers and 300 feet west of the transmitting tower. Figure 3 is a scale drawing "site plan" of the existing range.

*Published as *Goddard Space Flight Center Document X-525-64-223*, October 1964.

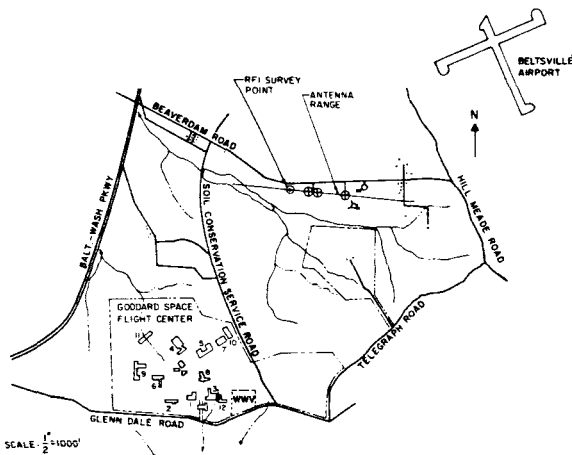


FIGURE 1.—Map showing location of antenna test range.

The Horizontal Range

The horizontal range presently consists of two 75-foot towers spaced 1,000 feet apart. Each tower contains a 20 by 20-foot room at the top for housing instrumentation; the rooms are heated and air-conditioned. The room on the westernmost transmitting tower (Figure 4) is fitted with fiberglass facings on the east and west sides. Field-generating antennas located in this room can thereby illuminate the long-range receiving tower located 1,000 feet to the east and a proposed short-range receiving tower to be located 300 feet to the west. A vertical track is attached to the east face of the transmitting tower; the track can be seen in the photograph of this tower (Figure 4). On a remotely controlled trolley on this track are mounted a 6-foot paraboloidal reflector, a log periodic dipole structure, and an 8-foot paraboloidal reflector. The parabolas are located 8 feet from each side of the centrally located log periodic structure. A power distributing transformer at the base of the tower

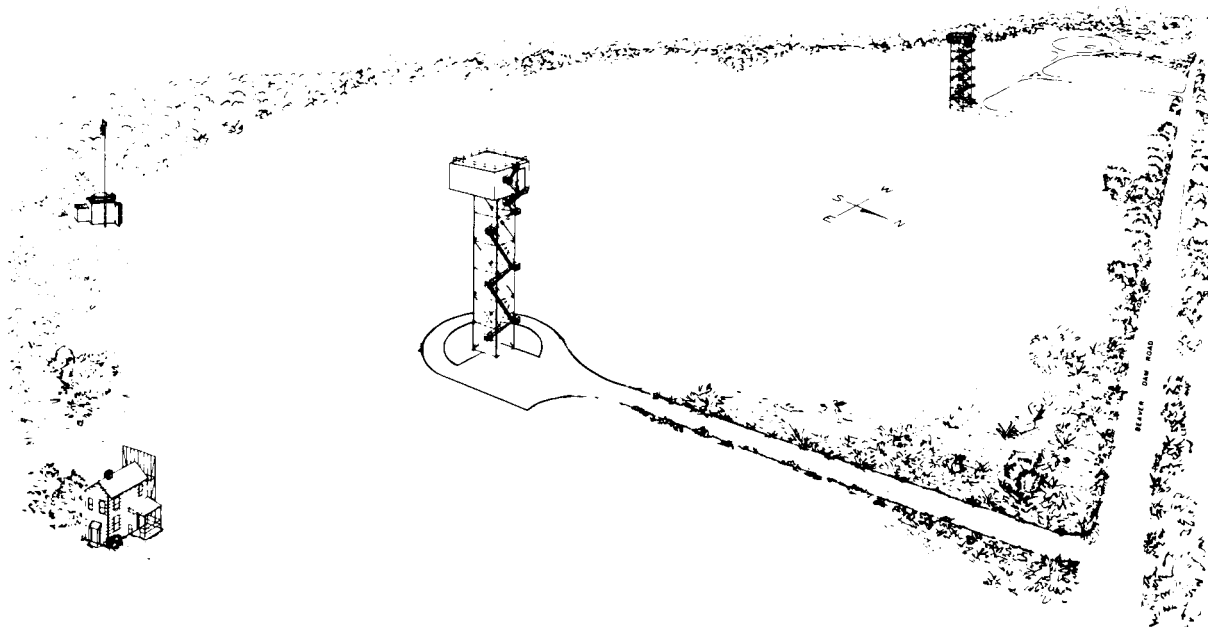


FIGURE 2.—Goddard antenna test range.

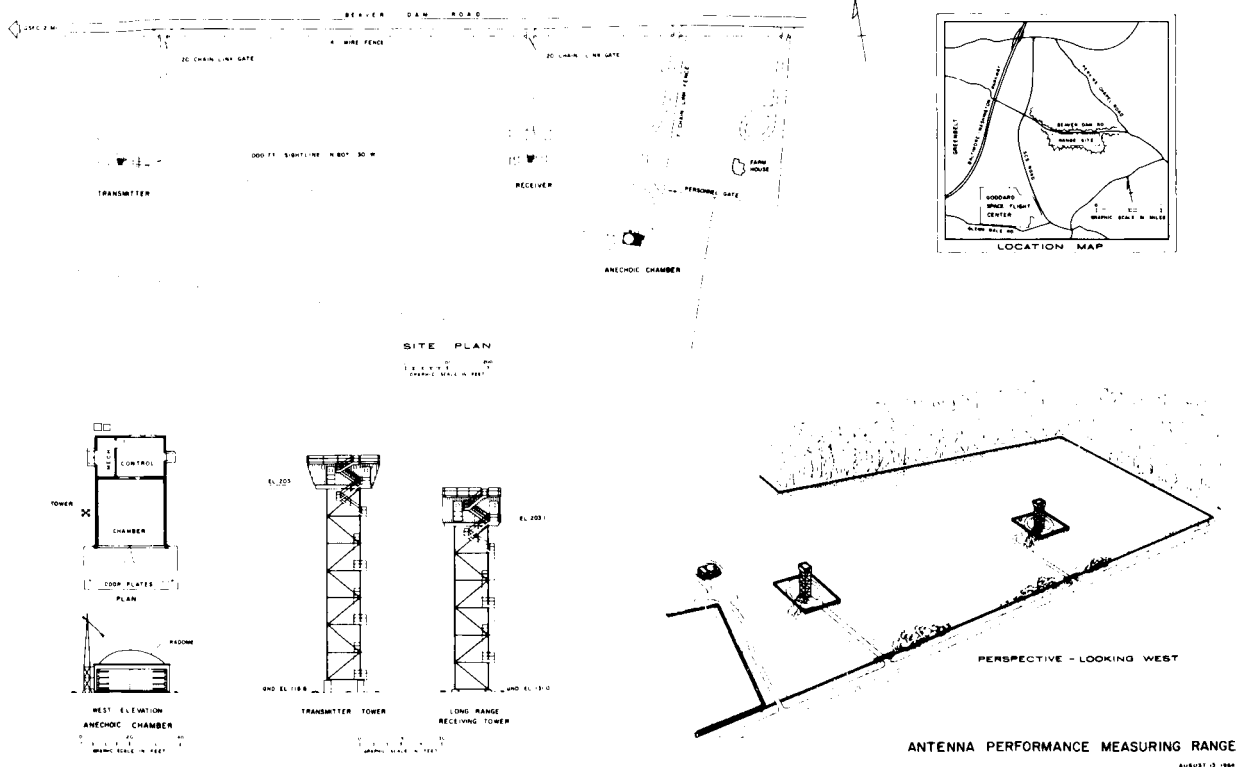


FIGURE 3.—Site plan, Goddard antenna test range.

fixes the lower limit of travel at 11 feet from the ground; the maximum upper point of travel is 63 feet (52 feet total available travel).

Each tower room is constructed with a flat roof and 8-inch steel I beams (Figure 5) for mounting heavy antennas. A one-ton jib crane



FIGURE 4.—Transmitting tower.

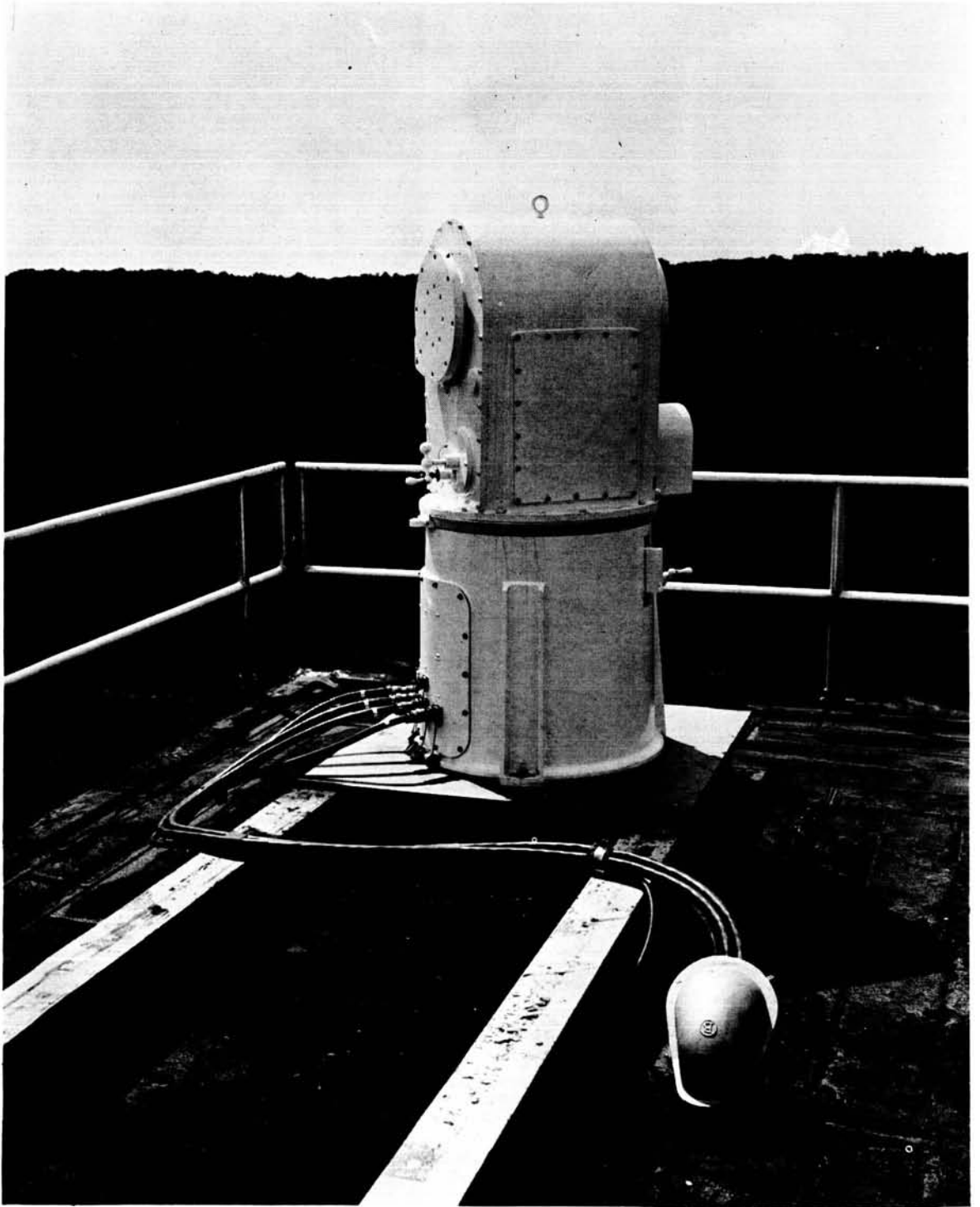


FIGURE 5.—Roof of transmitting tower.

(Figure 6) is mounted on the roof of each tower room. These cranes cause some spurious RF reflection and it may become necessary to remove

them. Nine 4-inch diameter underground cable ducts (Figure 7) connect the towers and carry power lines, telephone lines, and remote control

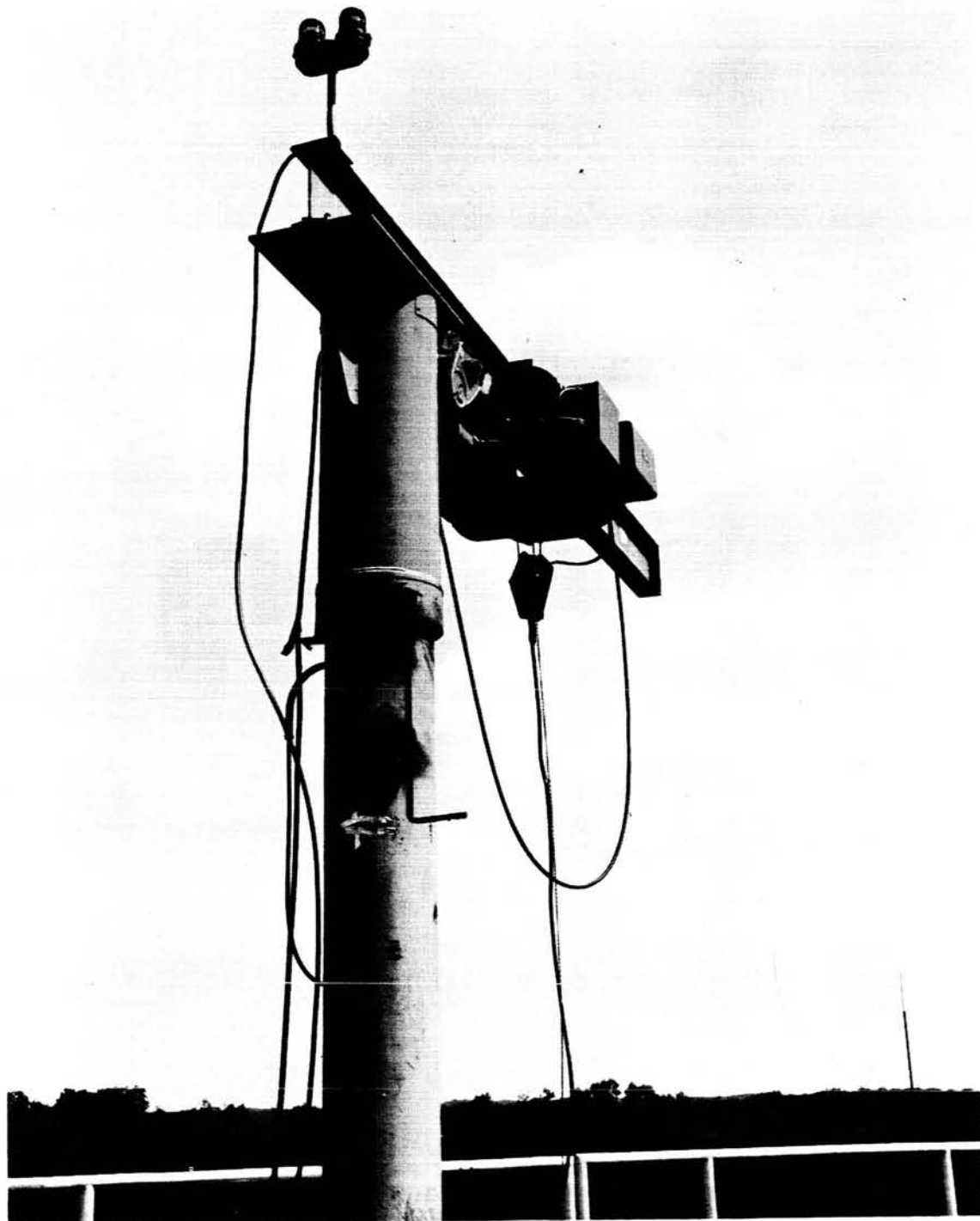


FIGURE 6.—Jib crane.



FIGURE 7.—Inter-tower conduit terminus.

cables. A continuous length of 1 $\frac{5}{8}$ -inch Styroflex, extending from one tower to the other through one of the ducts, can be used for phase pattern measurement.

The architectural and engineering design of the towers was done by Keller, Loewer, Sargent and Associates, Silver Spring, Maryland, under Goddard Space Flight Center contract NAS5-832. Construction drawings resulting from this design are numbered GF-176 through 189 and are dated April 1962.

The Vertical Range

Figure 8 is an exterior view of the vertical range.* The structure contains a 26 by 26-foot poured concrete, open-top chamber. This chamber is 16 feet deep, 6 feet being below grade. Two 8 by 13-foot doors, which form the west wall of the chamber, open out onto a concrete apron

*Designed by Keller, Loewer, Sargent and Associates, Silver Spring, Maryland, (contract NAS5-832) and constructed by Jack Bays, Inc., Alexandria, Virginia, (contract NAS5-4760). The architectural and engineering design drawings describing this facility are numbered GD-2000 through GD-2008 and are dated October 1962.



FIGURE 8.—Vertical range (exterior view).

and expose the interior of the chamber to the horizontal range transmitting tower for operation in the ground mode. Normal operation, however, is to an 8-foot circularly polarized log periodic structure suspended above a radome which covers the open top of the chamber. The distance from a phase center of this log periodic structure to the center of a quiet zone inside the chamber is 32 feet.

The four walls and the floor of the chamber are covered with B. F. Goodrich type VHP-70 RF absorbent material (contract NAS5-3573) as shown in Figure 9, which is a view looking toward the 8 by 13-foot doors. Figure 10 is the same

view prior to installation of the absorbent material. This absorbent material is 70 inches thick and reduces the usable interior to an area 14 by 14 feet in dimension. A 10-foot diameter spherical RF "quiet" zone is produced. An azimuth-over-elevation pedestal, nested in the RF absorbent material on the north wall (Figure 11), permits positioning of antennas in this quiet zone. The pyramidal absorber cones on the floor are covered with high-density vinyl foam planks supported on inverted pyramids of styrene foam. This arrangement can be seen in Figure 12, a view of the floor looking toward a 3.5 by 6.5-foot "pull out" door supported on tracks so

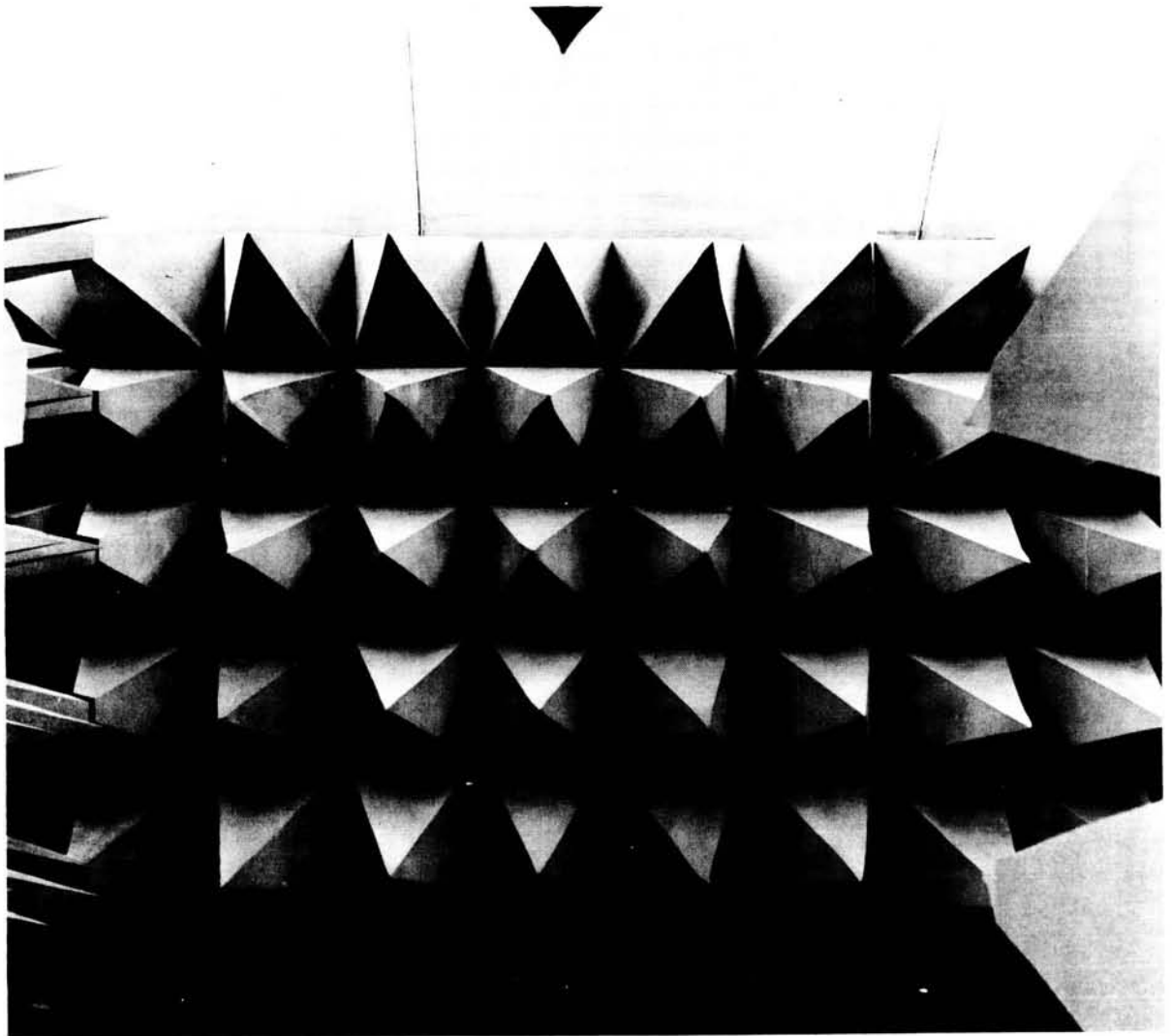


FIGURE 9.—Interior of anechoic chamber looking toward door after installation of absorbent material.

that it can be drawn out of the chamber wall into an appurtenant 18 by 15-foot control room.

Figure 13 illustrates the manner in which heating and air conditioning is admitted to the chamber through 2-inch holes in the absorber. Lighting is accomplished by floodlighting the interior surface of the white painted radome. In addition, a viewing port to the control room can be seen in Figure 13. The "pull out" door can be seen in Figure 14, which is an interior view of the appurtenant control room. Also in Figure 14 can be seen the doorway to the mechanical equipment room. This mechanical equipment

room, 8 by 15 feet, houses a heat pump and electric power distribution panels.

Figure 15 is an interior view of the sandwich-type radome* which serves as a roof for the chamber. The radome is composed of an inner core of dielectric foam 4 inches thick and is covered with a fiberglass outer shell 0.030 inch thick.

The House

The house located on the range is a remodeled farmhouse which provides two rooms on the first

*Supplied by Raymond Development Industries, Inc., Huntington Park, California (contract NAS5-4285).

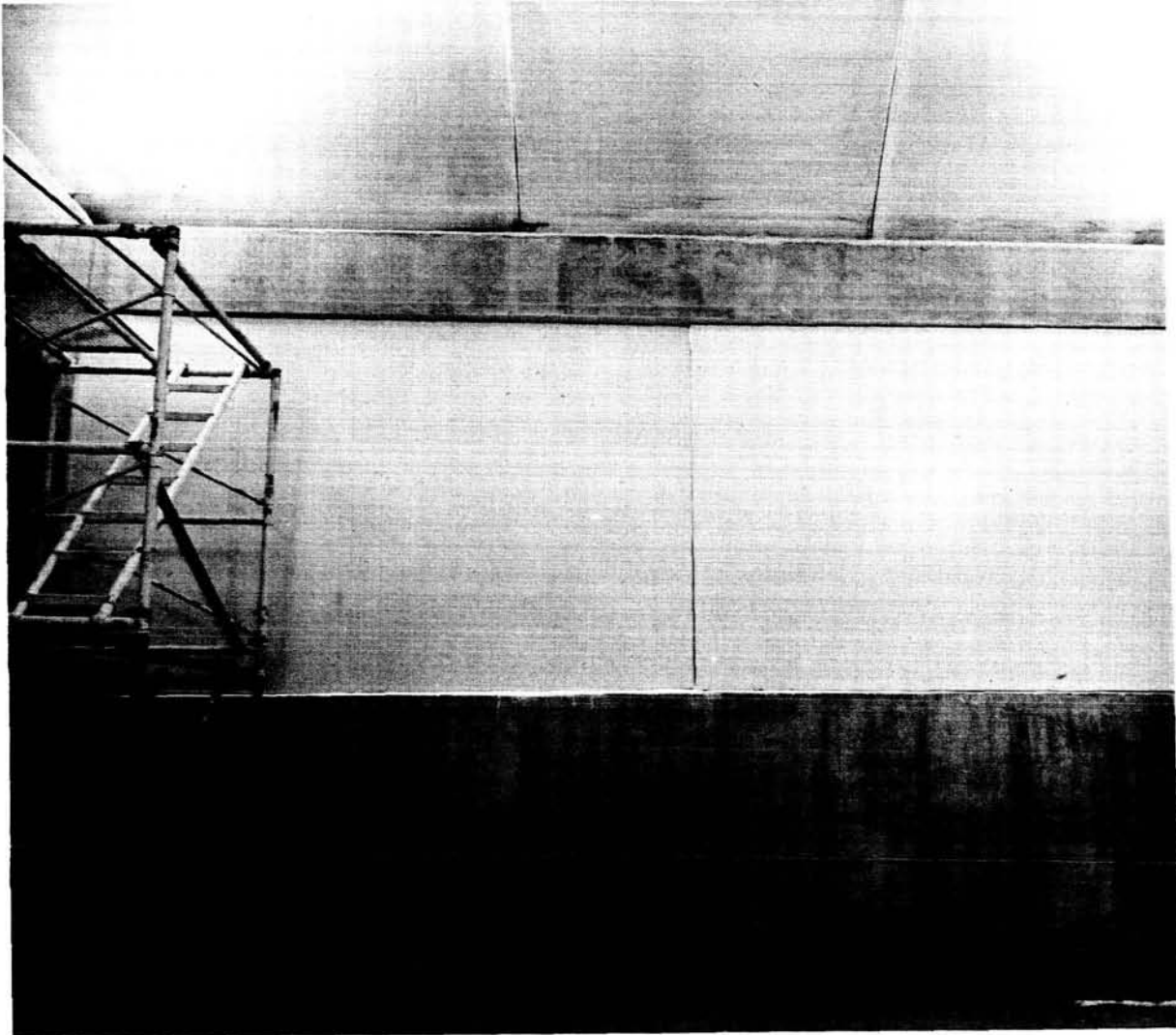


FIGURE 10.—Interior of anechoic chamber looking toward door before installation of RF absorbent material.

floor, a 27 by 15 foot office room and a 17 by 13 foot shop and storage space, as well as sanitation and bath facilities. On the second floor are two 13 by 15 foot rooms, one of which is used as an office and the other of which is used as a laboratory. The house has electrical heating and heavy-duty electrical service to power window air conditioners, power tools, and electronic test equipment. It is planned to install a closed-circuit television system with monitors located in the house to view the towers, tower rooms, and vertical anechoic chamber interior. This system

will serve as a safety device by indicating personnel in distress atop the towers (the 75-foot climb is a strenuous one), and will enable the range manager to observe tests underway at various locations at one time.

RANGE INSTRUMENTATION

The 1,000-foot horizontal test range is presently instrumented for measurements in the frequency range 50 Mc/s to 12.4 Gc/s. Instrumentation for the receiving and transmitting towers is listed below.

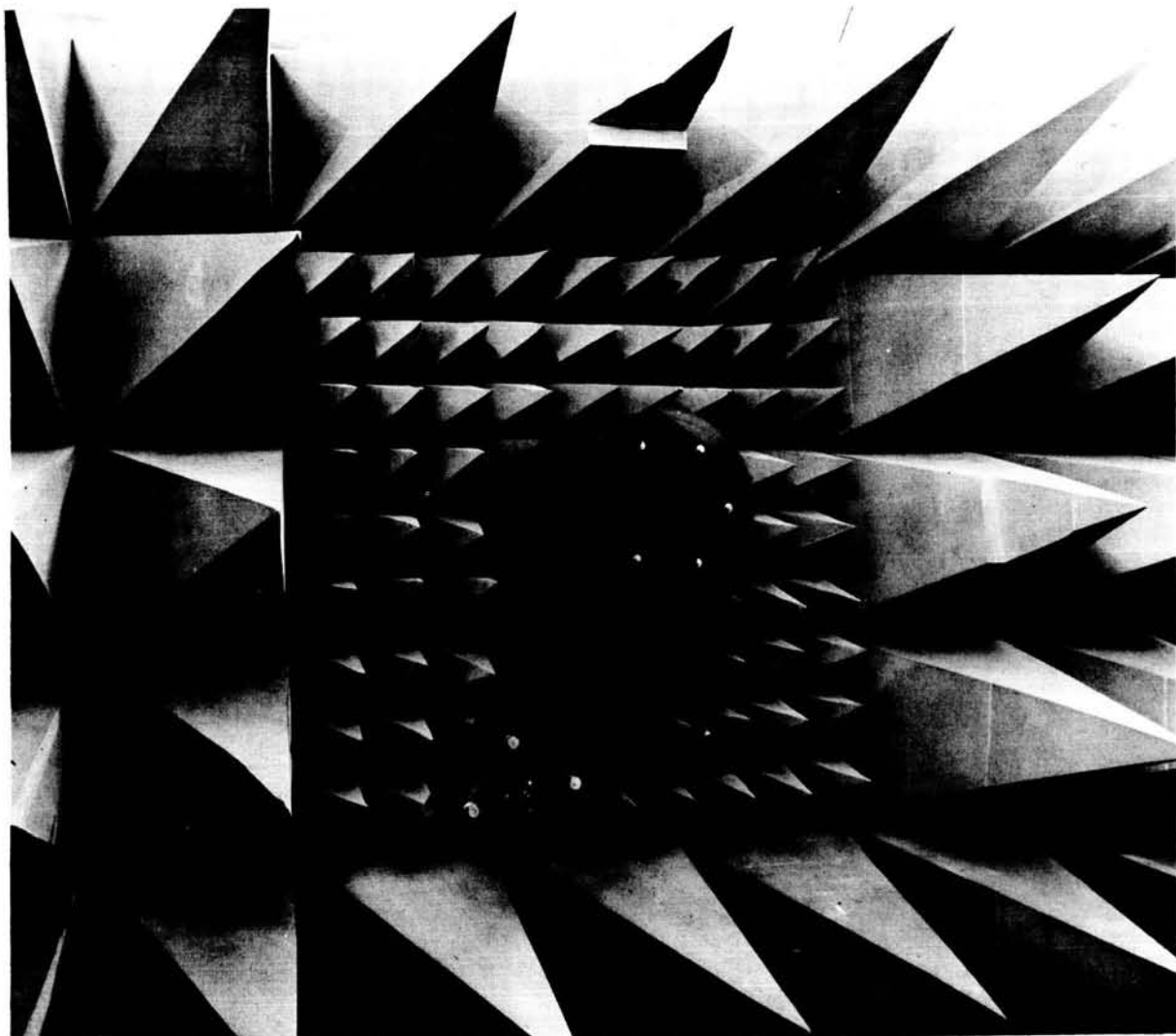


FIGURE 11.—Antenna mount inside vertical range anechoic chamber.

The Receiving Tower

Wide-range receiving system

Scientific-Atlanta Model 1640 APZ/crystal
mixer 14-2

Antenna pattern recorder (rectangular)

Scientific-Atlanta Model 1525

Antenna pattern integrator and spherical con-
verter unit

Scientific-Atlanta Model 2611 and Model 2621

Test antenna positioner programmer

Scientific-Atlanta Model 2004

Radiation distribution printer

Scientific-Atlanta Model 1801

Positioner control

Scientific-Atlanta Model 4112

Synchro isolation amplifiers

Scientific-Atlanta Model 4541

Position indicator (dual)

Scientific-Atlanta Model 4422-66

Positioner control

Scientific-Atlanta Model 4103

Differential synchro transmitter

Scientific-Atlanta Model 4572



FIGURE 12.—Floor of anechoic chamber looking toward pull-out door.

Remote tuning units

Scientific-Atlanta Model RT2D

Transmitting tower positioner control

Scientific-Atlanta Model PC4A

Transmitting tower position indicators

Scientific-Atlanta Model PI-3-444

Switch panel for PI-3-444

Scientific-Atlanta Model PAE-S-R

Azimuth positioner

Scientific-Atlanta Model PA-15/81791

Servo control unit

Scientific-Atlanta Model 3610

Servo amplifier

Scientific-Atlanta Model 3620

Position indicator unit

Scientific-Atlanta Model 3710

Calibrated gain standard, 100-1000 Mc, linear and circular

Litton Industries 12E1-10

Calibrated gain standard, 1000-5000 Mc, linear and circular

Litton Industries 13E1-5

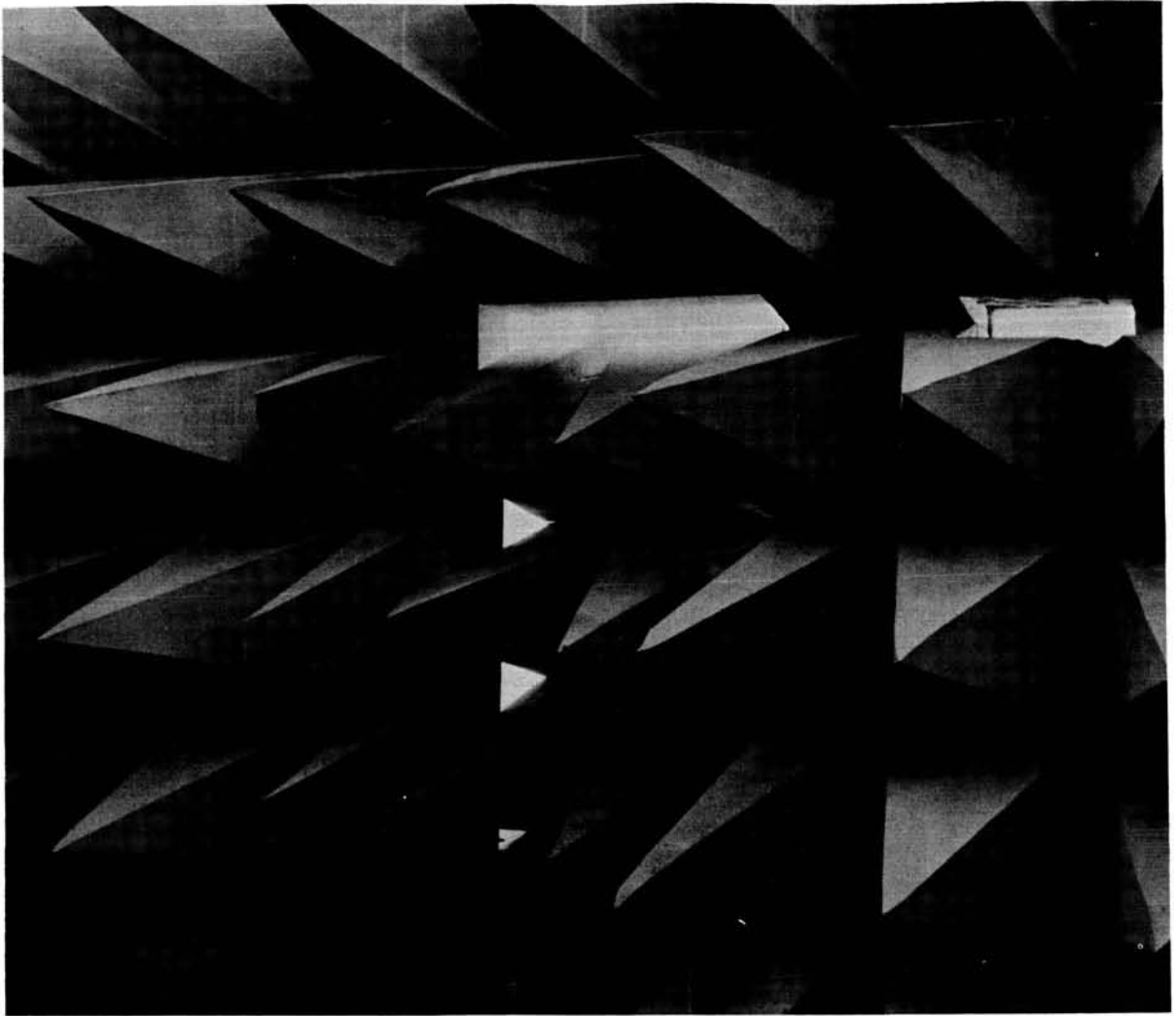


FIGURE 13.—East wall of vertical range anechoic chamber.

The Transmitting Tower

Remote tuned signal sources	Transmitting antenna (8-foot parabolic reflector)
Scientific-Atlanta Model SS21 (3 each)	Scientific-Atlanta Model 22-8
Remote tuned signal source	Transmitting antenna feed for 8-foot parabola
Scientific-Atlanta Model SS31, 50 to 2000 Mc	Scientific—Atlanta Model 27-1.0/8, 1 to 6 Gc
RF oscillators	Transmitting antenna (6-foot parabolic reflector)
Scientific-Atlanta Model RFO 2, 2 to 4 Gc	Scientific-Atlanta Model 22-6
Scientific-Atlanta Model RFO 4, 4 to 8 Gc	Transmitting antenna feed for 6-foot parabola
Scientific-Atlanta Model RFO 8, 8 to 12.4 Gc	Scientific-Atlanta Model 27-4.0/6, 4 to 12.4 Gc
Polarization positioner	Field-generating antenna hoist on track
Scientific-Atlanta Model 5601-S1	Scientific-Atlanta Model HQ-949
Transmitting antenna (log periodic structure)	Azimuth over elevation positioner
Scientific-Atlanta Model 26-1.0/81791, 50 to 2000 Mc	Scientific-Atlanta Model PAE-5 (2 each)

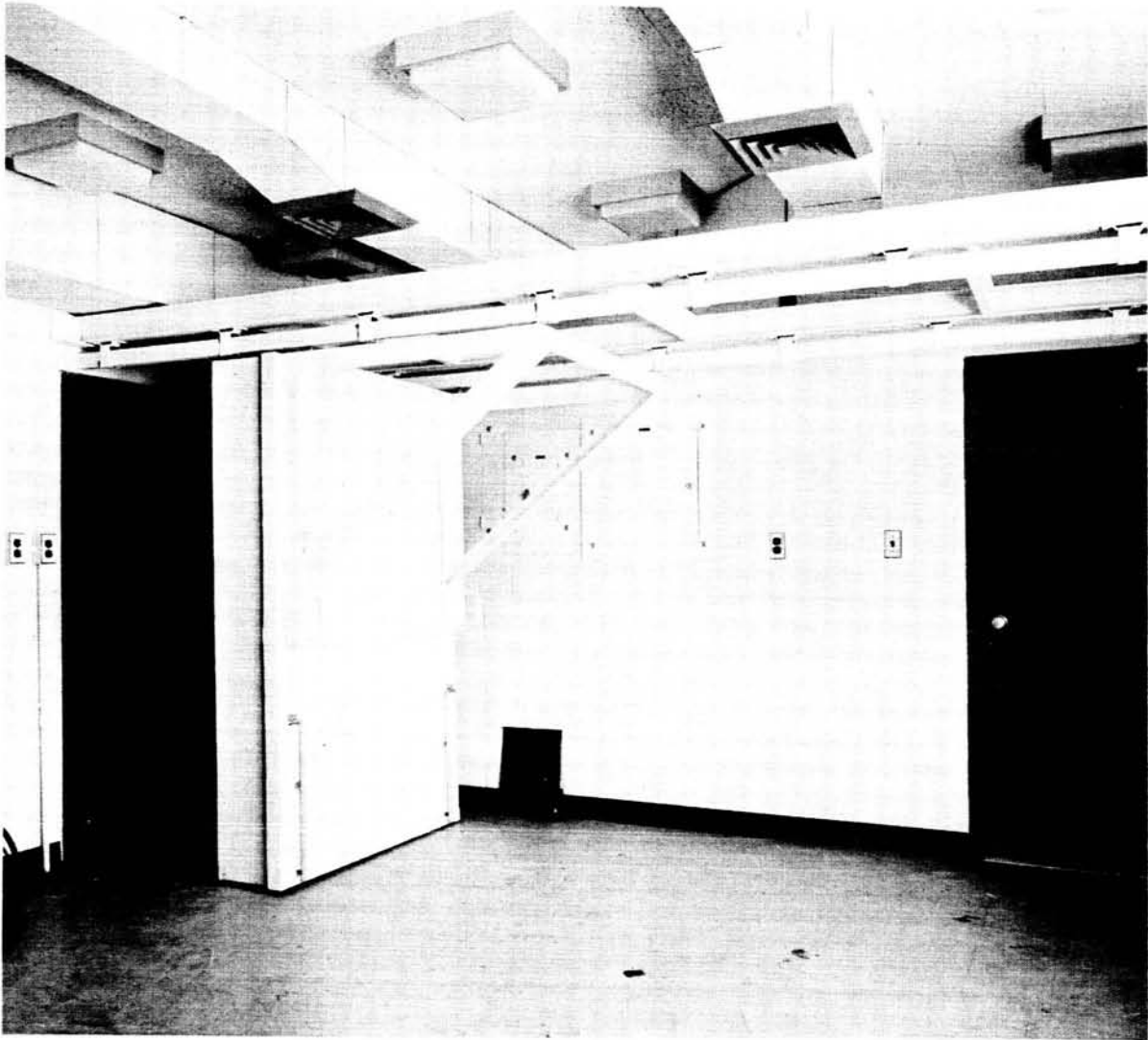


FIGURE 14.—Control room, vertical range anechoic chamber.

HORIZONTAL RANGE CAPABILITY

The Horizontal Range consists of a transmitting tower and a receiving tower separated by 1,000 feet. The ground area separating the two towers has been graded level to within ± 6 inches to minimize and control reflections. Figure 16 is a photograph of the receiving tower; Figure 17 is the transmitting tower as seen from the top of the receiving tower. Figures 18 and 19 are interior views of the receiving tower control room. Figure 20 shows the transmitting tower instruments. Figure 21 shows the three transmitting

antennas on the transmitting tower vertical positioning track.

This facility was designed as a dual mode range covering the frequency spectrum of from 50 Mc/s to 12.4 Gc/s. That is to say, it operates in the ground plane mode from 50 Mc/s to 1000 Mc/s and in the free space mode from 1000 Mc/s to 12.4 Gc/s.

To design the range, it was necessary to compute performance characteristics. As a part of this analysis, the incident field on an antenna under test was evaluated in terms of variations

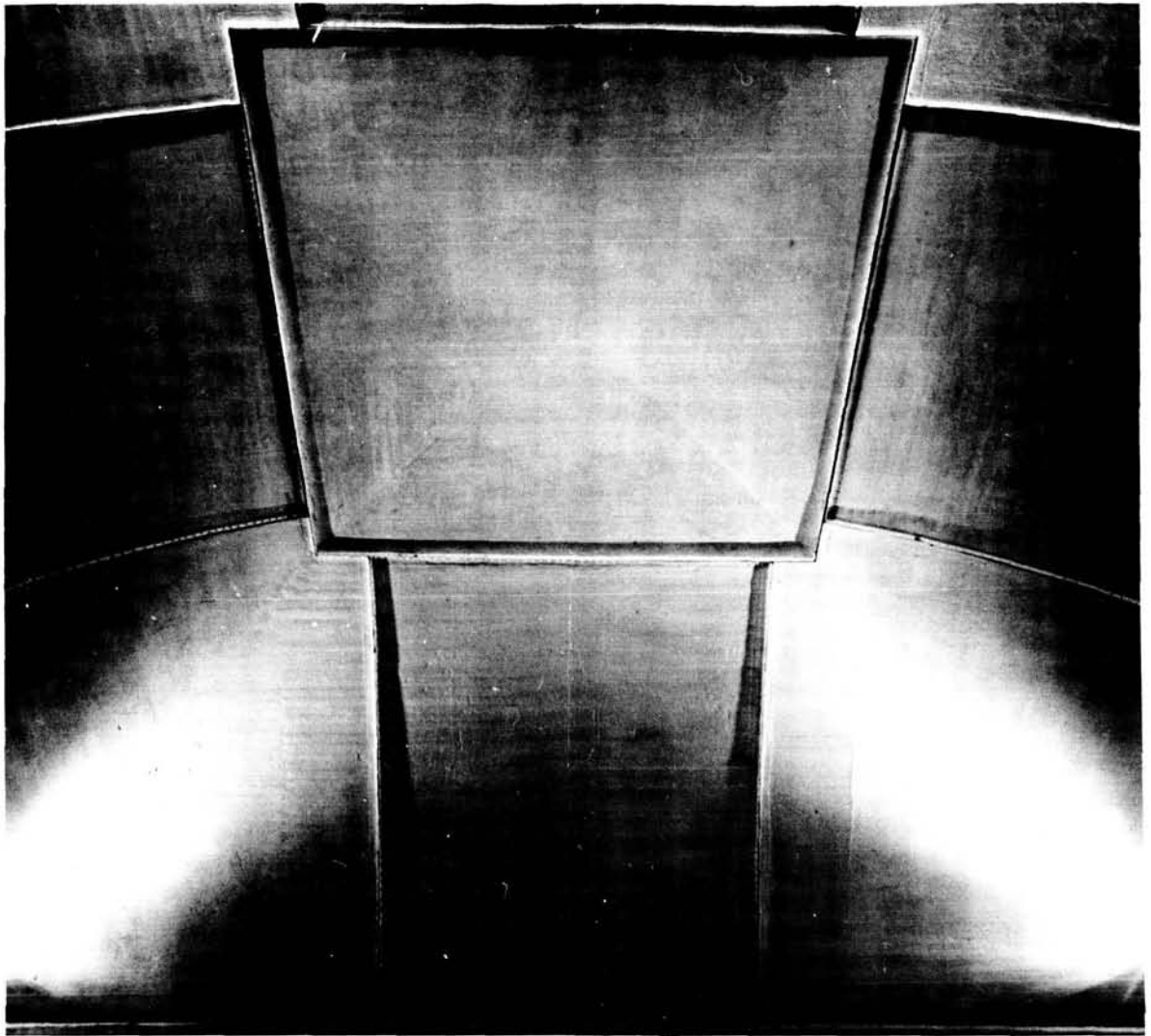


FIGURE 15.—Interior view of vertical anechoic chamber radome.

in phase and amplitude. To obtain accurate far-field, or Fraunhofer region, patterns it is necessary that the distance between the transmitting antenna and the antenna under test be sufficiently large. If this distance is too small, the antenna under test is illuminated with a spherical wavefront, and a near-field, or Fresnel region, pattern is obtained. The Fresnel pattern is a function of the distance at which it is measured. For accurate far-field measurements the antenna under test should be illuminated with a plane wavefront. Since plane wavefronts are

obtainable only at infinite distances, some limits must be specified. A commonly specified criterion (the Raleigh¹ criterion) is that the phase difference between the center and edge of the antenna under test shall be² no greater than $\lambda/16$. If this is the case, the distance defining the range, R , is

$$R \geq \frac{2d^2}{\lambda}$$

where d is the aperture dimension of the antenna under test. Figure 22 is a plot of this relationship.



FIGURE 16.—Horizontal range receiving tower.

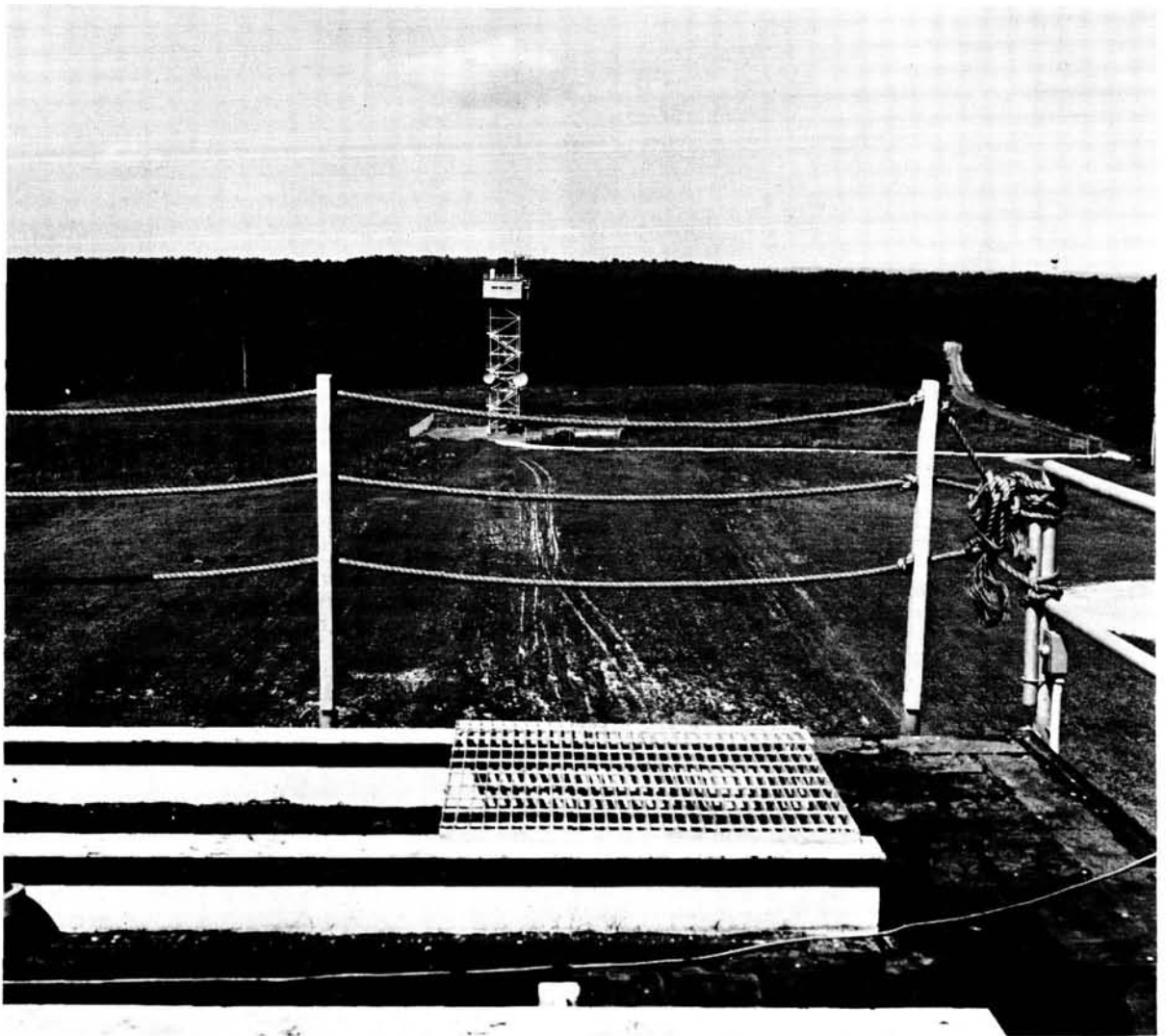


FIGURE 17.—Horizontal range transmitting tower viewed from receiving tower room.

If an attempt is made to measure the pattern of an antenna with an excessively large aperture (i.e., if $d = 10$ feet at 10 Mc on the 1000-foot range, Figure 22) minor distortions of the side lobe structure will result. Greater excesses of d will cause appreciable errors in the measured gain and side lobe structure.³ For a well-focused antenna at $2d^2/\lambda$, phase deviation of $\lambda/16$ will cause a gain error⁴ of about 0.01 db and negligible change in the pattern shape. At d^2/λ the gain error will be 0.06 db. At the highest frequency (12.4 Gc) now instrumented on the 1000-foot

range, Figure 22 shows the maximum allowable aperture, D , to be approximately 6 feet.

Excessive variations in the amplitude of the field over the aperture of an antenna under test will also cause significant errors in the measured gain and sidelobe level. Montgomery⁵ shows that the power received (across the aperture of an antenna under test) should not differ from that which would be received if there were no amplitude error by more than 5 percent or 0.25 db. In addition, Montgomery shows the relationship of the field-generating antenna aperture, D_t , to the

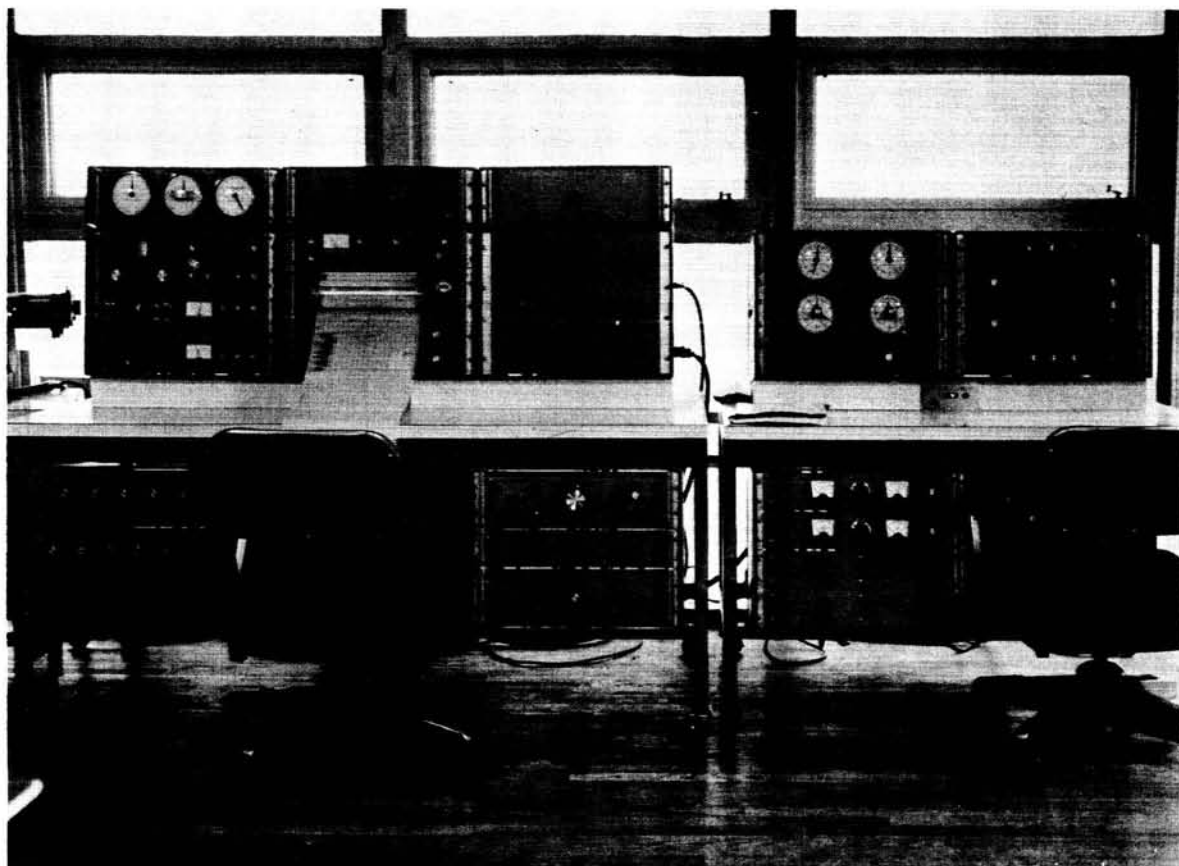


FIGURE 18.—Horizontal range receiving tower instruments.

antenna under test aperture, D_r , for this condition to be

$$0.25 \text{ db or } 0.95 \approx 1 - 0.05 \left(\frac{1}{\lambda R} \right)^2 \\ (D_t^4 + 6D_t^2 D_r + D_r^4).$$

Figure 23 has been constructed from this relationship and shows the maximum allowable antenna aperture which can be tested on the 1000-foot range for various field-generating aperture sizes.

The above discussion of phase and amplitude variations considers only the direct path of propagation. Variations of the field can also be caused by specular reflections from the ground. In the elevated or free-space mode of operation these effects are minimized by the vertical directivity pattern of the source antenna and the antenna

under test and, hopefully, by absorption and scattering rather than specular reflection from the range surface. The 1000-foot horizontal range uses the elevated or free-space mode at the higher frequencies between 1.0 and 12.4 Gc/s.

As opposed to the method of minimizing the effects of reflections by limiting their magnitude, the ground level technique is a method for a sufficiently uniform field over the aperture of an antenna under test. In this technique the transmitting antenna and the antenna under test are operated in close proximity to the ground and the antenna range surface is graded flat and smooth within Rayleigh's⁶ criterion. Under these conditions the magnitude of the reflection coefficient is nearly unity and the phase is nearly 180 degrees. The lower frequency limit of this technique depends upon the angle of incidence and the conductivity of the ground.

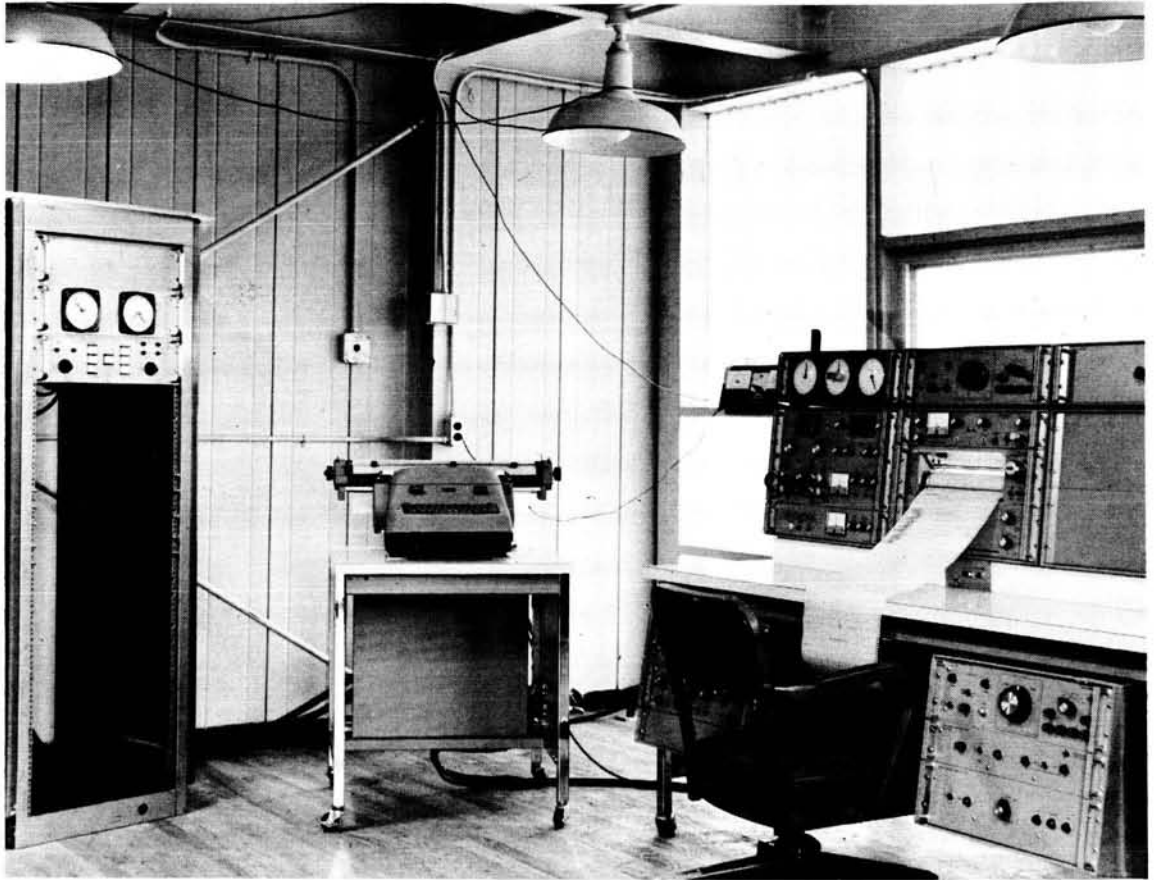


FIGURE 19.—Horizontal range receiving tower instruments.

In the 1000-foot horizontal range the distance R and the height h_2 of the center of the test aperture are fixed. The antennas located at the transmitting tower are mounted on a vertical carriage permitting height adjustment from 11 feet to 63 feet above ground level (52 feet total available travel). The height setting is a function of frequency. The proper height is given by

$$h_1 = \frac{\lambda R}{4h_2}$$

Adjustment of h_1 is accomplished from the receiving tower control room. The vertical tilt of the field-generating antennas is also adjustable remotely to provide signal "peaking" capability. This adjustment is especially important near ground level.

As previously mentioned, in the design of a ground level range the smoothness of the surface

is important in achieving specular reflection and Rayleigh's criterion must be satisfied. Under these conditions the surface irregularities must be less than $\sin \psi \lambda / 8$.

In the area between the towers on the 1000-foot horizontal range, the ground has been leveled to ± 6 inches.

VERTICAL RANGE CAPABILITY

The performance characteristics of the vertical range are discussed in considerable detail in a Goddard Technical Note⁷ now in preparation. The discussion of performance capability in this Facility Description is therefore limited to general observations.

Radome

The loss tangent and dielectric constant of the sandwich radome material were measured by the



FIGURE 20.—Horizontal range transmitting tower instruments.

manufacturer Raymond Development Industries. Using flat samples, one foot square, the characteristics listed in Table I were observed.

TABLE I.—Dielectric Constant and Loss Tangent Vertical Range Radome

Material	Dielectric Constant	Loss Tangent
Fiberglass skin.....	4.00	0.0050
Foam core.....	1.12	0.0005

The frequency at which these measurements were made was not indicated by the manufacturer.

The radome performance is reasonably good at low frequencies; however, it can be said

generally that it causes trapping and reflection when the wavelength becomes short enough to approximate the radome thickness (i.e., above 900 to 1000 Mc). The use of a sandwich radome for this application was probably not a good choice. It is planned to conduct a study, using scale models, to determine the most suitable radome.

RF Absorbent Material

The performance characteristics of the RF absorbent material were thoroughly measured by the supplier, B. F. Goodrich Company. Each block of material was checked by insertion in a waveguide large enough to propagate 120-Mc waves. In addition, the performance levels in the completed chamber were measured by moving tuned dipoles horizontally and vertically in the

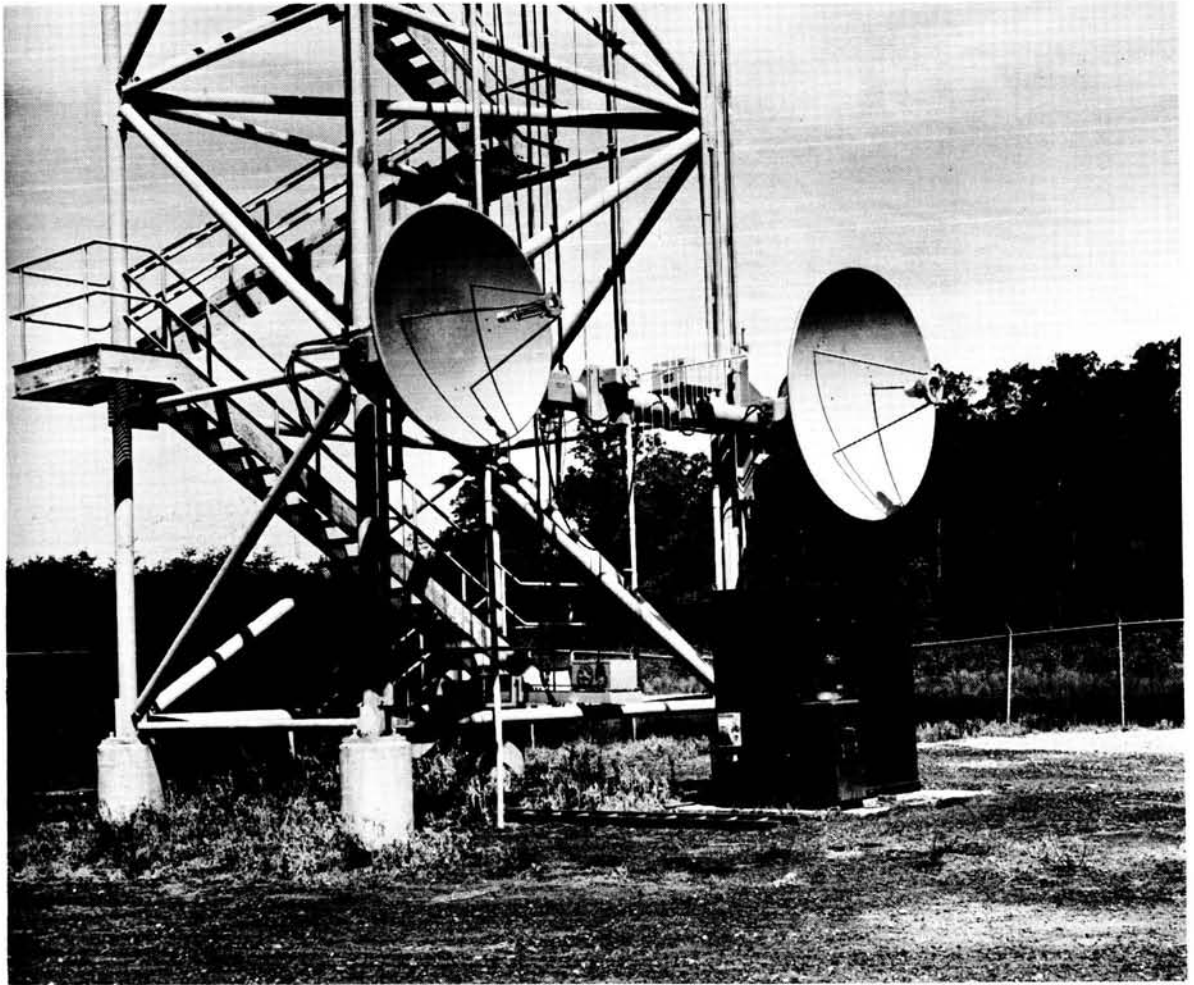


FIGURE 21.—Horizontal range RF field-generating antennas.

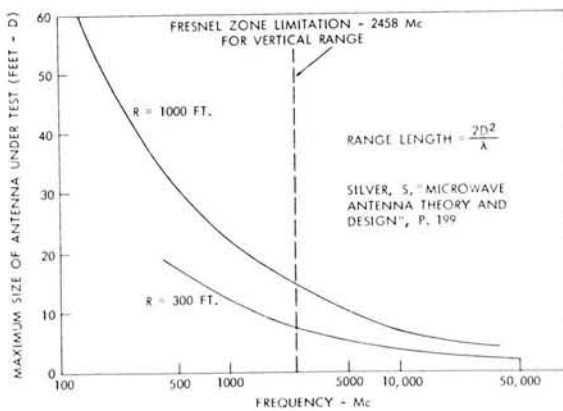


FIGURE 22.—Maximum size for antenna under test so that phase variation is less than 45 degrees (for 1000-foot range).

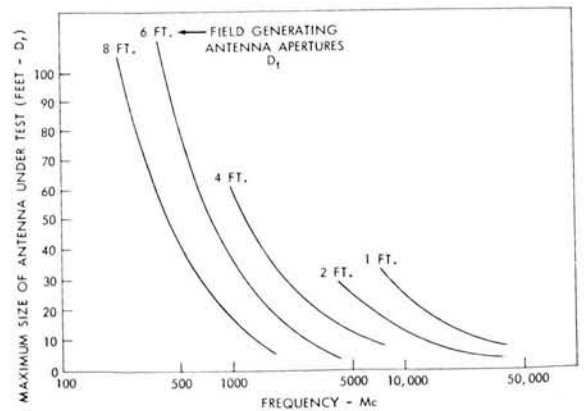


FIGURE 23.—Maximum size for antenna under test so that amplitude variation is less than 0.25 db (for 1000-foot range).

chamber. Orientation of the dipoles was also varied. Transmission from the chamber to the log-periodic structure suspended above the radome was used. Reflectivity levels were observed at three frequencies as shown in Table II.

TABLE II.—Attenuation of Reflected Energy in Vertical Range Anechoic Chamber

Frequency (Mc)	Attenuation (db)
125	28
400	32
1200	40

The chamber is usable over the frequency range 120 to 10,000 Mc and generally provides good performance over this range.

Initially it was feared that edge diffraction from the upper edge of the absorber-covered walls would present problems. Accordingly, the absorber supplier constructed a 1/9-scale model to measure this effect. Results were encouraging. When the full-scale chamber was evaluated, no edge diffraction was observed.

Operational Limits

The vertical range can be considered usable over the frequency band 120 to 10,000 Mc.

Spurious reflections do not occur from the floor, the model positioner, or the field-generating antenna mast. The radome limits performance to 1000 Mc at the edges of the radome especially. Signal variations of about $\frac{1}{2}$ db occur as a result of the log periodic structure swaying in the wind. Antenna models as large as 10 feet in a maximum dimension can be introduced into the anechoic chamber.

HORIZONTAL RANGE SYSTEM SENSITIVITY

Table III summarizes the system sensitivity for the 1000-foot horizontal range, assuming the gain of the antenna under test to be 0 db. Any gain in this antenna over that of an isotropic radiator will increase the usable dynamic signal range by that amount. The values shown for transmitter output and transmitting antenna gain are the published specification values for the Scientific-Atlanta, Inc., equipment used. The receiver sensitivity value are those specified for the Scientific-Atlanta, Inc., Model 1640 APZ Wide Range Receiver.

RADIO FREQUENCY INTERFERENCE SURVEY

To establish the suitability of the range location from a standpoint of radio frequency interference, a survey was made on February 2, 1960, by James F. Goodrich and William A. Morton of Department 474, Bendix Radio Division, Bendix Corporation. A field intensity meter (Empire

TABLE III.—Horizontal Range System Sensitivity

Frequency (Gc/s)	Transmitter Output (dbm)	Transmitting Antenna Gain	Space Attenuation	Receiver Sensitivity	Usable Signal Dynamic Range
0.1	19.5	8	62	100	65.5
0.2	19.5	8	68	100	59.5
0.3	21	8	72	100	57.0
0.4	21	8	74	100	55.0
0.6	21	8	78	100	51.0
1.0	22	27	82	100	67.0
2.0	18.5	31	88	110	67.5
3.0	18.5	34	91	110	71.5
4.0	18.5	36	94	110	70.5
6.0	13	37	98	100	52.0
9.0	13	40	101	100	52.0
10.0	13	40	102	100	51.0
12.0	13	42	103	100	52.0

Devices Model NF-105) covering the frequency range 20 Mc to 1000 Mc was used to measure field strength. This device is a compact, portable instrument which can be operated on a standard 12-volt automobile battery with an inverter. Four turning heads were used housing the IF and

RF circuits for four frequency bands. Table IV lists these heads together with the pick-up antennas used.

The survey was made at a location just south of Beaverdam Road and one half mile east of the intersection of this road with Soil Conservation

TABLE IV.—RF Detection Equipment used with Empire Devices Model NF-105 Field Intensity Meter

Frequency Band	Receiving Head	Antenna
20-30 Mc	Model T-A	Loop Model LP-105
20-200 Mc	Model T-1	Dipole Model DM-105-T-1
200-400 Mc	Model T-2	Dipole Model DM-105-T-2
400-1000 Mc	Model T-3	Dipole Model DM-105-T-3

Service Road. The location is indicated on Figure 1 where it can be seen to be 1300 feet west of the center of the range. This location represented the closest approach available to the range at the time, without damaging a crop of wheat planted in the fields. Table V lists the measured signal levels in microvolts per meter. Correction has been made for cable loss, impedance mismatch, and

effective antenna height. The direction from which the signal arrives is given in degrees from magnetic north and the nature of the signal is described. Continuous tuning over the band 20 to 1000 Mc was done; only those frequencies at which a signal was detected are listed. At all other frequencies in the band 20 to 1000 Mc, the spectrum was quiet.

TABLE V.—Radio Frequency Interference Survey on Antenna Range Facility

Frequency (Mc)	Signal Level (V/m)	Direction	Remarks
20.0.....	1000.....	120°	WWV—vertically polarized
20.5.....	100.....	060	Teletype
20.75.....			Facsimile (audible on phones only)
20.9.....	150.....	045	Mixed beeps
21.....			Foreign voice
21.0 thru 24.0.....	8 thru 200.....		Radio amateurs
24.6.....	1000.....	150	WWV—horizontal polarization
25.3.....	20.....	broad	Carrier
25.7.....	9.....	135	Teletype
26.3.....	35.....	045	Scrambled modulation
26.5.....	20.....	broad	Tunable noise
27.0 thru 29.6.....	8 thru 800.....		Radio amateurs and various voice
29.8.....	3.....	090	WWV—vertical
30.2.....	25.....	060	120-cycle buzz and background noise
32.0*	90.....	360	Mixed voice and noise
34.4*	13.....	broad	Garbled voice
35.0***	30.....	110	Garbled voice
39.8.....	30.....	170	WWV—vertical
43.2.....	250.....	360	Washington radio paging service
44.0 thru 49.0**	15 thru 55.....		Momentary voice (police)
55.0.....	6000.....	020	Sync pulse
59.0.....	6000.....	020	Voice channel

} TV Channel 2

TABLE V.—Radio Frequency Interference Survey on Antenna Range Facility (Continued)

Frequency (Mc)	Signal Level (V/m)	Direction	Remarks
59.2 thru 65.0	Especially quiet		
67.0	30,000	045	Sync pulse
72.0	20,000	045	Voice channel
75.0	180	135	1000-cycle beep, 1-sec intervals
75.3	28	120	Pulse, approx. 300 pps
77.7	14,000	040	Sync pulse
81.5	10,000	040	Voice channel
90.0	35	015	WWV—vertical
92.0	280	090	FM
94.0	5000	070	FM
95.0	35	050	FM
96.0	5500	070	FM
96.2	7000	100	FM
99.0	7500	090	FM
101.0	5000	090	FM
104.0	2200	090	FM
105.0	60	135	FM
107.0	220	080	FM
108.0	2000	080	FM
108.2 thru 150.0	Especially quiet		
133.0	2200		Tunable ignition noise
134.0	350	360	WWV—horizontally polarized
135.5	21,000		Momentary voice
137.0			Momentary voice (on phones only)
140.0			Weak carrier (on phones only)
150.0	55	150	1000-cycle beep, 1-sec intervals
153.0	110		Carrier
154.0	85	060	Sync pulse
158.0	100	080	Voice channel
164.0	140	050	Sync pulse
169.0	130	050	Voice channel
175.0	22,000	075	Sync pulse
179.0	14,000	075	Voice channel
188.0	25,000	060	Sync pulse
192.0	20,000	060	Voice channel
198.0	8000	030	Sync pulse
203.5	8000	030	Voice channel
213.0	39,000	030	Sync pulse
217.0	2000	030	Voice channel
252.0	100		
267.0			
273.0	190		
277.5	900	300	
278.0			
282.0	220		
283.0	1100		
294.0	90		
297.0			
307.0	1300		
315.0			
319.0	280	085	
320.0			
337.0			

Various intermittent airground voice communications, most of which were too brief to determine direction or intensity.

TABLE V.—Radio Frequency Interference Survey on Antenna Range Facility (Continued)

Frequency (Mc)	Signal Level (V/m)	Direction	Remarks
339.0			Various intermittent airground voice communications, most of which were too brief to determine direction or intensity.
351.0			
358.0			
367.0			
388.0			
452.0			
500.0			Intermittent weak carrier 200 to 400 prf short time pulse no noted signals.
500.1 thru 1000.0			

*In the band 32.0 to 34.0 Mc, 42-db ignition noise was noted during passage of aircraft overhead.

**In the band 44.0 thru 49.0 Mc, ignition noise from passing trucks was heard.

***In the band 37.0 to 38.0 Mc, a classified Department of Defense signal was read.

ACKNOWLEDGEMENTS

The assistance of personnel of the Scientific-Atlanta Corporation in the writing of this document is acknowledged and greatly appreciated.

REFERENCES

1. RALEIGH, LORD, "On Pin-Hole Camera Photography," Philosophical Magazine, 1891, Volume 31, p. 87.
2. CUMMING, W. A., "Radiation Measurements at Radio Frequencies," Proc. IRE, 1959, Vol. 47, p. 705.
3. HARVEY, A. F., "Microwave Engineering," Academic Press, 1963, p. 587.
4. SILVER, S., "Microwave Antenna Theory and Design," McGraw-Hill Book Co., New York, 1949, pp. 574-576.
5. MONTGOMERY, C. G., "Technique of Microwave Measurements," McGraw-Hill, 1947, p. 902.
6. STONE, JOHN M., "Radiation and Optics," McGraw-Hill, 1963, p. 143.
7. KORVIN, W., and STECKEL, J., "A Vertical Test Range for Antenna Radiation Measurements," NASA/GSFC Report X-632-64-234, July 1964.

HANDBOOK OF NASA/GSFC TRACKING DATA-ACQUISITION, AND COMMUNICATIONS ANTENNAS*

PAUL A. LANTZ

FOREWORD

Research workers in many diverse areas require a generalized compilation of the characteristics of antennas located at facilities which provide tracking and data acquisition for National Aeronautics and Space Administration (NASA) satellites. It is hoped that this handbook, which includes relevant portions of a volume previously published in November 1961, will serve as a reference handbook for those persons planning projects which will utilize these antennas as system components.

This publication combines and replaces GSFC Document X-525-61-00, "Handbook of Antennas at NASA Satellite Tracking Stations," and GSFC Document X-525-64-168, "Handbook of Antennas at NASA Satellite Tracking and Space Data-Acquisition Network (STADAN) Facilities."

Because of the time lag involved in publishing a document of this type, it is possible to approach only the most recent information regarding antennas now in operation, or those planned for operational use. From time to time, it will be necessary to publish revisions to include information relative to antenna modifications and new antenna types.

A NOTE ON ANTENNA GAIN

Antenna gain can be calculated with considerable accuracy by somewhat complicated and laborious methods; some conventional forms of antennas can be measured also, though not always accurately. On the other hand, approximate gain can be calculated from rather simple formulas which are useful for estimating range capability when antenna beamwidths are specified. For practical antennas (excluding the impractical

super-gain types) the gain and beamwidth are related as follows

$$G = k \frac{41,253}{\theta_v \times \theta_h}$$

where G is the power gain; 41,253 is the volume of a hemisphere in square degrees; θ_v and θ_h are the horizontal and vertical half-power beamwidth; and their product is the volume of the solid-angle pattern in square degrees. The constant k , called the "aperture efficiency," theoretically has a maximum possible value of 1; in practice, it may reach values as high as 0.9 (though this is rare), and for well-designed antennas of conventional beam shape it is usually equal to, or greater than 0.6. A typical value for a horn-fed paraboloidal reflector is 0.65, and a dipole array of good design (e.g., Dolph-Tschebyscheff† current distribution) may have $k = 0.85$. For a given general type and size of antenna, the gain of the main beam and the suppression of unwanted sidelobes are inversely related; for example, the extreme aperture efficiency of 0.9 would be obtained only at the expense of rather high sidelobes. Sidelobes are suppressed by tapering the illumination of a reflector or the current distribution of an array; this naturally results in a reduced efficiency of use of the available antenna area (i.e., wider beamwidth and lower gain, as compared with uniform illumination).

Gain measurements were made for most of the antennas described in this report. In several cases, gain was calculated using the gain-beamwidth relationship. For this purpose, the value $k = 0.85$ was used for the planar arrays, and the value $k = 0.76$ was used for the yagi arrays. The

†Dolph, C. L., "A Current Distribution for Broadside Arrays which Optimizes the Relationship between Beam Width and Side-Lobe Level," Proc. I.R.E., Vol. 34, June 1946, pp. 335-48.

*Published as *Goddard Space Flight Center Document X-525-64-222*, October 1964.

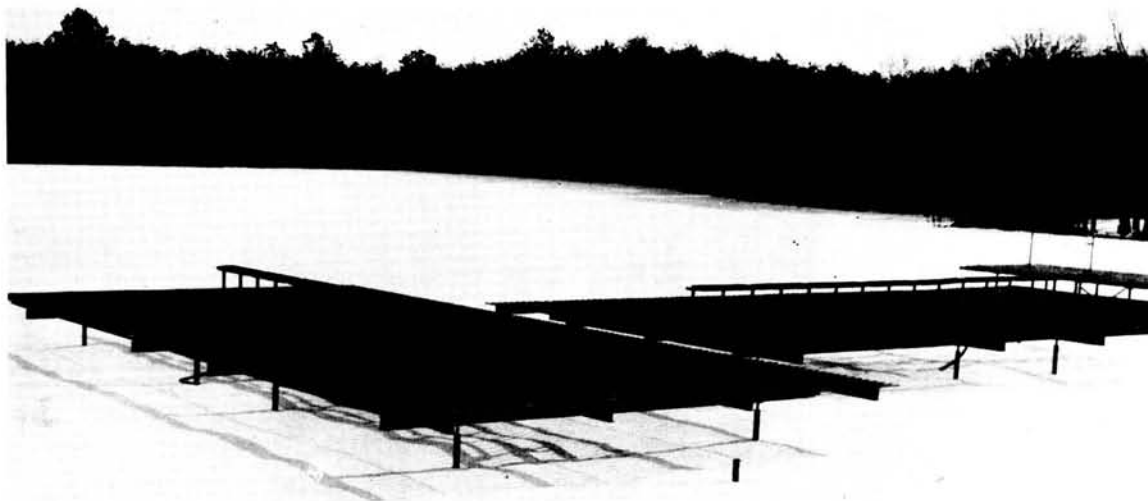


FIGURE 1.—136-Mc minitrack antenna at Blossom Point.

assumed (in some cases, computed) aperture efficiency of the other antenna types is indicated in the body of the report.

I. MINITRACK ANTENNA

136-MC Tracking Antenna

Technical Appliance Corporation
 Sherburne, New York
 NAS5-18(W)

General Description

The 136-Mc tracking antenna is a colinear array of eight 10.5- \times 49.7-inch skeleton slots spaced 55.6 inches (0.64λ) apart and 23.7 inches (0.27λ) above a common 25- \times 50-foot ground screen. Figure 1 shows antennas installed at Blossom Point, Maryland. The current distribution is a 22-db Dolph-Tschebyscheff taper, shown in Figure 2. Figure 2 shows one-half of the symmetrical feed system, which is constructed of $\frac{7}{8}$ -inch rigid coaxial line. Also shown are branch-point impedances to produce the desired power distribution and quarter-wave transformer impedances required to match them to a 50-ohm line. A modified British split balun¹ is used to convert from balanced to unbalanced line, eliminating troublesome RF outside the coaxial lines. The characteristic line impedance in the balun is designed to match to the feed-point impedance,

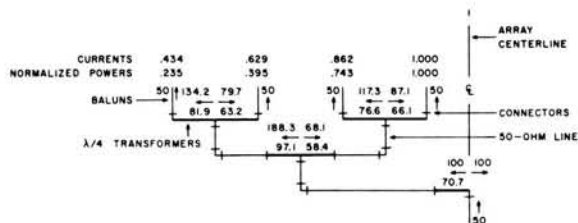
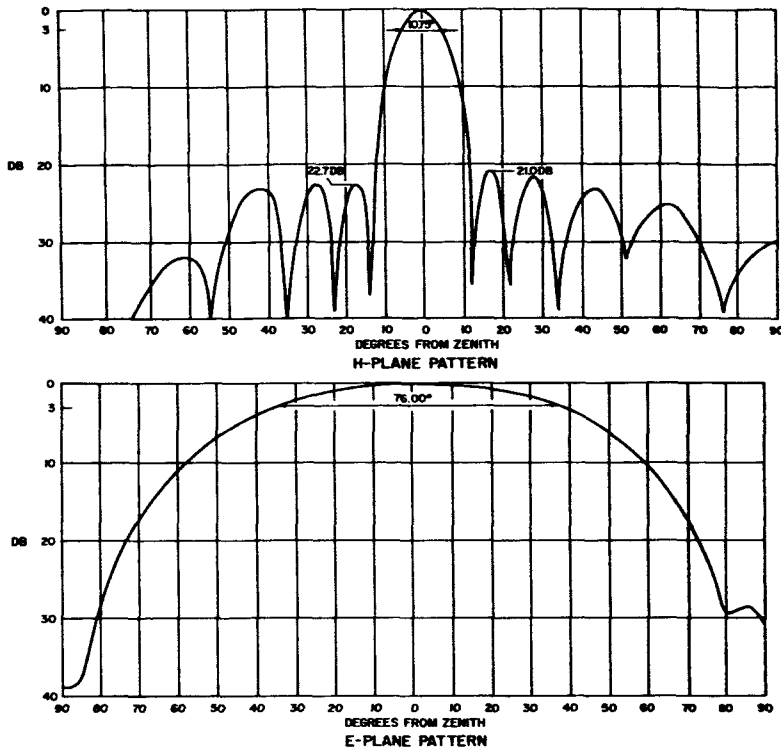


FIGURE 2.—Schematic of coaxial feed system, 136-Mc minitrack antenna.

which is approximately 924 ohms. The antenna was designed at the Physical Sciences Laboratory, New Mexico State University, and manufactured at the Technical Appliance Corporation, where it was designated Model G-1284.

Pattern Performance

The H-plane pattern (Figure 3) of this antenna was measured by rotating the full-scale antenna located on a range with a sufficient depression angle (7 degrees) to minimize reflections from the ground transmission path. The beamwidth was found to be 10.75 degrees. The first sidelobes were below 21 db, which is commensurate with design criteria. To measure the E-plane patterns (Figure 3), a single slot was mounted on an 18- \times 25-foot ground screen. The beamwidth was 76 degrees wide, and the sidelobes were



GAIN: 16.3 DB ABOVE AN ISOTROPIC SOURCE

FIGURE 3.—E- and H-plane patterns, 136-Mc minitrack antenna serial no. 001.

measured below 22.5 db. The cross-polarized component of radiation was below 30 db.

Gain

The manufacturer computed the gain of this antenna by means of the gain-beamwidth relationship

$$G = k \frac{41,253}{\theta_E \times \theta_H} \quad (1)$$

and found it to be 36.3, or 15.6 db. However, an efficiency factor of 0.7 was applied in this computation in an effort to correct for losses. This reduction is considered to be too stringent; in fact, volume integration² of the pattern indicates a gain of 42.8, or 16.3 db, which is considered to be more valid. This represents an 85 percent aperture efficiency, which is reasonable for a tapered array.

Input Impedance and Voltage/Standing-Wave Ratio (VSWR)

Figure 4, a plot of the input impedance to this antenna, shows that the bandwidth over which

the VSWR is less than 1.15 (1.2 db) is approximately 3 Mc.

Phase and Amplitude Distribution

The amplitude of currents in the slot elements of this array is tapered outward from the array center in accordance with a Dolph-Tschebyscheff distribution for 22-db sidelobes. Table 1 gives the theoretical distribution for an eight-element array. The developmental model of the antenna coaxial system was hand-built to provide the

TABLE 1.—Optimum Current Distribution for Eight-Element Array to Yield 22-db Sidelobes

Slot No.	Current Coefficient
4 and 5.....	1.0000
3 and 6.....	0.8618
2 and 7.....	0.6287
1 and 8.....	0.4844

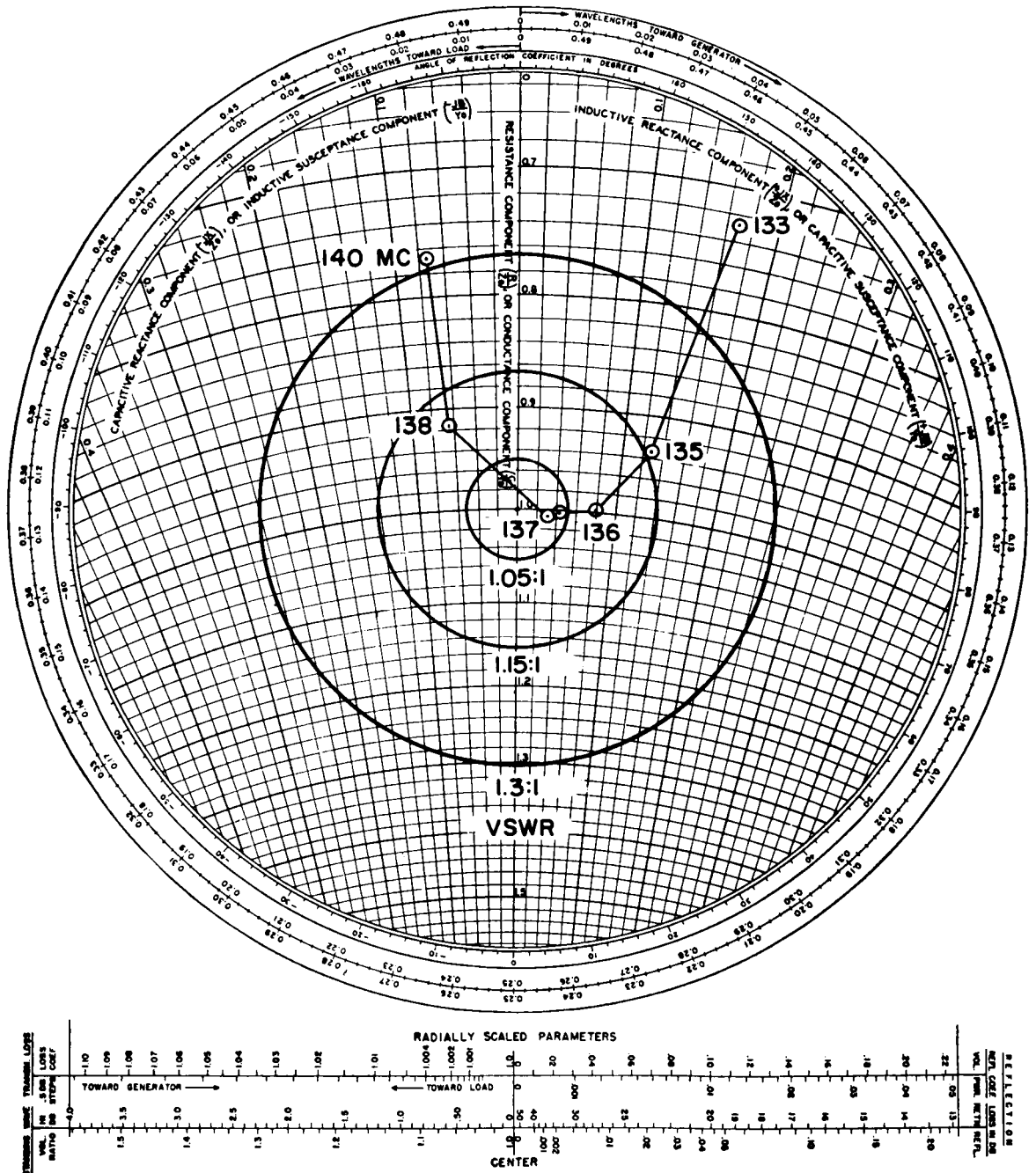


FIGURE 4.—Input impedance, 136-Mc minitrack antenna serial no. 001.

theoretical current distribution; however, it was found economical to relax tolerances in production. This theoretical amplitude distribution is plotted in Figure 5, bounded by arbitrarily selected limits. Computer analysis of the pattern resulting from any distribution falling within these tolerances indicated a suitable performance.

None of the antennas was observed to exceed these limits.

The phase variation across the aperture was specified to be ± 3 degrees. Figure 5 illustrates a typical phase variation measured on an antenna which, in most cases, was observed not to exceed ± 2.0 degrees.

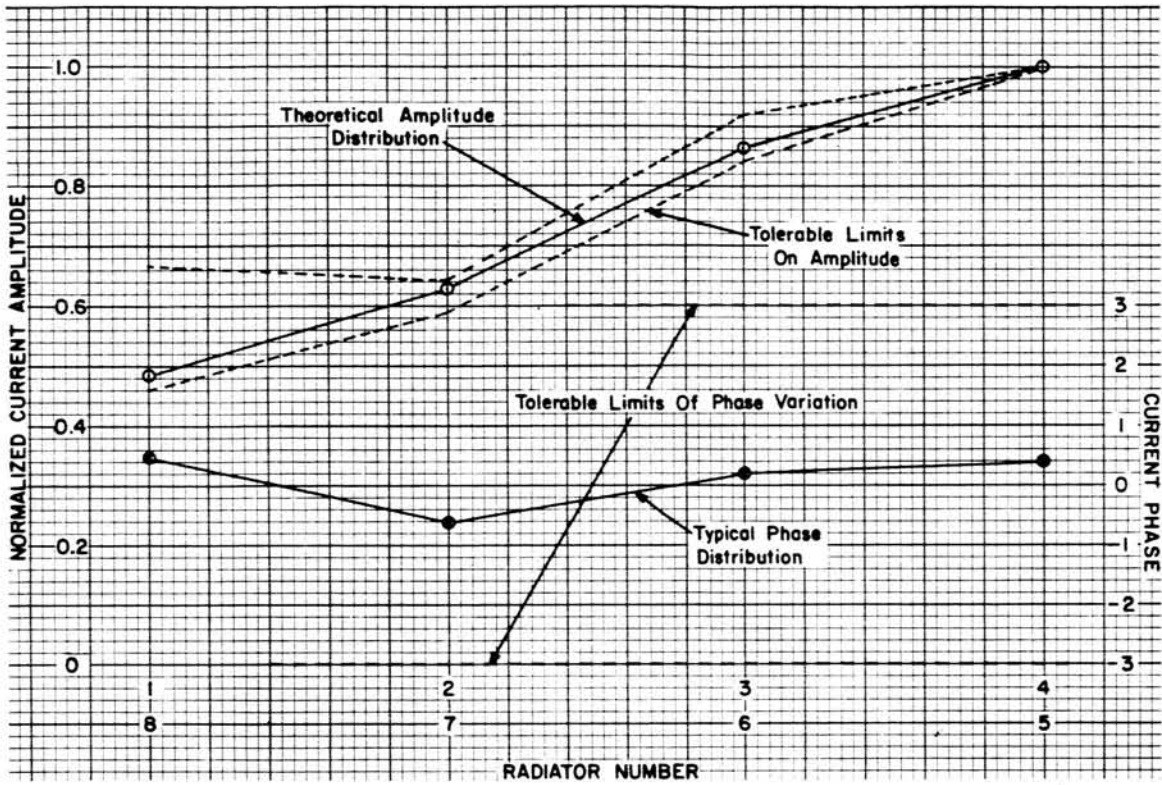


FIGURE 5.—Phase and amplitude variation, 136-Mc minitrack antenna.



FIGURE 6.—136-Mc ambiguity-resolving antennas.

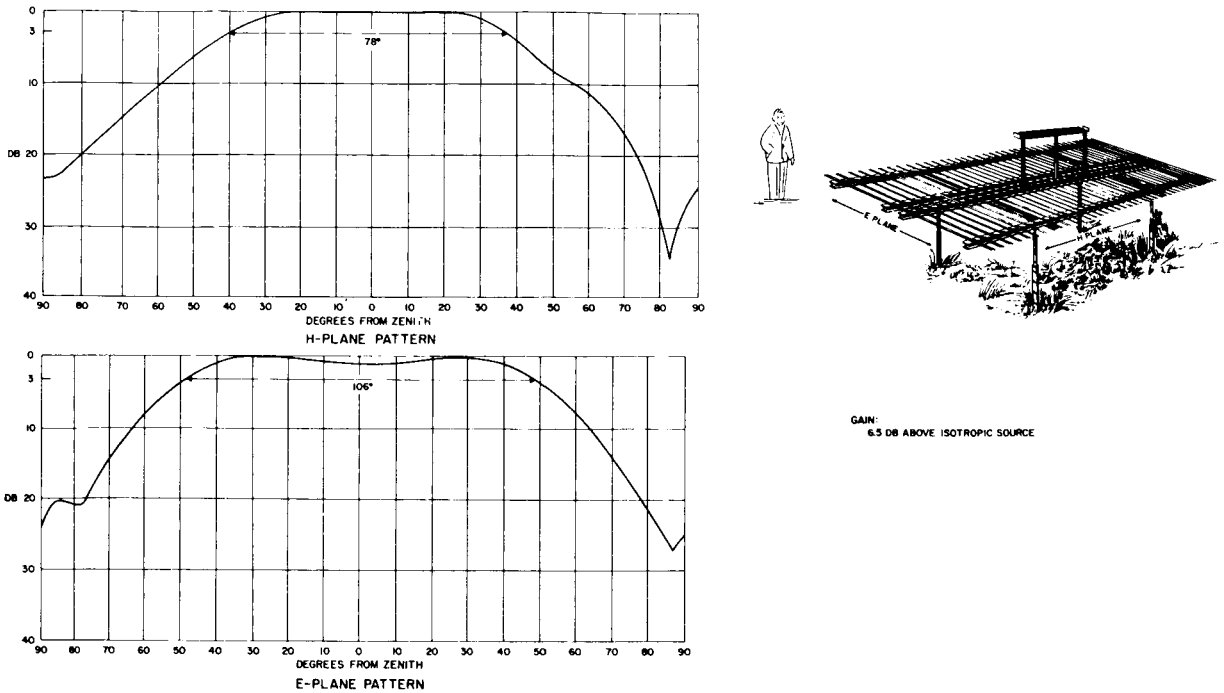


FIGURE 7.—E- and H-plane patterns, 136-Mc ambiguity-resolving antenna serial no. 001.

Ambiguity-Resolving Antenna

Technical Appliance Corporation
 Sherburne, New York
 NAS5-18(W)

General Description

The ambiguity antenna for the 136-Mc system consists of a single skeleton slot mounted on an 18- \times 18-foot ground screen. The antenna is illustrated in Figure 6. The screen spacing was reduced to 0.25λ , instead of the 0.35λ spacing used in the Minitrack antenna, to increase the H-plane beamwidth and thereby to assure sufficiently early reception from each satellite pass. Figure 7, a plot of the patterns of the ambiguity antenna, indicates the E-plane beamwidth to be 106 degrees, and the H-plane beamwidth to be 78 degrees; the theoretical gain, based on 85 percent efficiency, is 6.4 db. The input impedance is plotted in Figure 8, which shows that the antenna met the specified VSWR 1.15:1 (1.2 db) over the 136 ± 0.5 -Mc band.

Low-Angle Tracking Antenna

Technical Appliance Corporation
 Sherburne, New York
 NAS5-1527

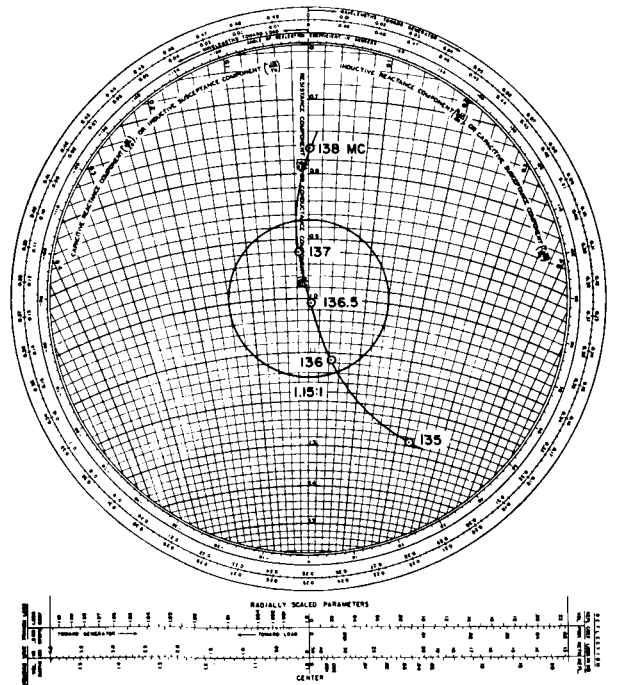


FIGURE 8.—Input impedance, 136-Mc ambiguity-resolving antenna serial no. 001.

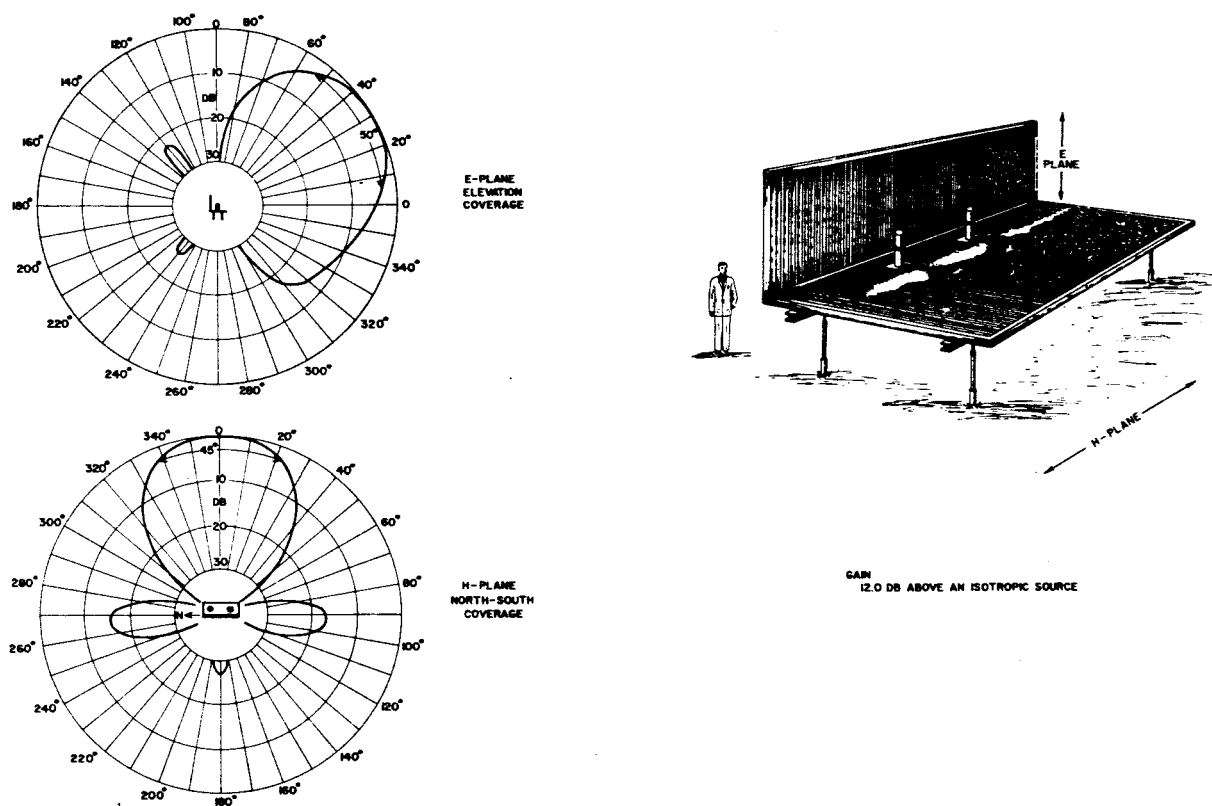


FIGURE 9.—E- and H-plane patterns, 136-Mc low-angle minitrack antenna.

General Description

The Fort Myers satellite tracking and data-acquisition network (STADAN) facility provides tracking information on Air Force Eastern Test Range before third-stage injection. To obtain immediate post-launch tracking data, a specialized Minitrack antenna was developed for use in an interferometer at that location. The antenna, Figure 9, consists of a pair of vertical linear elements (short dipoles) mounted 0.5λ apart on a ground plane 15 feet wide and 14 feet deep, and 0.25λ in front of a vertical reflector plane 8 feet high.

Figure 9 shows the pattern performance of this antenna. The H-plane (azimuth) beamwidth is shown to be 45 degrees. Hence, two antennas located end to end, looking away at a separation angle of 45 degrees, will cover a 90-degree azimuth sector up and down the coast. Eight of these antennas form two orthogonal baselines 50 wavelengths long. Elevation coverage is good up to approximately 50 degrees, an angle sufficient for

tracking first- and second-stage vehicles. The theoretical gain of the antenna, based on an 85 percent aperture efficiency, is 12.0 db.

II. PARABOLOID REFLECTOR ANTENNAS

40-Foot Mojave Antenna

Philco Western Development Laboratories
Palo Alto, California
NAS5-1757

General Description

This antenna (Figure 10) is an X-Y mounted paraboloidal reflector ($f/D=0.40$) with a circular aperture 40 feet in diameter, for communications by means of the Relay satellite. It is equipped with a prime-focus amplitude-monopulse feed system operating at 4080 Mc for tracking, and at 4170 Mc for receiving communications data. This feed also transmits communications data to the satellite at 1725 Mc with average 10-kw continuous-wave (CW) power. A 148-Mc command-transmitting element was planned but deleted.



FIGURE 10.—40-foot paraboloid reflector, Mojave Satellite Tracking and Data-Acquisition Network (STADAN) Facility.

The monopulse system operates in either sense of circular polarization, but both senses cannot be supplied simultaneously. Transmission is also accompanied in either of the two circular senses.

The transmitter is mounted above the Y or upper (east-west) axis, eliminating the need for rotary joints. Aluminum waveguide (type RG-105/U) connects the transmitter to the feed.

The reflector surface is solid. Aluminum honeycomb panels, doubly curved, provide a contour accuracy not exceeding 1/32 inch rms from the least-square best-fit paraboloid, permitting operation up to 10,000 Mc. The feed is supported at the focus on a quadripod built to carry a 400-pound load (feeds, parametric amplifiers, comparators, etc.). Lateral and axial adjustment of the feed is provided up to ± 2 inches.

The dual-purpose feed consists of four centrally located square-aperture waveguides for monopulse operation at 4080 Mc. These are heavily loaded with teflon and spaced 0.408λ on centers to permit proper location of the surrounding transmit feed. These horns taper to a 2- \times -2-inch square waveguide and are matched to freespace by square metal plates inserted in the teflon loading. Four horns, spaced 0.504λ on centers, surround the monopulse feed and transmit at 1725 Mc. These are ridge-matched to an RG-105/U waveguide which in turn is fed with a 1 $\frac{5}{8}$ -inch coaxial line. Glass-reinforced teflon plates located diagonally across square sections of the horns are manually rotated within 90 degrees to change polarization. Table 2 shows operational frequencies.

The pedestal (built by Todd Shipyard for the prime contractor, Philco) is a hydraulically driven X-Y mount capable of positioning the RF beam and reading out this position to an accuracy of ± 60 seconds of arc. The tracking rate is variable from 0 degrees to 5 degrees per second, and the drive is capable of acceleration or deceleration up to 5 degrees per second.² Sky coverage is limited only by a 10-degree (radius) solid-angle keyhole about the X-(lower) axis. The X-axis is oriented

to be north-south, and the X-Y-axis orthogonality is held to 1.3 seconds. This antenna is located at the Mojave STADAN station.

Pattern Performance

Primary patterns of the feed were measured by the prime contractor, Philco Western Development Laboratories, Palo Alto, California. Monopulse E- and H-plane primary sum patterns (Figure 11) were measured at 4080 Mc. The angle subtended at the focal point by a paraboloidal reflector can be computed from the relationship given by Silver:³

$$\tan \theta = \frac{1}{2} \left[\frac{D/f}{1 - D^2/16f^2} \right], \quad (2)$$

and θ is found to be 64 degrees when $f/D=0.4$. The edges of the reflector are indicated on the primary pattern (Figure 11). The indicated levels on the primary patterns must be corrected by the space attenuation to derive the reflector-edge illumination. Jasik⁴ relates space attenuation to subtended angle

$$A = 20 \log \sec^2 \frac{\theta}{2}, \quad (3)$$

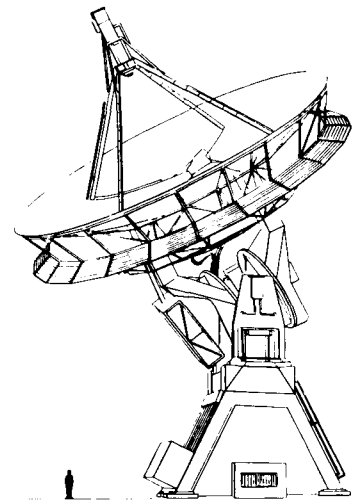
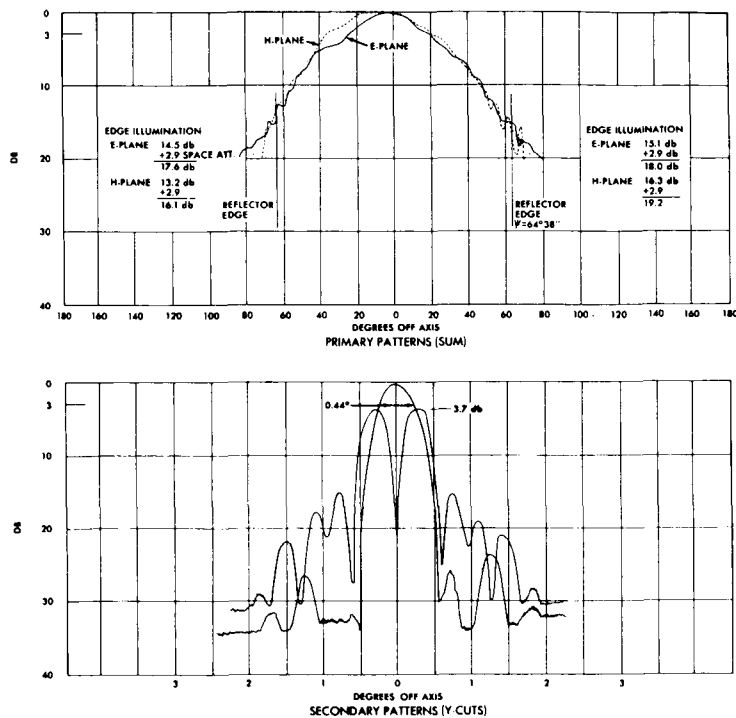
and for a 64-degree subtended angle, the space attenuation is 2.9 db. Hence, the average edge illumination is 11.8 db, approximately the optimum (11.5 db) for good monopulse characteristics established by Crompton.⁵

Although primary patterns of the 1725-Mc transmit feed were also measured, the patterns are not shown in this report. Edge illumination was approximately 18.0 db in the E-plane and approximately 8.0 db in the H-plane. No reason is known for this discrepancy.

Secondary patterns of the 4080-Mc monopulse feed measured with the feed installed in the reflector at Goldstone, California, are also shown in Figure 11. The sum channel half-power beamwidth is 0.44 degree with sidelobes down to approximately 26.0 db. The difference-pattern peak gain is 3.7 db below the sum-pattern peak. The difference slope, normalized to an isotropic radiator, is 23.5 db per 0.1 degree, or 15.0 voltage ratio per 0.1 degree.

TABLE 2.—Relay Operational Frequencies

Frequency (Mc)	Function
1725.....	10-kw transmit (one way, i.e. television)
1723.33....	10-kw transmit (two way, i.e. telephone)
1726.67....	
4170.0.....	One-way reception
4165.0.....	Two-way reception
4175.0.....	
4080.....	Phase-monopulse tracking



NOTES

1. PATTERNS AND GAIN MEASURED AT 4080 Mc
2. MEASURED GAIN SUM: 31 db ABOVE ISOTROPIC
DIFFERENCE: 48 db ABOVE ISOTROPIC

FIGURE 11.—Patterns, 40-foot data-acquisition paraboloid, Mojave STADAN Facility.

Gain

The sum channel of the monopulse feed was measured at 4080 Mc by comparison with a calibrated standard horn and found to be slightly greater than 51.0 db. This represents an aperture efficiency of approximately 55 percent.

Noise Temperature

The noise temperature at 4080 Mc of the antenna (exclusive of the receiver and other system components) was computed to be 42°K at 30-degree elevation angle, at which point the temperature can be expected to rise sharply because of the increased atmospheric contribution. Waveguide loss, receiver noise, and the variation of reflector coefficient of the ground will increase these temperature levels. Heavy rain can be expected to increase noise temperature an additional 50°K at 4080 Mc. This estimate of the antenna temperature is to be confirmed by direct measurement with a Dicke radiometer.

Miscellaneous Characteristics

The axial ratio for the monopulse feed was measured to be less than 0.9 for both senses over

the entire range 4080 to 4180 Mc. The VSWR of the transmitting feed was measured to be 1.05. Dry nitrogen fills the entire feed system at a minimum internal pressure of 0.5 psi.

40-Foot Data-Acquisition Facility Antennas

Philco Western Development Laboratories
Palo Alto, California
NAS5-1757

General Description

These antennas are basically identical to the 40-foot Mojave antenna previously described. Three antennas have been constructed for use at the data-acquisition facilities (DAF) at Santiago, Chile; Johannesburg, South Africa; and Quito, Ecuador. The feed of the three antennas differ entirely from the feed of the Mojave antenna, and is designed for operation at the frequencies shown in Table 3. The feed box weighs 800 pounds and contains much equipment identical to that used in the feed box of the 85-foot paraboloid reflector antenna at Fairbanks, Alaska. To carry this load, struts have been enlarged greatly beyond those used in the previously described 40-foot antenna.

TABLE 3.—40-Foot DAF Antenna Feeds

Frequency (Mc)	Function
136.5...	Receives tracking beacon and telemetry
400.0...	Receives tracking beacon and telemetry
1705.0...	Reserve, data acquisition (not instrumented)
2270.0...	Transmits 10-kw average for Goddard range and range rate
148.0...	Appended command element (described later)

Feeds for the three receiving frequencies are monopulse feeds. At 136 Mc, four crossed-dipole turnstiles are spaced 44 inches (0.509λ) apart on a square expanded metal screen 87 inches on a side. The resultant 4.2 percent aperture blockage is moderately high but not excessive. Slight loss in gain (about 0.1 db), slight increase in side-lobe level, and slight null filling are the consequences of the resultant quadratic-phase cancellation in the blockage area. The turnstile is swept back in the usual manner to improve the primary form factor and to pull in the phase center toward the focal point, 6.5 inches in front of the screen.

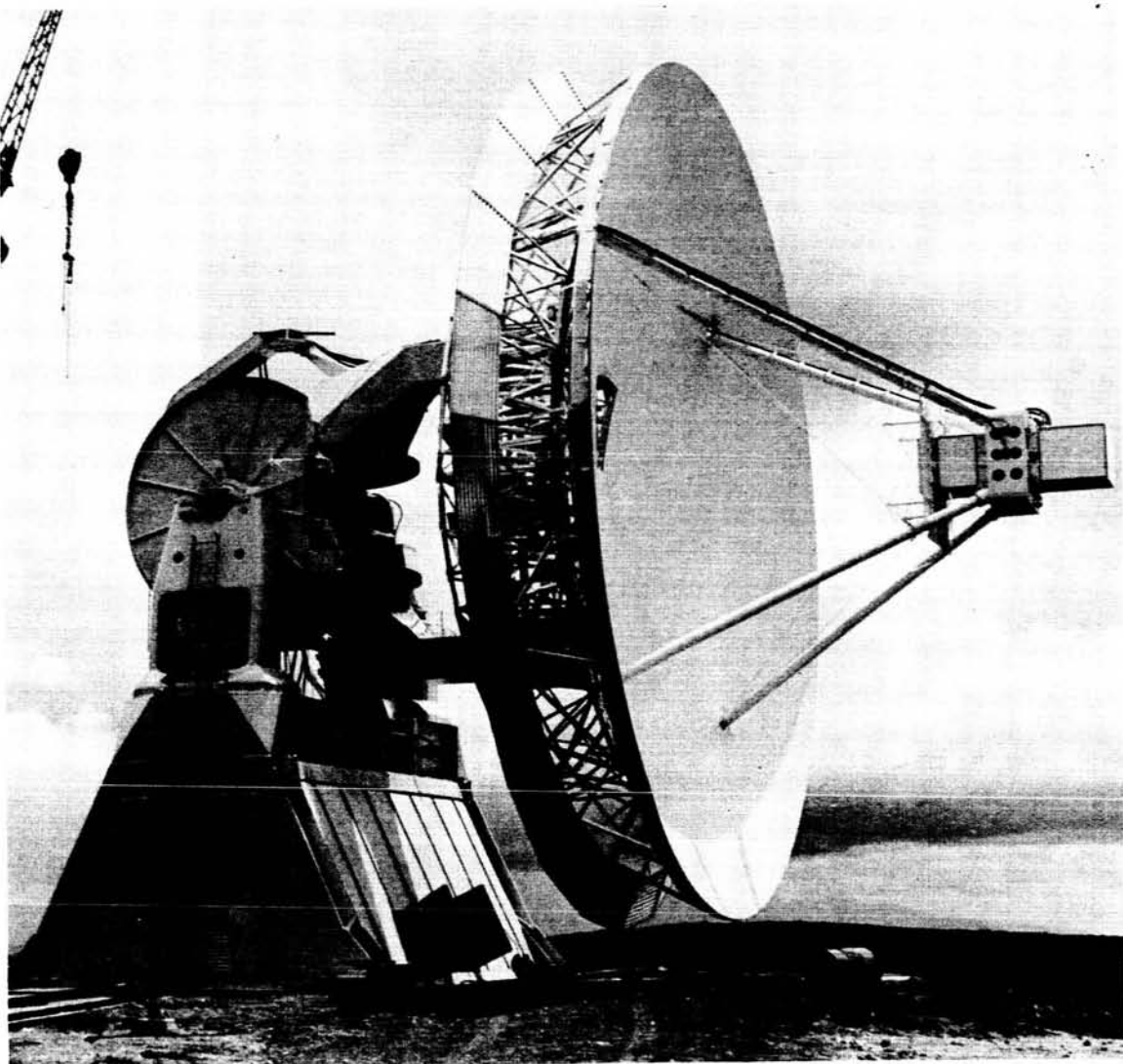


FIGURE 12.—40-foot data-acquisition paraboloid, Quito, Ecuador.

In addition, the back sweeping removes the 136-Mc dipole from the vicinity of the support strut. Except for scaling, the 400-Mc feed is identical to the 136-Mc feed. The 1705-Mc feed consists of four crossed slots backed with coaxial-fed cavities; these slots are oriented 45 degrees with respect to the other feeds and are spaced 9 inches (1.3λ) on centers. No preamplifier or similar instrumentation is provided with these feeds. The 2270-Mc transmit feed is a crossed slot which forms the aperture of an on-axis horn, tapered down from round waveguide to end in this cross-shaped slot. Manually adjusted polarizers in the round guide permit establishment of either sense of circular polarization. The transmit crossed slot separates the 1705-Mc cavities. All the slots are covered with a single flat-sheet radome and are located on focus. Figure 12 is a photograph of the antenna installed at Quito, Ecuador. Figure 13 is a schematic diagram of the feed arrangement. Primary patterns were measured for the 2270-Mc transmit feed only.

Pattern Performance

An attempt was made to measure the secondary patterns of the antenna installed at Quito, Ecu-

ador; however, the results demonstrate the futility of attempting to make such a measurement by using signal sources in ground-based calibration towers. At Quito, the tower from which the 136-Mc RF field is generated is located 500 feet away from the antenna ($2D^2/\lambda = 442$ feet) at an elevation angle of 19 degrees above the horizon. The tower from which the 400-Mc RF field is generated is located 1300 feet from the antenna ($2D^2/\lambda = 1302$ feet); however, because the terrain drops off, it is on the horizon. However, both locations introduce serious multipath reflections from the ground, especially the location of the long-range tower. For this reason, only patterns measured for 136 Mc are included in this handbook; the others are not considered valid.

Figure 14 is a plot of measured 135-Mc sum patterns for right-circular polarization about the upper Y-axis and lower X-axis; the measured half-power beamwidth is 12 degrees, and sidelobes are down only about 10.0 db. (On a later occasion, when a calibration was made by an aircraft flyover at 6000 feet, sidelobes appeared to be down 16.0 db.) High shoulders on the X-axis pattern, and the cutoff a few degrees off-axis on the lower side are consequences of ground reflec-

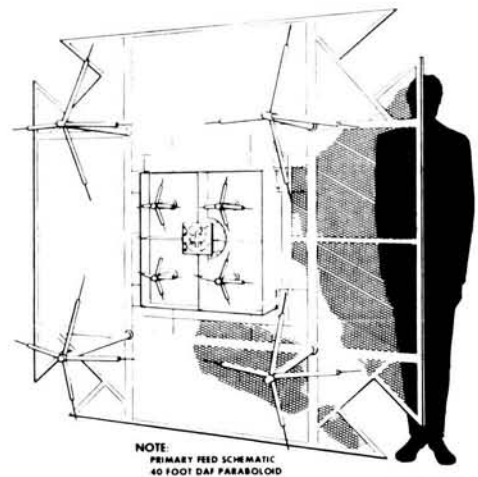
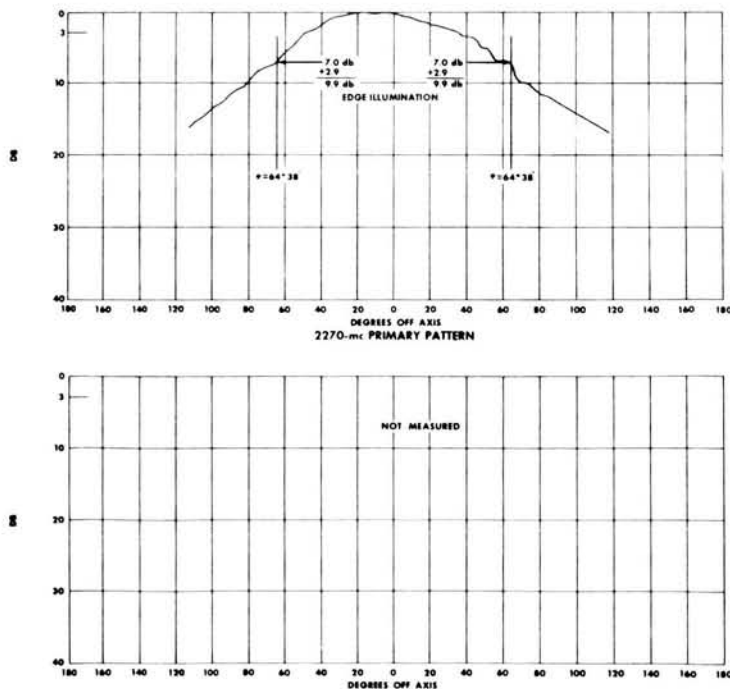


FIGURE 13.—Primary patterns, 40-foot DAF paraboloid.

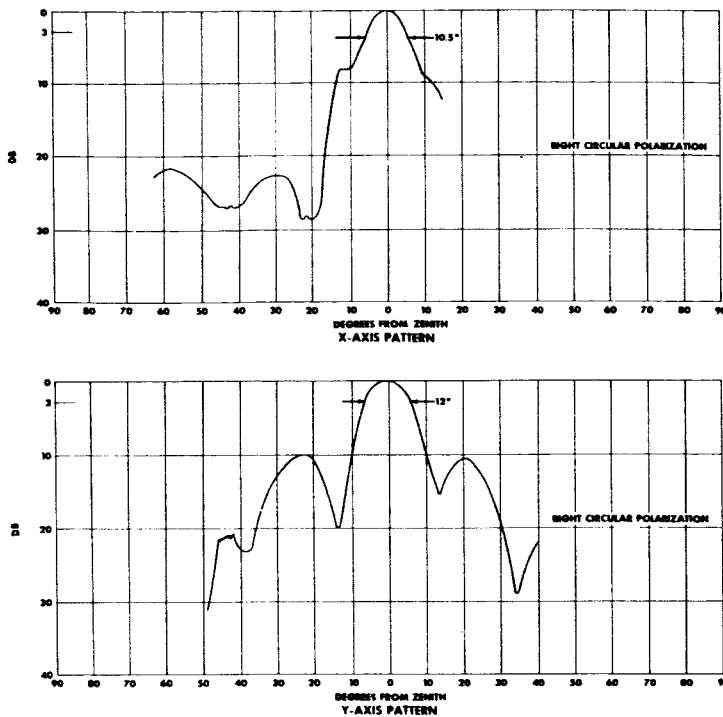
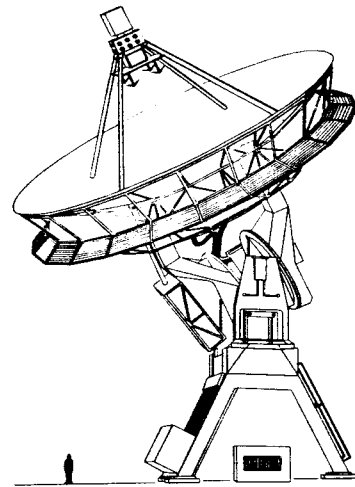


FIGURE 14.—136-Mc sum patterns, 40-foot paraboloid for Data-Acquisition Facility, Quito, Ecuador.



NOTE:
MEASURED GAIN 19.0 db ABOVE AN ISOTROPE

tion. Figure 15 is a plot of difference patterns measured at 136 Mc for right circular polarization. The null depth is 35.0 db and the slope is 4.8-volts ratio per degree, an excellent result. Again, the ground return was an impediment to measurement of the X-axis.

Patterns measured at 400 Mc (Figure 16) are not presented because of serious ground-return distortion. The antenna is fitted with a feed to receive 1705 Mc, although no RF equipment is provided at this time.

Gain

Gain was measured by flying the calibration aircraft at 6000 feet directly over the antenna, which was locked in the zenith position. Gain standards were located beside the antenna, and the signal received in all antennas was recorded on pattern recorders. The standards were Technical Appliance Corporation, 6-turn helices over ground planes; these devices are designated at TACO Model G-1229 at 136 Mc, and G-1257 at 400 Mc. The gain of the antenna at 400 Mc was 30.9 db, representing a 47 percent aperture efficiency. At 136 Mc, a gain of 19.0 db was observed; however,

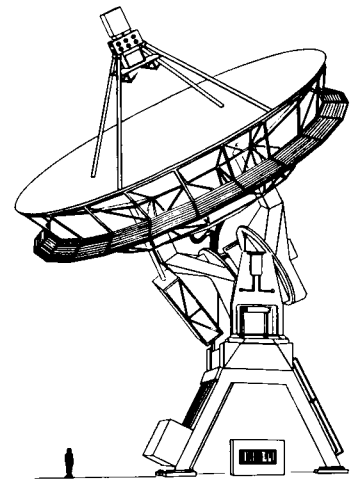
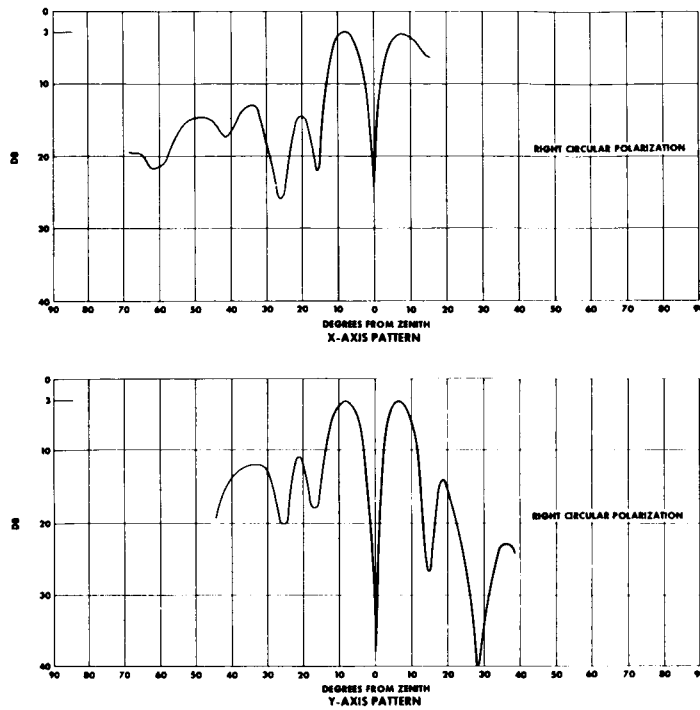
the calibration of the standard was observed to be 2.0 db too low. The exact gain value will not be determined until after the other two antennas (Santiago and Johannesburg) are measured, and the standard is recalibrated.

85-Foot Antennas

Blaw-Knox Company
Pittsburgh, Pennsylvania
NAS5-484
NAS5-3543
Rohr Aircraft Corporation
San Diego, California
NAS5-2065

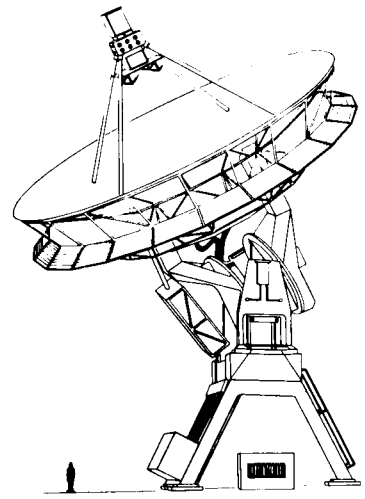
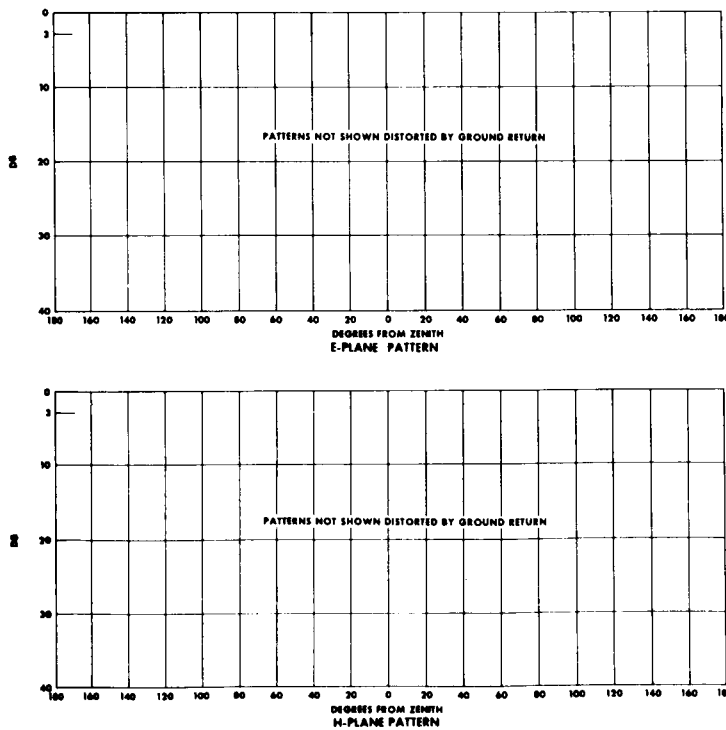
General Description

These antennas (Figure 17) are focal-point-fed solid-surface aluminum paraboloidal reflectors ($f/D=0.423$) with circular apertures 85 feet in diameter. Each section of the reflector surface can be adjusted individually. A hydraulic drive moves the X-Y-mounted antenna about the upper Y-(north-south) axis as well as the orthogonal, lower X-axis. A gimbal-lock zone occurs in the direction of the primary (X) axis, as is



NOTE:
MEASURED GAIN: 19.0 db ABOVE AN ISOTROPE

FIGURE 15.—136-Mc difference patterns, 40-foot paraboloid for Data-Acquisition Facility, Quito, Ecuador.



NOTE:
MEASURED GAIN: 30.9 db ABOVE AN ISOTROPE

FIGURE 16.—400-Mc patterns, 40-foot paraboloid for Data-Acquisition Facility, Quito, Ecuador.

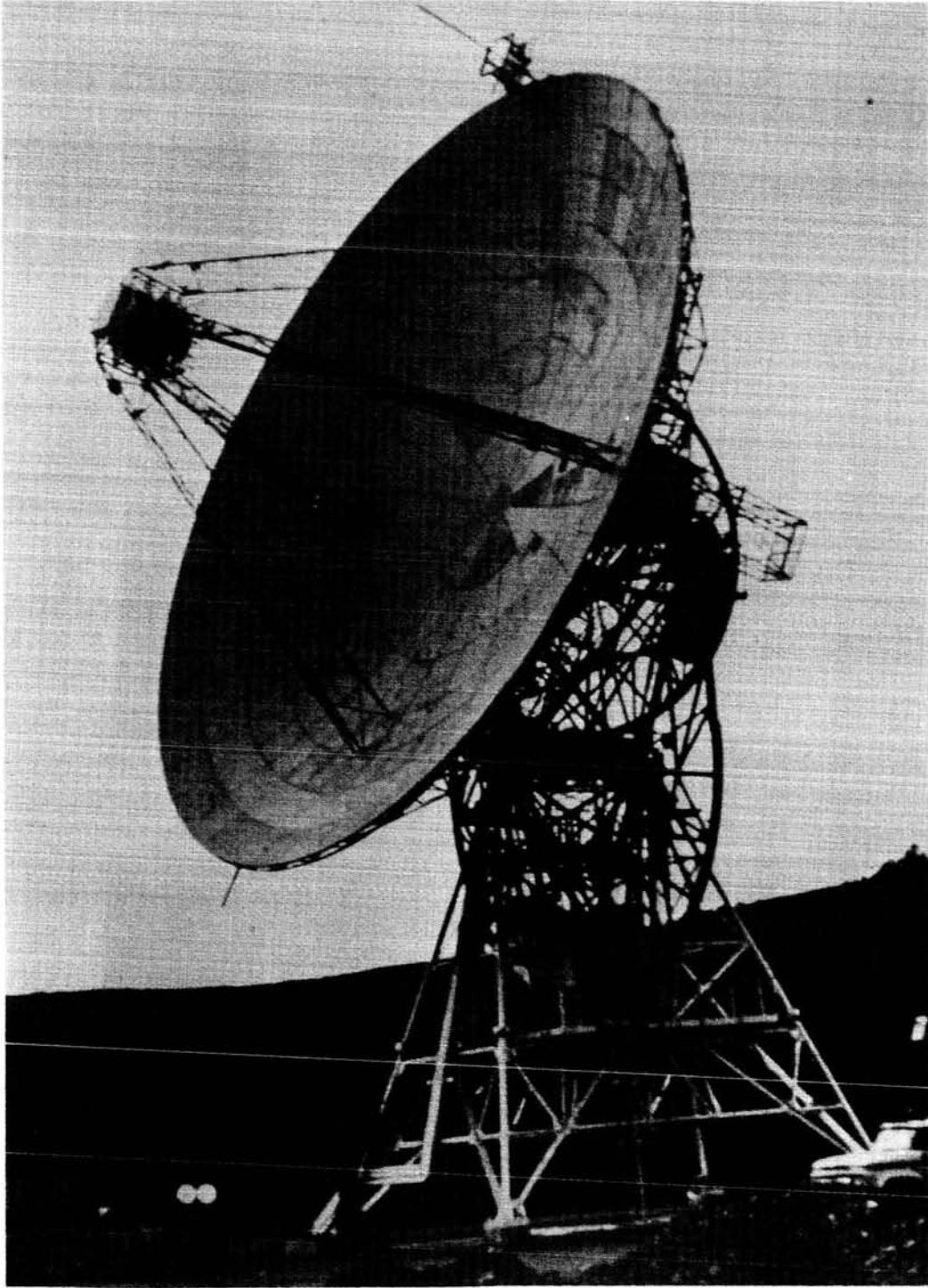


FIGURE 17.—85-foot paraboloid reflector antenna, Fairbanks, Alaska No. 1.

typical of any two-axis antenna mount. In some cases this gimbal-lock zone determines the lowest limit of tracking motion; in other cases, the interference of the reflector with the bearing yokes and

other physical perturbances is the limiting factor; in still other cases, the geographic horizon is the limiting factor. Figures 18, 19, 20, and 21 are coverage-limit plots for four 85-foot paraboloids.

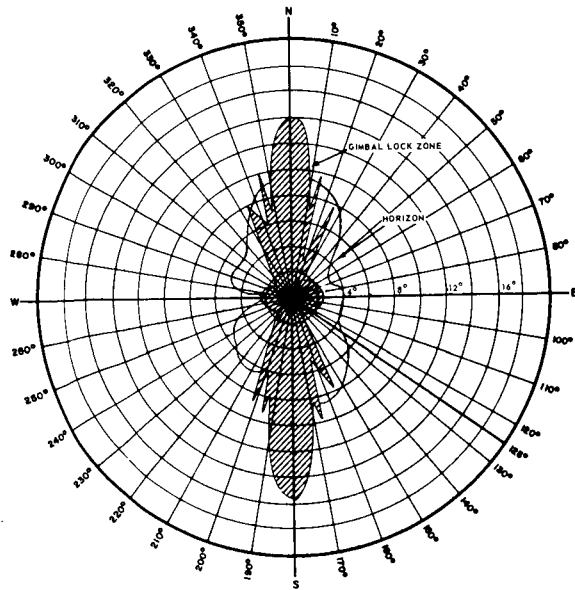


FIGURE 18.—Effective horizon profile (85-foot paraboloid, Fairbanks, Alaska No. 2).

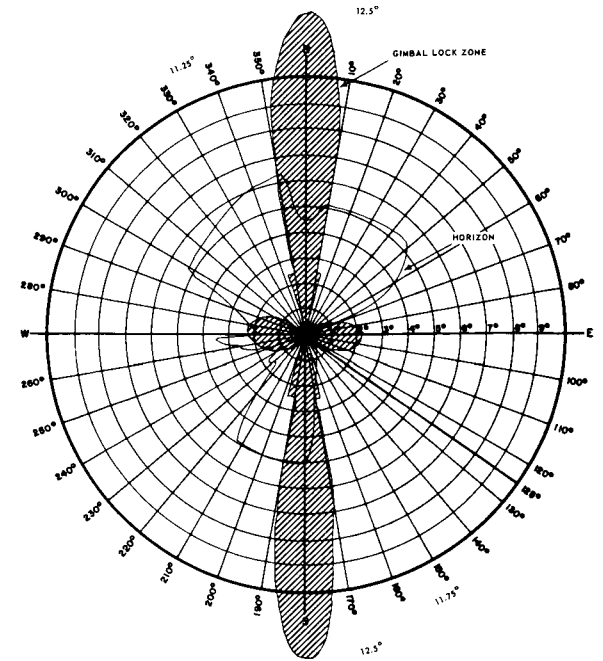


FIGURE 20.—Effective horizon profile (85-foot paraboloid, Rosman, N. C. No. 1).

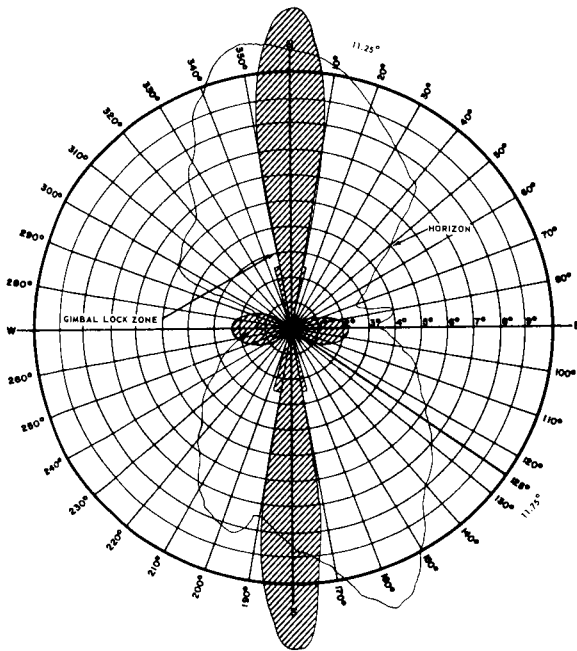


FIGURE 19.—Effective horizon profile (85-foot paraboloid, Fairbanks, Alaska No. 2).

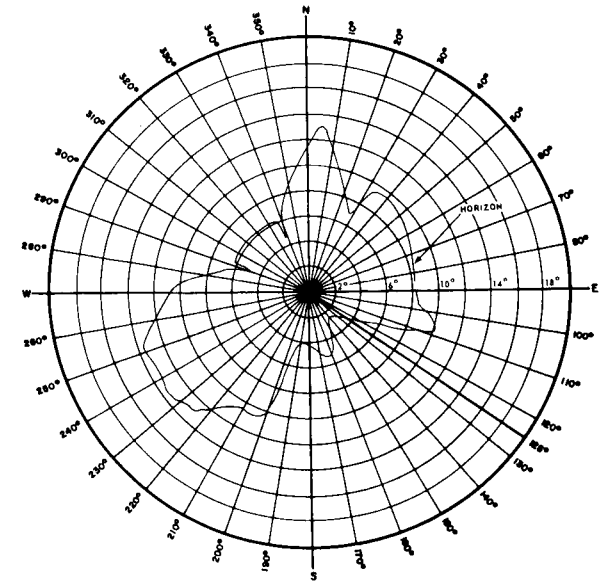


FIGURE 21.—Effective horizon profile (85-foot paraboloid, Canberra, Australia).

The entire antenna weighs about 300 tons and stands about 120 feet high. The 85-foot-wide-band antennas are part of a network to provide tracking, data-acquisition, and communications support for future satellite programs such as Nimbus, Eccentric Orbiting Geophysical Observatory (EGO), Orbiting Astronomical Observatory (OAO), Polar Orbiting Geophysical Observatory (POGO), and Advanced Orbiting Solar Observatory (AOSO). Six of these antennas either are in existence, or are under contract, as shown in Table 4. Three additional antennas, similar to item 6 in the table, will be procured, and will contain feeds for transmitting 2105 Mc and receiving 2285 Mc, the Apollo range and range-rate frequencies. The design of all reflectors, with the exception of antenna 1, provides a 10-foot circular section at the vertex of the reflector, which is removable for retrofit with Cassegrain feeding.

A contract has been awarded to RANTEC Corporation to retrofit the Rosman II antennas with a 400- to 4200-Mc Cassegrain feed, for use with the Advanced Technological Satellite; the feed will be delivered in September 1964. All six antennas are fully instrumental for data-acquisition and monopulse tracking in the 136- to 137-Mc, 400- to 406-Mc, and 1700- to 1710-Mc bands. In addition, 148-Mc command transmit elements are attached to the edges of the reflectors. These antennas are described in a subsequent section of this report.

The antennas are capable of tracking (in either linear or either circular polarization) at rates from 0 to 3 degrees per second with accelerations up to 3 degrees per second per second. A study is now under way to determine the pointing accuracy of

these antennas. Preliminary data indicate that the accuracy lies between ± 90 and ± 120 seconds of arc.

Five operational drive modes are selectable:

- Phase-monopulse tracking
- Perforated paper drive tape
- Manual
- Slave

Initial acquisition search, superimposed on the first and second modes

The antenna data system provides for the measurement, digital encoding, and readout of shaft angles for feeding into the servosystem, the visual readout, and the paper tape punch. These positional data, with data-quality codes, are punched on recorder paper in teletype code every 10 seconds. A small computer, with the teletype receiver, obtains satellite orbital predictions from Goddard Space Flight Center, examines the received data for transmission errors, and generates 1-second drive-tape predictions by computation and interpolation.

The support for the feed system is a quadripod, the head of which is a hollow structural square designed to hold a feedbox 4×4×6 feet long. This box contains, the coaxial hybrid junctions required for the phase-monopulse operation, the coaxial switches required to provide polarization diversity, the parametric amplifiers, the frequency converters, and similar equipment. The 1700-Mc feed consists of four crossed dipoles on a modified Roberts⁶ balun, mounted on the end of the feedbox under a radome cap. These dipoles are spaced 0.47λ apart and 0.21λ above a flat ground screen. The 400-Mc feed also consists of four crossed dipoles mounted on Roberts baluns 0.25λ

TABLE 4.—85-Foot Paraboloid Specifications

Location	Contractor	Surface Tolerance (in.)	Feed Deflection* (in.)	Axis Alignment (sec)
1. Fairbanks I.....	Blaw-Knox	0.062	0.125	6
2. Fairbanks II.....	Rohr	0.031	0.062	3
3. Rosman I.....	Rohr	0.031	0.062	3
4. Canberra.....	Rohr	0.031	0.062	3
5. (Unassigned).....	Rohr	0.031	0.062	3
6. Rosman II.....	Blaw-Knox	0.040	0.090	5

*Axial and lateral.

above a ground screen on the end of the feedbox, and spaced 0.56λ apart. The 136-Mc feed, which is similarly constructed, is mounted 0.25λ above a ground screen at an interelement spacing of 0.58λ .

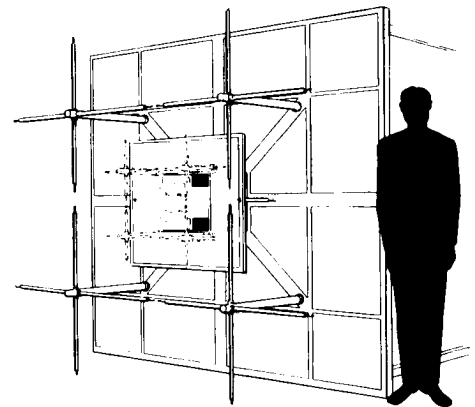
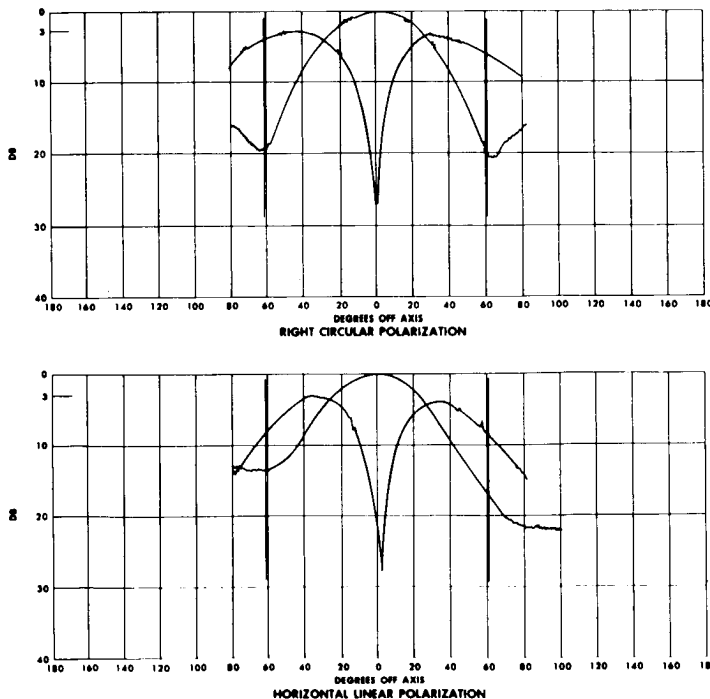
Boresighting data can be recorded on film by means of motion-picture cameras photographing television monitors at all sites and by an additional motion-picture camera mounted in the optical cage on the antenna at Fairbanks I.

Pattern Performance

Primary patterns of the three feeds, using a scale (0.0637) model at 136 Mc and full-scale models at 400 and 1700 Mc, were measured for both senses of circular polarization and for both linear polarizations. Table 5 summarizes characteristics for two representative polarizations. Figures 22, 23, and 24 show representative primary patterns for each frequency.

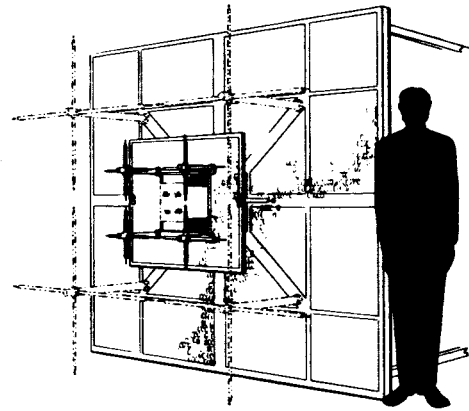
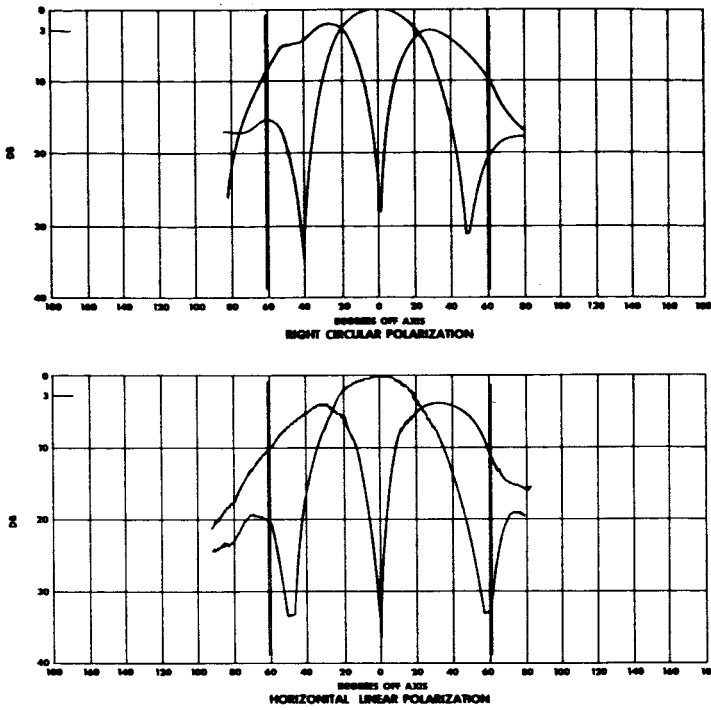
TABLE 5.—Primary Pattern Characteristics, Fairbanks I Paraboloid E-Plane, Referred to Field-Generating Test Dipole (Angle Subtended by Reflector 61.10°)

Frequency (Mc)	Polarization	Edge-Illumination Sum Pattern (db)	Null-Position Sum Pattern ($^\circ$)	Difference Slope (v ratio)
136-----	RC	20.0	62	1.4
136-----	HL	20.0	00	1.4
400-----	RC	19.0	40	1.4
400-----	HL	30.0	50	2.0
1700-----	RC	15.5	00	1.2
1700-----	HL	18.5	00	1.4



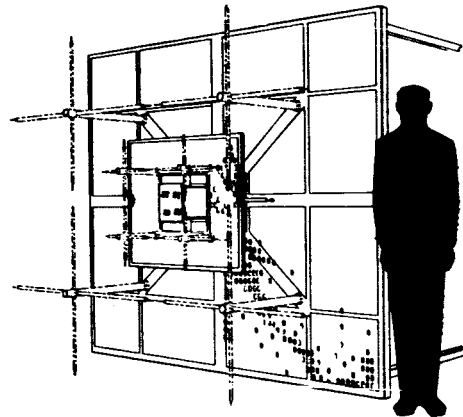
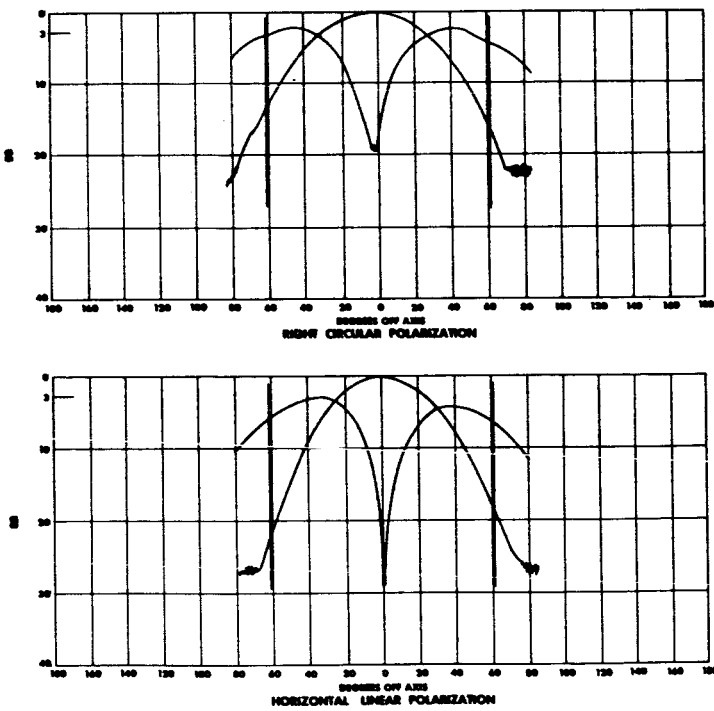
NOTE: E-PLANE FOR CIRCULARLY-POLARIZED PATTERN IS RELATIVE TO FIELD GENERATING TEST DIPOLE

FIGURE 22.—136-Mc primary sum and difference patterns (85-foot paraboloid, Fairbanks, Alaska No. 1).



NOTE: E-PLANE FOR CIRCULARLY-POLARIZED PATTERN IS RELATIVE TO FIELD GENERATING TEST DIPOLE

FIGURE 23.—400-Mc primary sum and difference patterns (85-foot paraboloid, Fairbanks, Alaska No. 1).



NOTE: E-PLANE FOR CIRCULARLY-POLARIZED PATTERN IS RELATIVE TO FIELD GENERATING TEST DIPOLE

FIGURE 24.—1700-Mc primary sum and difference patterns (85-foot paraboloid, Fairbanks, Alaska No. 1).

The nulls fell well inside the reflector at 400 Mc, placing the sidelobes on the reflector. The difference-pattern null slope appears to change with polarization, and is quite good at 400 Mc for horizontal polarization. The gain of the difference pattern relative to the sum pattern was down from 3.0 to 4.0 db at 136 Mc, from 2.0 to 3.8 db at 400 Mc, and from 2.9 to 3.3 db at 1700 Mc.

After installation of the first antenna near Fairbanks at Gilmore Creek, Alaska, secondary patterns were measured by generating: (a) a 1700-Mc and a 400-Mc RF field at a tower 18,662 feet from and 12.6 degrees north of the antenna, and, (b) a 136-Mc RF field at a tower 2095 feet from and 50 degrees north of the antenna. Failure to align a tower carrying an X-Y mount with the axis of rotation will result in a tilt (with respect to the ground) of the plane in which any RF pattern of radiation from the paraboloid is measured. This tilt of the plane can be corrected to be parallel with the earth's surface by moving the antenna about each axis simultaneously; however, this correction causes the polarization to roll. It was decided that the plane's tilting with respect to the ground was the lesser of these two evils. Although the ground effects vary

somewhat from one limit of pattern excursion to the other, this is a minor effect, and the measured patterns are considered to be reasonable representations of those which would be experienced when working toward a zenith-oriented satellite. Figures 25, 26, and 27 are plots of the secondary patterns at these three frequencies for right-circular polarization. Measured pattern performance about the Y-axis is summarized in Table 6. X-axis performance should be similar when directed near zenith; but, when operating with the calibration-tower RF source, multipath reflections from the ground distort the pattern.

TABLE 6.—Secondary Pattern Characteristics—Fairbanks I Paraboloid Right-Circular Polarization—Y-Axis

Frequency (Mc)	Beamwidth (°)		Sum-Pattern Sidelobe (db)
	Theoretical	Measured	
136.....	6.8	6.7	18.5
400.....	2.5	2.4	14.8
1700.....	0.5	0.5	18.0

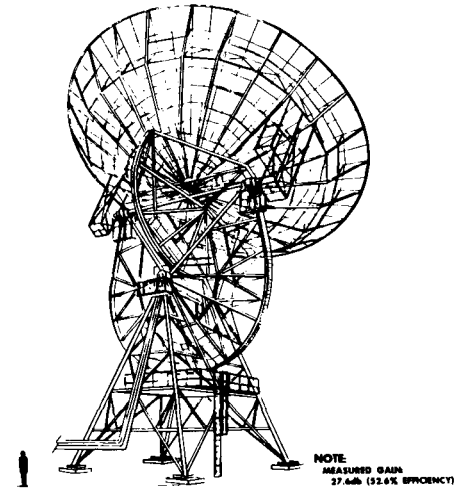
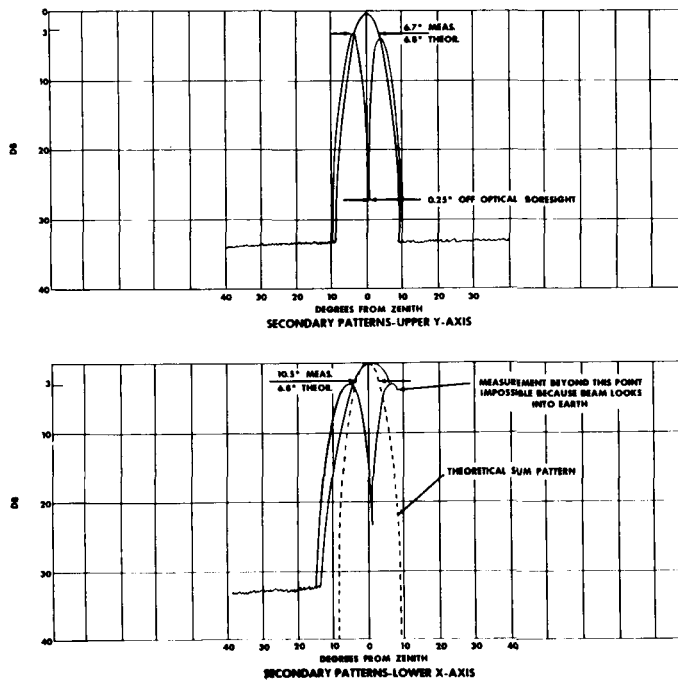


FIGURE 25.—136-Mc right-circular polarized sum and difference patterns (85-foot paraboloid, Fairbanks, Alaska No. 1).

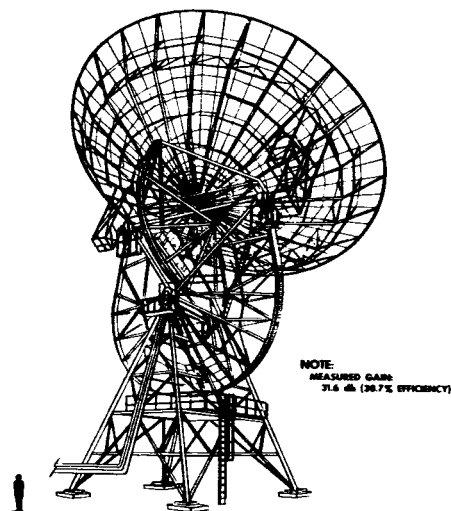
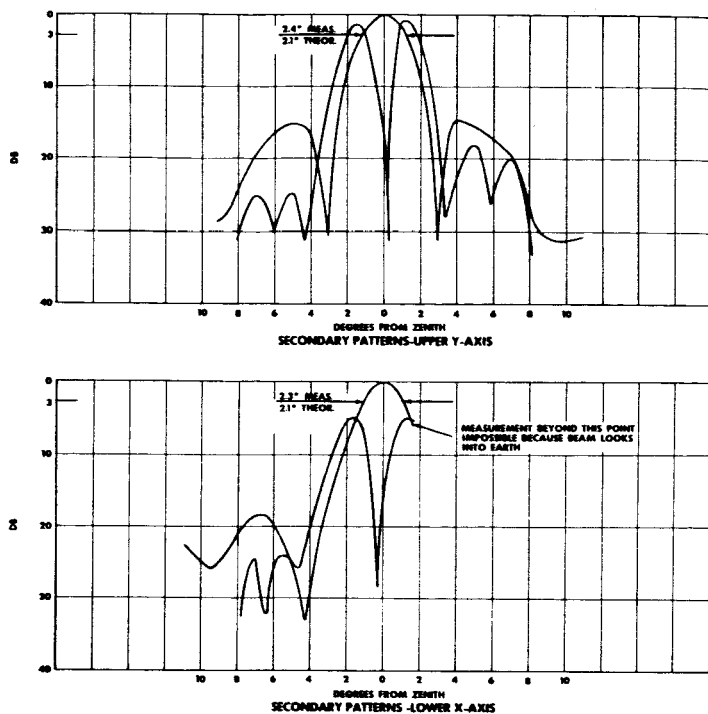


FIGURE 26.—400-Mc right-circular polarized sum and difference patterns (85-foot paraboloid, Fairbanks, Alaska No. 1).

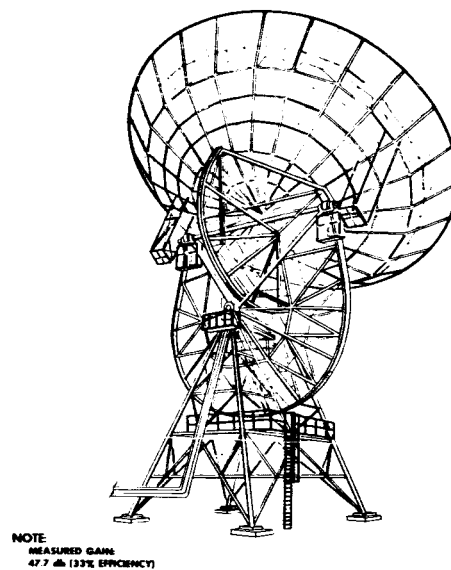
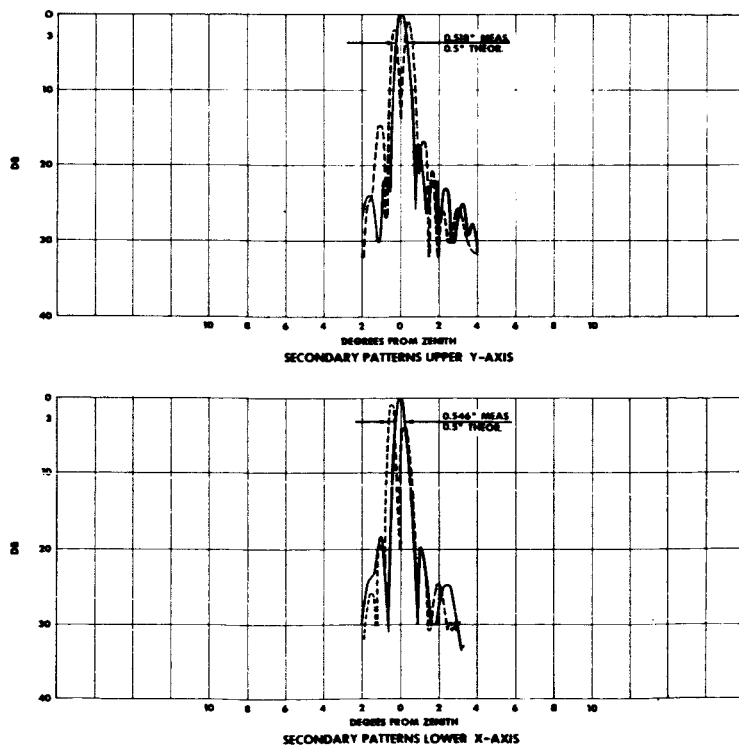


FIGURE 27.—1700-Mc right-circular polarized sum and difference patterns (85-foot paraboloid, Fairbanks, Alaska No. 1).

Measured pattern performance of the Rosman I antenna was similar to that indicated in Table 6, except that the sidelobe level of the Rosman antenna at 136 Mc was 8.0 db higher.

Gain

The gain of the Fairbanks I paraboloid was measured by comparison with calibrated standards. These standards were a Litton Model 13E 1-10 log-periodic dipole structure and an Andrew Type D-6 6-foot paraboloid. The theoretical gain for 100 percent efficiency can be computed from

$$\frac{4\pi A}{\lambda^2}, \quad (3)$$

and when this value is compared with the measured gain, the antenna efficiency is determined. Table 7 contains these calculations.

TABLE 7.—Measured Antenna Gain Fairbanks I, 85-Foot Paraboloid

Frequency (Mc)	Gain (db)	Efficiency (percent)
136.....	27.6	52.6
400.....	35.6	38.7
1700.....	47.75	33.0

Because of the equal primary illumination for all polarizations, gain did not change with polarization at 136 Mc. Gain change with polarization at 400 and 1700 Mc was slight. Although the poor efficiency at 1700 is unexplained, the presence of primary pattern sidelobes on the reflector at 400 Mc is considered to contribute to efficiency loss at that frequency.

Command Elements for 85-Foot Paraboloids Radiation Systems, Incorporated, Antenna

Radiation Systems, Incorporated
Alexandria, Virginia
NAS5-2462

General Description

The command-transmitting antenna (Figure 28) for use with 85-foot paraboloids is a disc-on-rod structure attached to the outer rim of the reflector. It operates over the 120- to 155-Mc frequency band and is capable of supporting 3-kw

average power. This disc-on-rod assembly consists of 10 discs, 29 inches in diameter, spaced 14.5 inches apart on a 4.5-inch-diameter shaft. A crossed dipole (approximately $\frac{1}{2}$ wave at 148 Mc) is located 28 inches above 123-inch-long reflector rods and launches the RF wave on the structure. A round basket, 53 inches in diameter and 36 inches deep, surrounds the crossed-dipole launcher and equalizes the E- and H-plane beamwidths. A switchbox mounted directly under the reflector rods provides capability for transmitting linear or circular polarization. Impedance matching is accomplished by reducing the spacing between the first disc and the dipole launcher from 14.5 to 9 inches. In addition, it was found that the radiation-pattern form factor could be improved by reducing the diameter of the first three and the last three discs from 29 to 27 inches.

The antenna is constructed of aluminum alloy. The total weight of the antenna, the polarization-switching equipment, and the attachment tube is 221 pounds.

Pattern Performance

The disc on rod radiates a beam symmetric along the longitudinal axis of the structure. E- and H-plane patterns (Figure 29) were measured over the 120- to 155-Mc band. The E- and H-plane symmetry is well preserved over the entire band. Sidelobes remain at approximately 13.0 db over the 120- and 145-Mc band and drop sharply to 18.0 db from 145 to 155 Mc. Beamwidths at frequencies throughout the band are shown in Table 8.

TABLE 8.—Gain and Beamwidth RSI Command Transmit Disc on Rod

Frequency (Mc)	Beamwidths (°)		Gain (db)
	E-Plane	H-Plane	
120.....	40	42	11.6
125.....	41	42	12.2
130.....	36	42	12.6
135.....			13.8
145.....	33	35	13.5
150.....	31	32	13.7
155.....	29	31	12.4

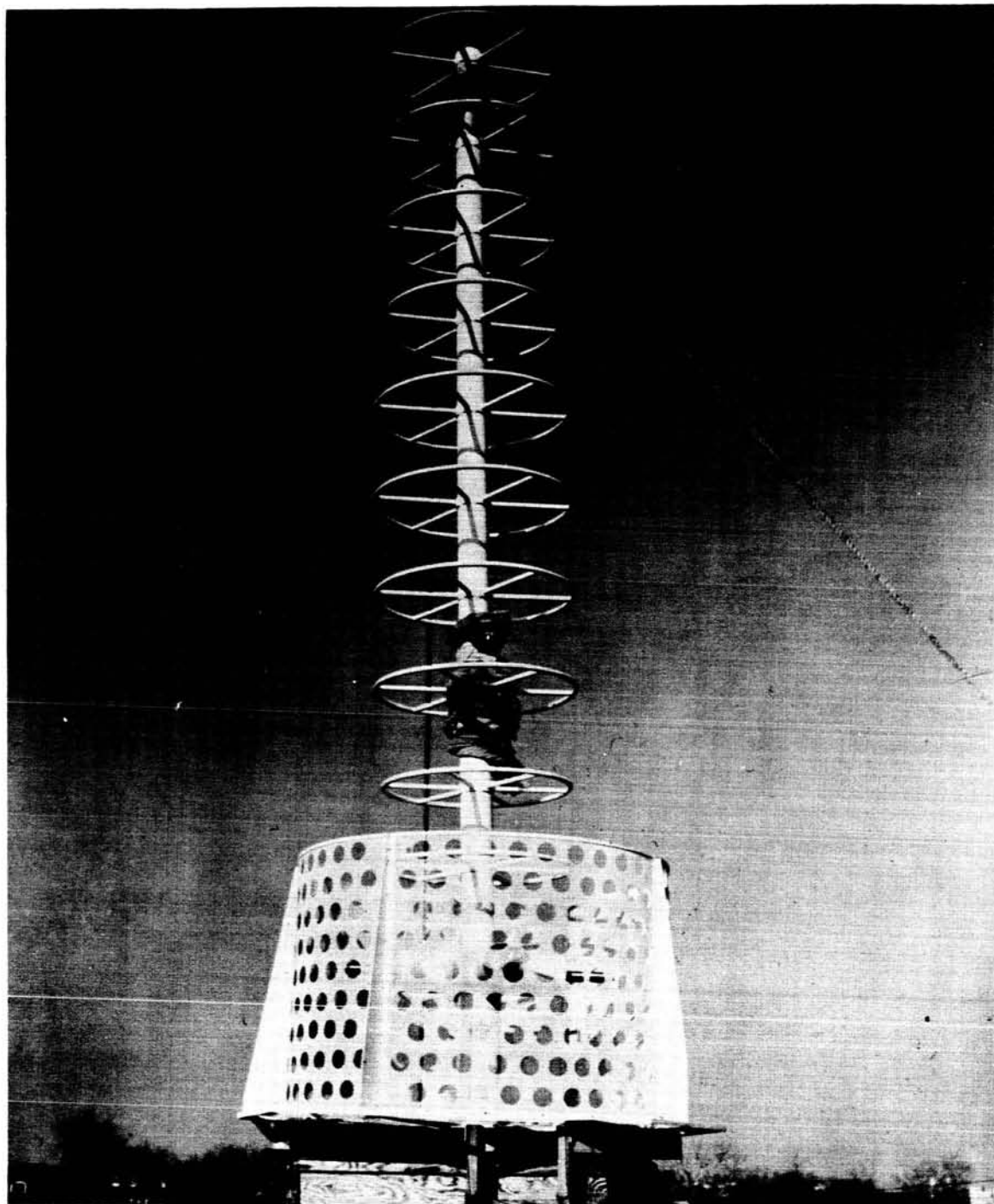


FIGURE 28.—Command element for 85-foot paraboloids.

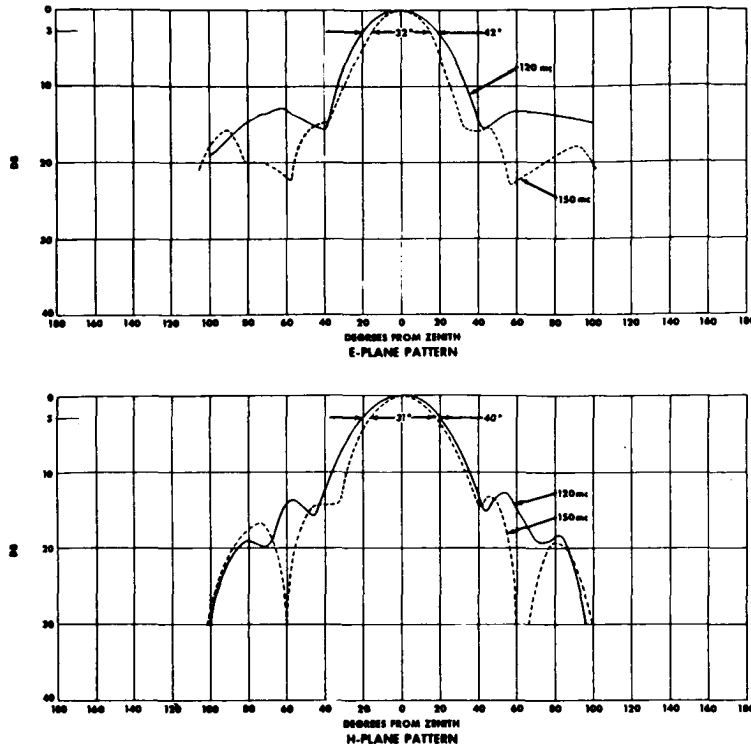


FIGURE 29.—Patterns-command element for 85-foot paraboloids.

Gain

Gain (above an isotropic source) measured by the similar-antenna substitution method is indicated in Table 8, which gives the average of measured parameters for three antennas. It can be seen that gain performance at the 148-Mc operational frequency is about 13.6 db.

Input Impedance and VSWR

The nominal input impedance of the disc-on-rod assembly is 50 ohms. The VSWR was measured and found to be 1.23 at 148 Mc; this value is typical of performance at other frequencies. At the low end of the band, VSWR improves to 1.18; at the high end of the band it is as high as 1.84.

Avien, Incorporated, Antenna

The command element used on the Fairbanks I antenna was provided by Avien, Inc., and like the RSI antenna, contains 10 discs. The driver consists of four $\frac{1}{2}$ -wave dipoles arranged in a square pattern above a square expanded-aluminum

ground screen. This antenna, designated as Avien Model 300A, performs similarly to the Radiation Systems antenna. Gain at 148 Mc is better than 12.0 db, the beamwidth is 39.5 degrees (wider than the RSI antenna), sidelobes are 12.0 db, and VSWR is 1.20. The antenna is constructed of aluminum alloy and weighs 218 pounds.

225-Mc Auxiliary Feed for 85-Foot Reflectors

Technical Appliance Corporation
 Sherburne, New York
 Purchase Order S-32965-3

General Description

To equip the two 85-foot paraboloidal reflectors at Fairbanks, Alaska, to function as a TIROS primary command and data-acquisition facility, four 235-Mc dipoles were procured and installed at the focus of each reflector, as shown in Figure 30. These dipoles are Technical Appliance Corporation Model D-1489 gamma-matched half-wave dipoles, and are oriented 45 degrees with

NOTE:
 MEASURED GAIN:
 120 mc - 11.6db
 150 - 13.7
 RELATIVE TO AN ISOTROPE

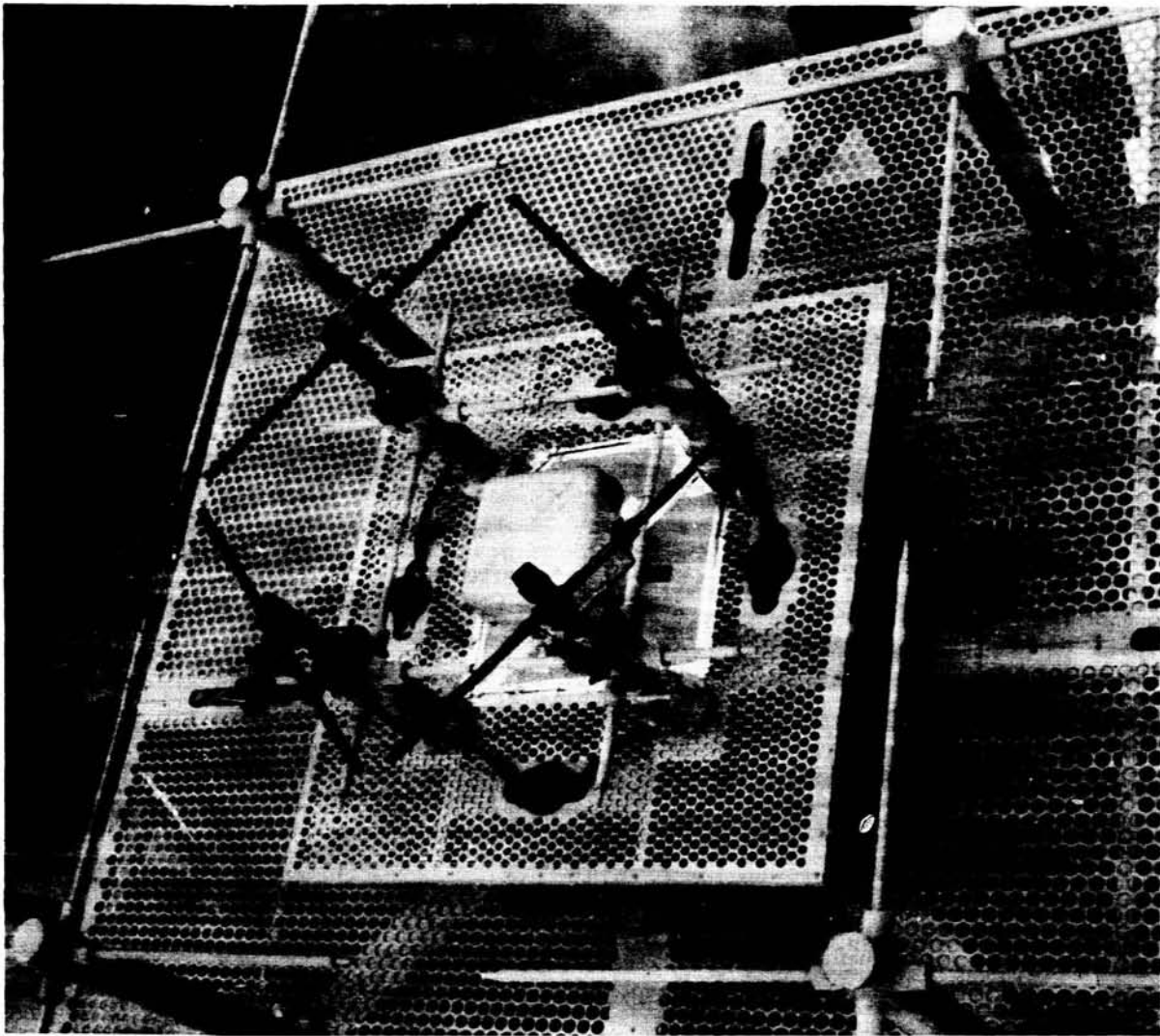


FIGURE 30.—235-Mc feed for 85-foot paraboloid, Fairbanks, Alaska.

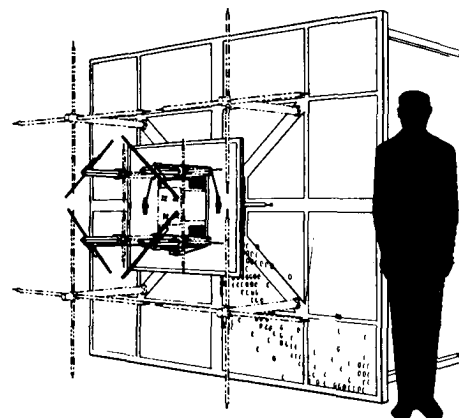
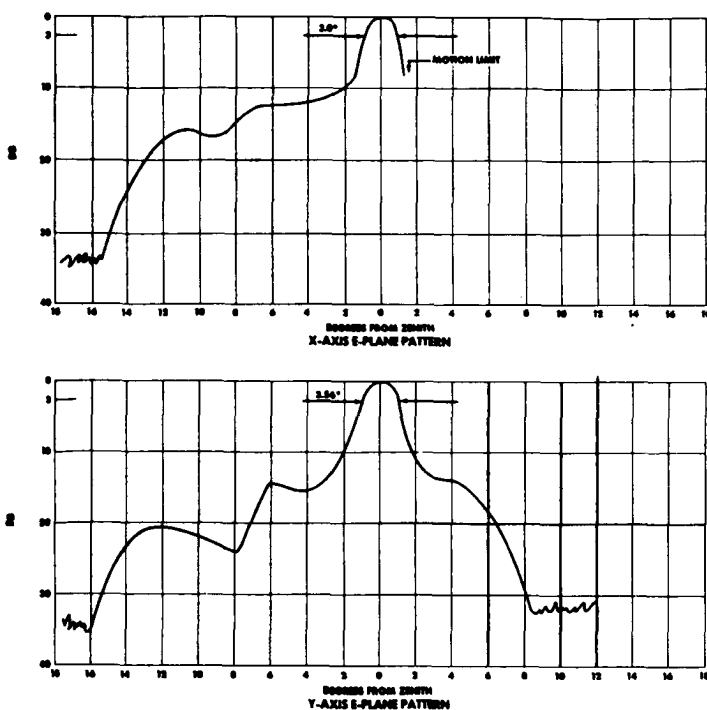
respect to the 400-Mc feed onto which they are mounted. This results in a broadside spacing of 26.7 inches (0.511λ at 235 Mc) and a spacing of 0.4λ above the 400-Mc ground screen. A coaxial hybrid junction, constructed of RG-59/U, together with remotely controlled coaxial switches, provide capability for polarization selection of this listening feed. Tracking of the satellite is accomplished by means of the 136-Mc monopulse feed in the paraboloid which tracks on the satellite beacon.

Pattern Performance

The pattern performance of the 235-Mc feed, as well as the resultant degradation of pattern

performance of extant feeds, was measured at Fairbanks after installation. Figure 31 is a plot of the X- and Y-axis patterns for the right-circularly polarized feed arrangement. It can be seen that the beamwidth is slightly wider than 3 degrees and that 12.0 db shoulders occur as a result of improper illumination of the reflector. The pattern is usable for the intended purpose.

The pattern performance at 136 and 400 Mc was essentially unaffected by the addition of the 325-Mc feed except for a slight broadening of the sun pattern beamwidth. Difference pattern null depths and sidelobe levels were unchanged. At 1700 Mc the same slight broadening of the sum-



NOTE
GAIN ABOVE AN ISOTROPE—APPROXIMATELY 31 db

FIGURE 31.—Patterns of 235-Mc feed for 85-foot paraboloid, Fairbanks, Alaska.

pattern beamwidth occurred and, in addition, the null depth was 8.0 db shallower and sidelobes were 3.0 db higher in the difference pattern.

Gain

The gain of the dipole-fed reflector at 235 Mc was not measured, but was computed to be approximately 31.0 db for the rather inefficient illumination.

VSWR

The VSWR, looking into the input to the hybrid junction, was measured with the feed installed in the reflector. It was observed that the VSWR was 1.1 at 235 megacycles.

III. ANTENNA ARRAYS

Nine-Yagi Array

All Products Company
Mineral Wells, Texas
NAS5-492

General Description

These antennas (Figure 32) are located at all STADAN Minitrack facilities and consist of an

array of nine, seven-element crossed yagi antennas spaced 0.77λ and mounted atop a 13.5-foot steel tower. Eight of the yagis are used in the acquisition of data and the centrally located yagi is used to transmit command signals. Azimuth-elevation rotators at the top of the tower, drive the antenna ± 80 degrees from zenith, and ± 360 degrees in azimuth. These rotators are composed of one-half horsepower, reversible, brushless motors, and Winsmith Model 9CVD RU-LU speed reducers which provide 10,000 inch-pounds of torque and move the antenna at a speed of 1 rpm about either axis. Antenna motion is controlled by remote toggle switches, and remote antenna position readout is given to an accuracy of ± 2 degrees.

Electric Design

The receiving portion of the antenna consists of nine crossed yagis corporately arrayed as indicated in Figure 33 and uniformly fed without the use of baluns. Branch point impedances are indicated on the figure as well as the characteristic impedances of the lines. Line impedance was compromised in an effort to economize and simplify the design. For example, at the input

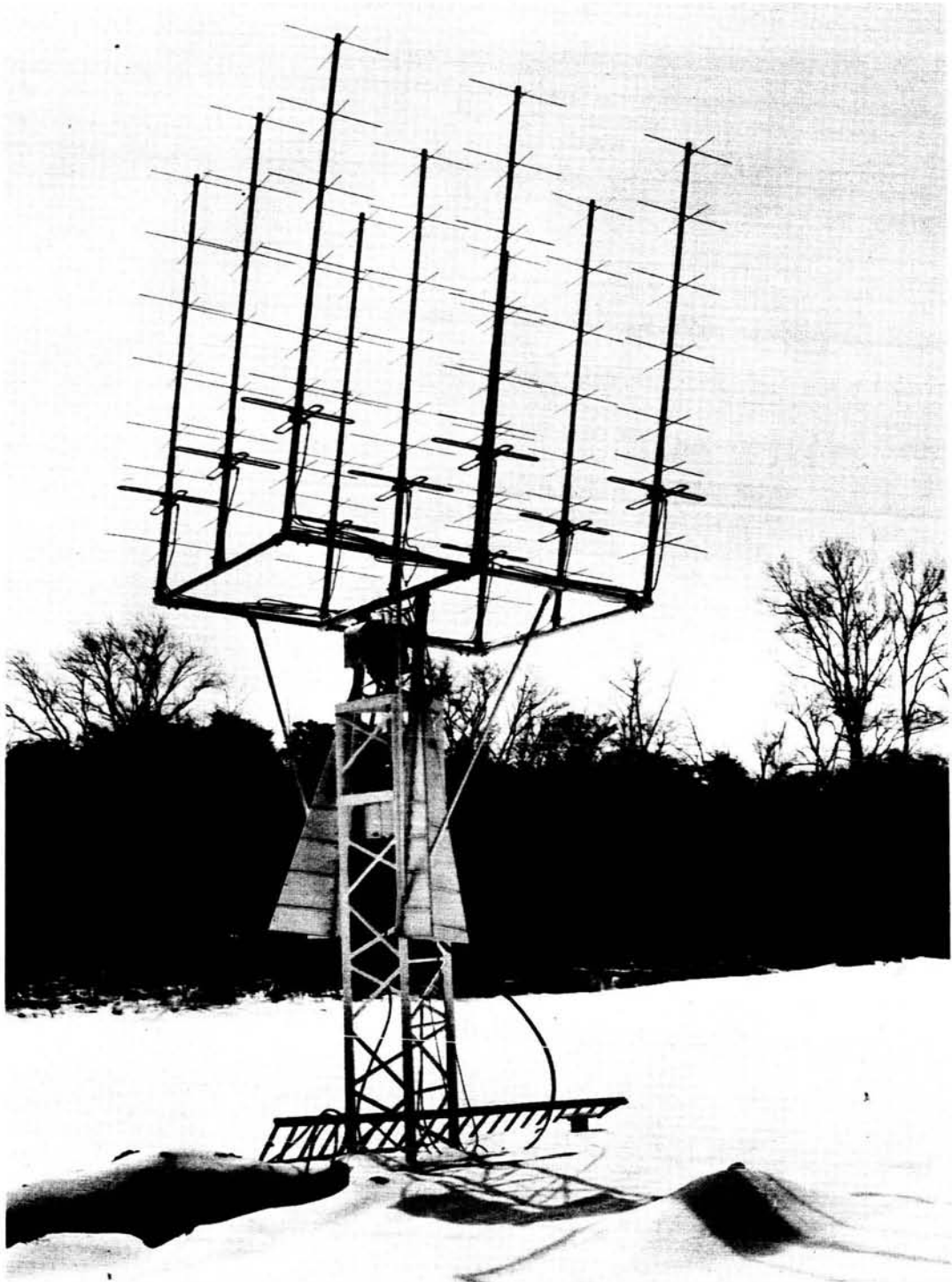
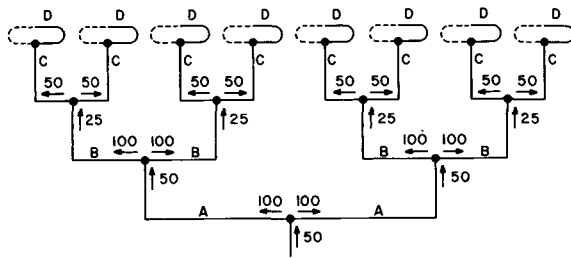


FIGURE 32.—136-Mc switchable nine-yagi array.



Impedance Transforming Lines

- A— 0.75λ Length of 75-ohm Foamflex
- B— 0.75λ Length of 50-ohm Foamflex
- C— 2.07λ Length of 50-ohm Foamflex
- D— 0.75λ Length of 95-ohm RG-133/U Inside Dipole

FIGURE 33.—Feed system for nine-yagi array.

branch point the impedance is 100 ohms, and a quarter wavelength of 75-ohm line is required to match back to 50 ohms; however, the line (line A) continues as a 75-ohm line (instead of a 50-ohm line) beyond the transformer. Consequently, a relatively high standing wave occurs on this line. The same compromise is made on lines B and D. The feed-point impedance of the folded dipole is 180 ohms and the match to a 50-ohm line is accomplished by inserting a length of 95-ohm RG/133/U cable inside the tubing. Half-inch diameter Foamflex is used in constructing the

feed system shown in Figure 33 and a separate feed system is used for each linear polarization. Circular polarization is accomplished by joining the two with an unbalanced coaxial "Y" which introduces a 90-degree phase delay in one arm. The direction of circularity is a function of orientation of the "Y" section. A line stretcher is provided to correct line-length errors and thereby improve circularity. Although the circularity was never measured, it is assumed that the lines have been cut to a sufficient accuracy to yield good performance.

Each individual yagi consists of a half-wave folded driver 0.22λ above a reflector element cut in length to 0.51λ . Six directors spaced 0.30λ apart and cut in length to 0.43λ collimate the energy. To improve the form factor, the length of the first director (adjacent to the dipole) is reduced to 0.42λ ; to improve the impedance match, the spacing between the first two directors is only 0.10λ . Aluminum construction, throughout, results in a 30-pound total weight for the 18-foot-long yagi.

Pattern Performance

The radiation patterns for linear polarization were measured at the factory and are plotted in Figure 34. Figure 34 shows that the E- and H-plane beamwidth is 19.5 degrees and that the sidelobes do not exceed 12 db.

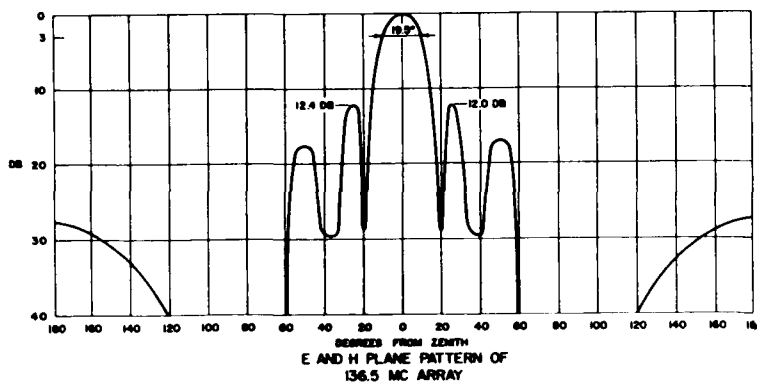
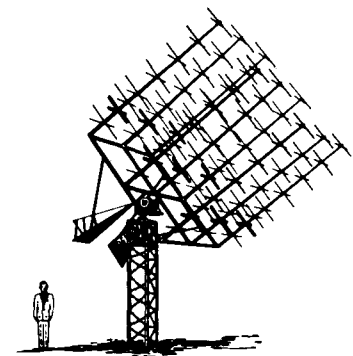


FIGURE 34.—E- and H-plane linearly polarized patterns, 136-Mc nine-yagi array.

GAIN (MEASURED)
19.2 DB ABOVE AN ISOTROPIC SOURCE

Gain

The gain of the prototype model was measured at the All Products Company, by comparison with a standard half-wave dipole, and found to be 19.2 db. Reference to the gain-beamwidth relationship (expression 1) indicates this measured value to represent a 76.6 percent aperture efficiency.

Input Impedance

The input impedance of the array was measured at the factory and is plotted in Figure 35. It can be seen that the antenna is matched at

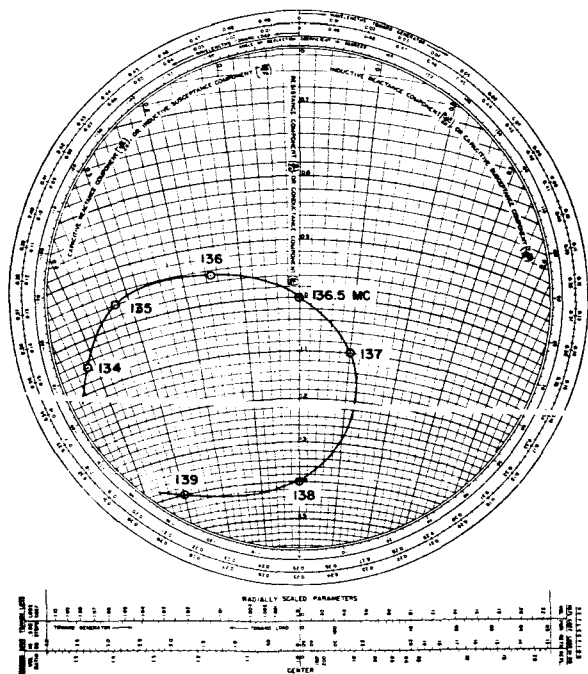


FIGURE 35.—Input impedance, 136-Mc 16-yagi array.

136.5 Mc and meets a design specification that the VSWR should not exceed 2.0 (6 db) over the 135- to 139-Mc band.

Sixteen-Yagi Array

All Products Company
 Mineral Wells, Texas
 NAS-5-492 (amendment 2)

General Description

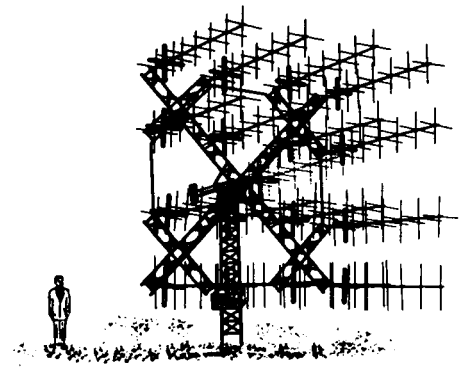
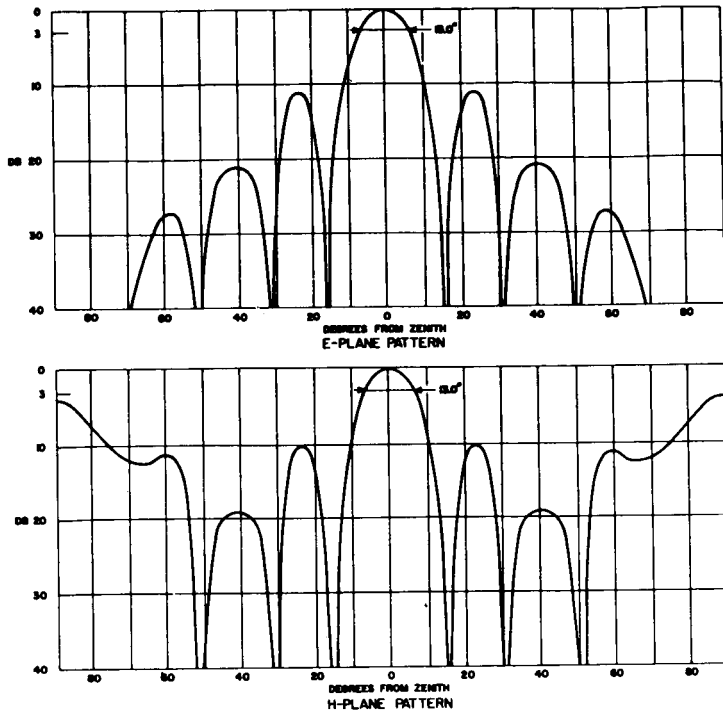
To provide a high-gain antenna for acquiring data from Explorer XII, the yagi elements were removed from surplus 108-Mc 16-yagi arrays, and 136-Mc yagi elements were attached at the same locations. A new Foamflex feed harness was cut to the proper dimensions to provide equal power in phase to each yagi. The resultant inter-yagi spacing (0.97λ) is not the best arrangement, as it causes some gain reduction and an increase in sidelobe level. In addition, the 136-Mc yagi designed for use at 0.77λ spacing in the 9-yagi array was used without modification. Although the first director adjacent to the dipole driver should have been relocated for 0.97λ spacing to improve the yagi form factor, time did not permit this. As a result, the individual yagi elements have a high H-plane sidelobe, which appears off-axis in the theoretical array pattern (Figure 30).

Pattern Performance

The pattern performance was measured by the firm of Silliman, Moffet, and Rohrer at Santiago, Chile, and by Goddard personnel at Woomera, Australia. It was found that performance deviated significantly from theoretical and was not uniform at both locations. Measured findings are shown in Table 9.

TABLE 9.—Measured Pattern Performance, 136-Mc 16-Yagi Array

Location	Beamwidth (°)		Location and Height of Far-Out H-Plane Sidelobe	
	E-Plane	H-Plane	(°)	(db)
Theoretical	13.0	13.0	90	6.0
Santiago	10.0	11.4	60	6.4
Woomera	12.0	14.8	60	14.0



GAIN: 22.4 DB ABOVE AN ISOTROPIC SOURCE

FIGURE 36.—E- and H-plane patterns, 136-Mc 16-yagi array.

Gain Performance

The theoretical gain performance based on 76.6 percent aperture efficiency is 22.4 db. Gain was measured at all antenna installations by comparing a half-wave dipole with a calibrated log-periodic standard. Some variation was observed: At Woomera the gain was observed to be 22.3 db, at Johannesburg 21.2 db, and at Santiago 19.0 db.

Impedance and VSWR

The input impedance was measured at Santiago by Silliman, Moffet, and Rohrer who matched the antenna with a two-stub tuner. The measured data is given in Figure 37, where the antenna is matched to a VSWR of 1.1 over the 136.4- to 137.0-Mc band.

Command Element for Yagi Arrays

All Products Company
 Mineral Wells, Texas
 NAS5-492

General Description

The command element for the 136-Mc yagi arrays is a crossed yagi, located at the center of

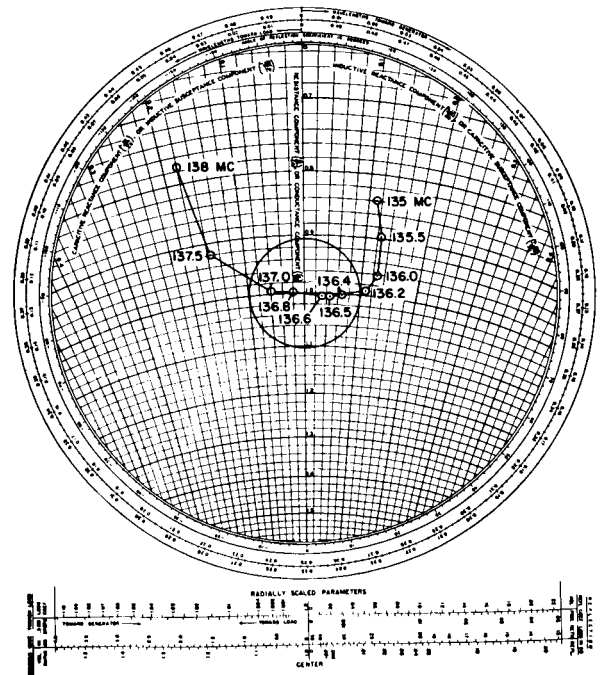


FIGURE 37.—Input impedance, 136-Mc nine-yagi array.

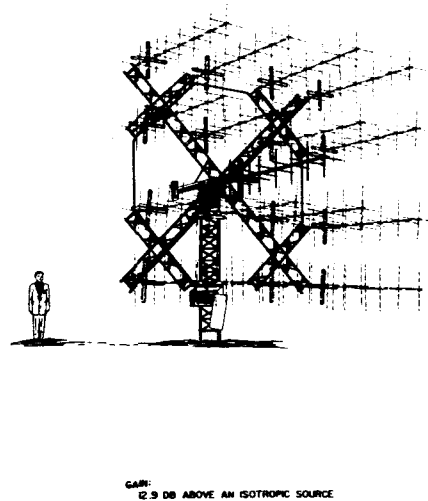
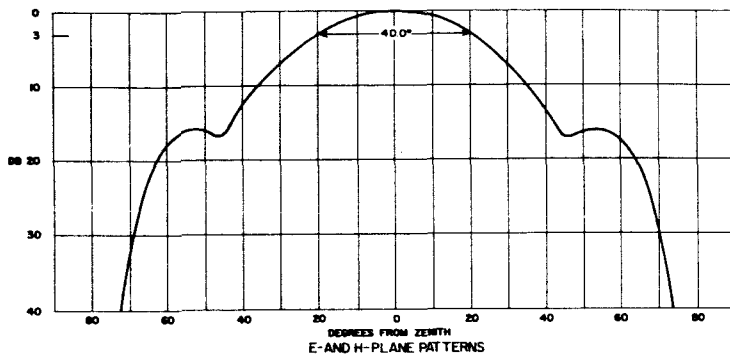


FIGURE 38.—Linearly polarized patterns for command antenna, 136-Mc yagi arrays.

the array, capable of transmitting 0.5-kw-average power. Figure 38 shows the command antenna in the center of a 16-yagi array in the same manner in which it is located in the 9-yagi array. Figure 38 also shows the beamwidth in the E- and H-planes to be 40 degrees and the gain to be 12.9 db, based on 76 percent aperture efficiency. The command yagi consists of seven elements over a reflector and is identical to the data-acquisition yagi except that it is scaled in frequency.

Satellite Automatic-Tracking Antenna (Satan)

Dalmo-Victor Company
Belmont, California
NAS5-3041

General Description

The SATAN antenna (Figure 39) is a 136- to 137-Mc phase-monopulse array of 16 yagi structures on a square frame attached to a hydraulically driven, tower-mounted X-Y pedestal. The yagi array was designed, developed, and constructed by Technical Appliance Corporation on a sub-contract from Dalmo-Victor Company. It consists of 16 yagis, 161 inches (1.91λ) long and spaced 86.6 inches (1.0λ) apart on centers. Each yagi is provided with a reflecting screen instead of the usual parasitic reflecting element. The dipole drivers are sleeve-type elements mounted on a pressurized, split, British baluns. Six parasitic

director elements are above each driver. The outputs of the 16 end-fire yagi elements are connected through a coaxial line network to provide polarization selection, polarization diversity, tracking outputs, and telemetry outputs. Figure 40 shows this network. All coaxial lines are $\frac{7}{8}$ -inch, 50-ohm Styroflex (Phelps-Dodge ST 78-50), and coaxial hybrid junctions are $\frac{1}{2}$ -inch, 50-ohm Styroflex (Phelps-Dodge ST12-50). The coaxial switches are Thompson-Ramo-Wooldridge, Inc., Model C 16T20B1, four-port transfer switches, with remote control indication. Figure 40 also shows the RF network.

The received signal enters a hybrid junction which sums the H-plane broadside pair of yagi elements in each subarray of four yagi elements. The two outputs from this subarray pass through coaxial line switches which change the polarization characteristics. The output from each subarray enters another hybrid junction, the sum output of which is the desired polarization and the difference output of which is the orthogonal polarization. The four sum signals pass to a monopulse bridge which provides a sum output and two orthogonal error-signal inputs. The sum signal is split into two components to provide both the monopulse reference signal for tracking and the telemetry signal. Each output contains a preselection filter and a preamplifier mounted on the array frame.

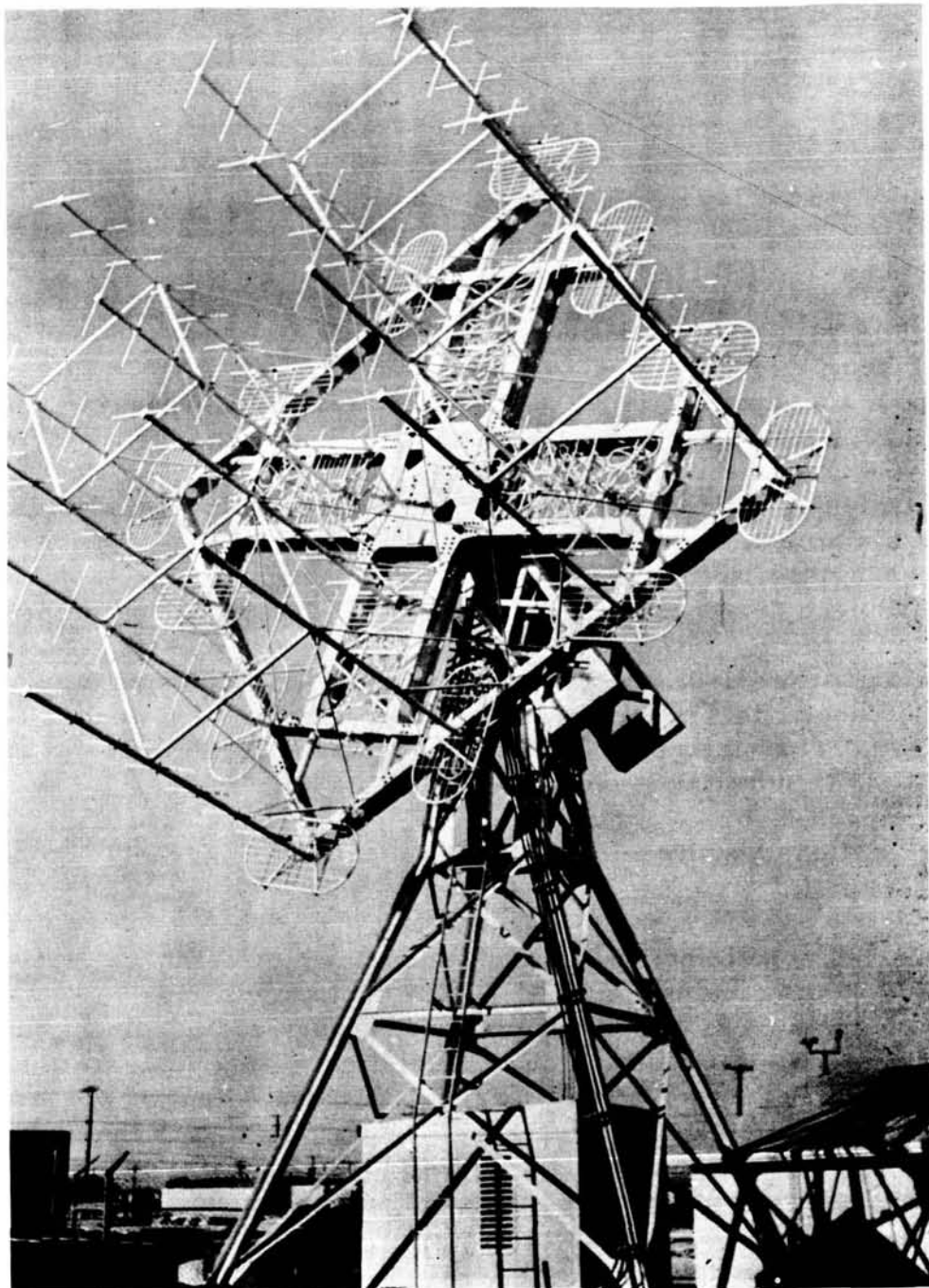


FIGURE 39.—Satellite automatic-tracking antenna (SATAN).

The specified tracking accuracy of the SATAN array is ± 0.10 degree dynamic boresight error. Because of this rigid requirement, washers are inserted in the RF line to correct line lengths in the feed system for phase alignment.

Pattern Performance

The pattern performance of the first three production models of the SATAN array was measured extensively. With the array mounted on an azimuth-over-elevation-over-azimuth pedestal,

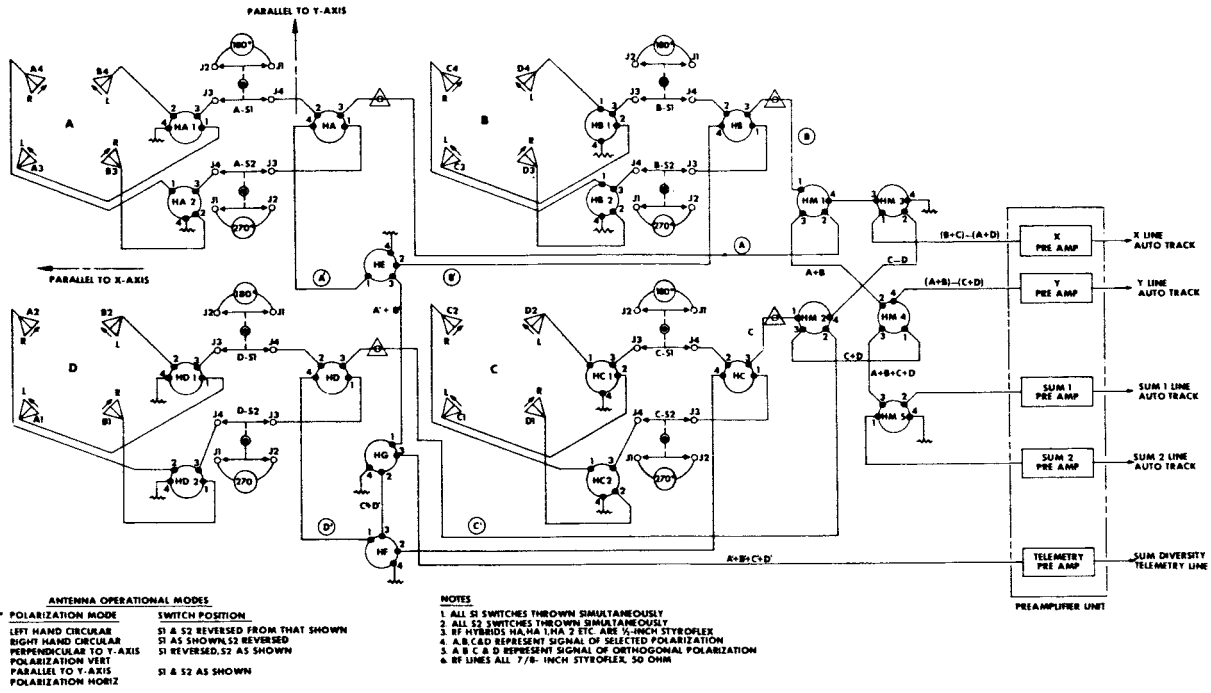


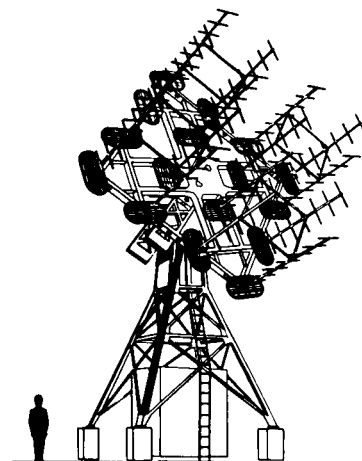
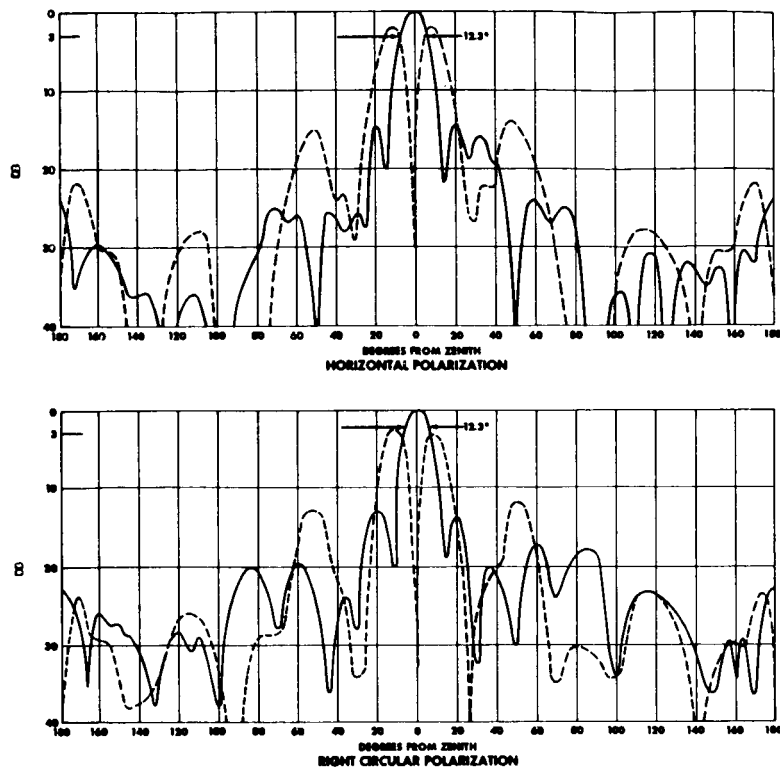
FIGURE 40.—Coaxial line schematic, SATAN antenna.

the array was rotated about the (effective) X-axis and the Y-axis as well as a simulated 45-degree cut. In this manner true major axis cuts were made through the patterns in a manner which could not be accomplished with the array on an X-Y pedestal. The pattern recorder was attached to the sum channel, X-difference channel, Y-difference channel, and sum-diversity channel of a coaxial terminal junction box. Patterns were measured for horizontal, vertical, and right- and left-circular polarization with both a horizontally and a vertically polarized field. Hence the cross-polarized signal component was measured. In this manner, 22 patterns were measured on each array at both 136.5 and 131.5 Mc. Representative patterns are shown in Figure 41 where the sum-pattern beamwidth can be seen to be 12.3 degrees. Specified beamwidth was less than 18.5 degrees; however, to meet the gain requirement, it was necessary to design for the 12.3-degree beam. Sidelobes were specified to be at least 11.0 db below the main beam and measured sidelobes were observed to be generally 12.0 db or more down.

The ellipticity of the pattern, for either sense of circular polarization, was specified to be less than 2.0 db and was measured to be no greater than 0.4 db. In addition, it was observed that the difference-pattern null depths were generally more than 28.9 db below the peak intensity of the sum-pattern main beam.

Boresight shift with polarization was meticulously measured because this monopulse antenna may be required to drive other antennas to great precision. Two characteristics were measured: (1) the extent of null shift with polarization change, and (2) the mean deviation of the electrical null from the optical boresight axis. The null shift was measured to be 0.18 degree about the X-axis and 0.20 degree about the Y-axis. The mean null deviation from the optical boresight axis was observed to be 0.25 degree for the X-axis and 0.19 degree for the Y-axis.

The cross-polarized component of the antenna pattern was measured and found to be well suppressed. For example, the cross-polarized component in the X-axis sum pattern was found to be suppressed 26.3 db for horizontal polarization and 26.1 db for vertical polarization.



NOTE:
GAIN: 22.3 ± 0.25 db ABOVE AN ISOTROPIC SOURCE

FIGURE 41.—136-Mc X-axis patterns, SATAN antenna.

Gain

To measure the gain of the SATAN array, a four-yagi SATAN subarray was calibrated using the identical antenna substitution method for use as a standard. A calibrated standard was mounted back-to-back on the array and gain was measured by rotating the mount and alternately reading the signal level. Table 10 shows these results where the specified 22.0-db gain is met at 136.5 Mc. The measured gain at 136.5 Mc is considered to be correct to 0.25 db. Note the efficiency of this antenna array. Application of

the approximate gain expression based on beamwidths

$$G = 0.76 \frac{41,253}{\theta_v \times \theta'_n}$$

where 41,253 is the number of square degrees in a hemisphere, θ is the beamwidth, and the constant 0.76 is the factor ordinarily applied to yagi arrays, yields a theoretical gain of 23.3 db. Comparison with the measured gain yields an 80.7 percent efficiency. However, the sidelobes of the SATAN array are rather high and the aperture distribution is uniform; hence, it can be expected that the efficiency is somewhat greater than this value. If the rectangular aperture area of the array is converted to the equivalent circular area of a paraboloid and the computed gain of this paraboloid is compared to the measured gain of the SATAN array, an 82.2 percent efficiency is obtained. This latter figure is probably the more valid one and it can be assumed that the SATAN array is highly efficient.

TABLE 10.—Measured Gain, SATAN Antenna

Frequency (Mc)	Gain (db)
131.5.....	21.9
136.5.....	22.24
141.5.....	22.6

Input Impedance and VSWR

The VSWR and impedance seen looking into the sum channels of the input hybrid mixer (HM-5 in Figure 40) were measured for all polarization combinations over the 10-Mc frequency band centered about 136.5 Mc. It was found that VSWR did not exceed 1.16 for any circumstance over the 136- to 137-Mc band. Figure 42 is a typical impedance plot. Impedance was measured looking into the sum channel for right-circular polarization. The excellent impedance match of the antenna can be seen in Figure 42.

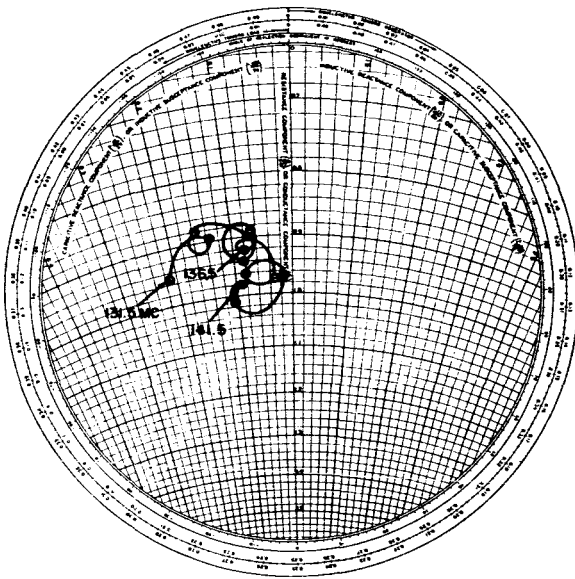


FIGURE 42.—SATAN array impedance.

Broadband High-Power Command Antenna

Radiation Systems, Incorporated
Alexandria, Virginia
NAS5-3142

General Description

This transmitting antenna (Figure 43) is an array of nine disc-on-rod elements designed for mounting on the SATAN pedestal, and for use with a 5-kilowatt CW transmitter (ITA Electronics Corporation Model 2500H) operating over the 120- to 155-Mc frequency band. Two nine-way power dividers (one for each of two orthogonal

linear polarizations) split the power equally into the nine radiators. A switching mechanism is provided to set up two orthogonal linear and both circular polarizations; however, switching cannot be accomplished while transmitting.

The radiating element is similar to the command element used on the 85-foot paraboloids described in this handbook. All discs are 29 inches in diameter and are spaced 26 inches apart on a step-tapered shaft designed to have constant strength (tension and compression) along its entire length. Spacing between the first disc and dipole driver is 11.5 inches and spacing between the first disc and the second disc is 22.5 inches to effect a good impedance match. Dipoles are oriented so that one set is parallel to the Y-axis and another set is normal to the Y-axis. The shaft is 242 inches long (3λ at 148 Mc) from crossed-dipole driver to tip, providing an element gain of 11.2 db. An open metal-frame tub, 44.25 inches deep and 64.25 inches in diameter, surrounds the crossed-dipole driver to equalize E- and H-plane beamwidths and to reduce back-lobe radiation. The bottom of this tub forms a reflector for the crossed-dipole driver and is located 22 inches behind the dipole. A dual-line balun is contained in the shaft of each disc-on-rod element.

Figure 44 shows a plan view of the arrangement of the disc-on-rod elements in this array. This unique arrangement of elements, in a coadunated triangle pattern, combines the individual apertures of all the disc-on-rod elements with the minimum of lost space or overlap. As a result, 90.8 percent of the aperture area is utilized whereas only 78.5 percent would result from a checkerboard arrangements of elements. Figure 44 shows that elements 1, 3, and 9; 3, 4, and 6; and 6, 7, and 9 lie on the points of three contiguous triangles. Elements 2, 5, and 8 lie on the mid-points of these base lines and on the points of the equilateral triangle. The elements are erected on three crossed-structural beams, with three elements on each beam. In order to describe the aperture size, it can be said that elements 1, 4, and 7 lie 120 degrees apart on a circle of 160-inch radius; elements 3, 6, and 9 lie 120 degrees apart on a circle of 120-inch radius; and elements 2, 5, and 8 lie 120 degrees apart on a circle of 60-inch radius. The total weight carried by the pedestal is 2450 pounds.

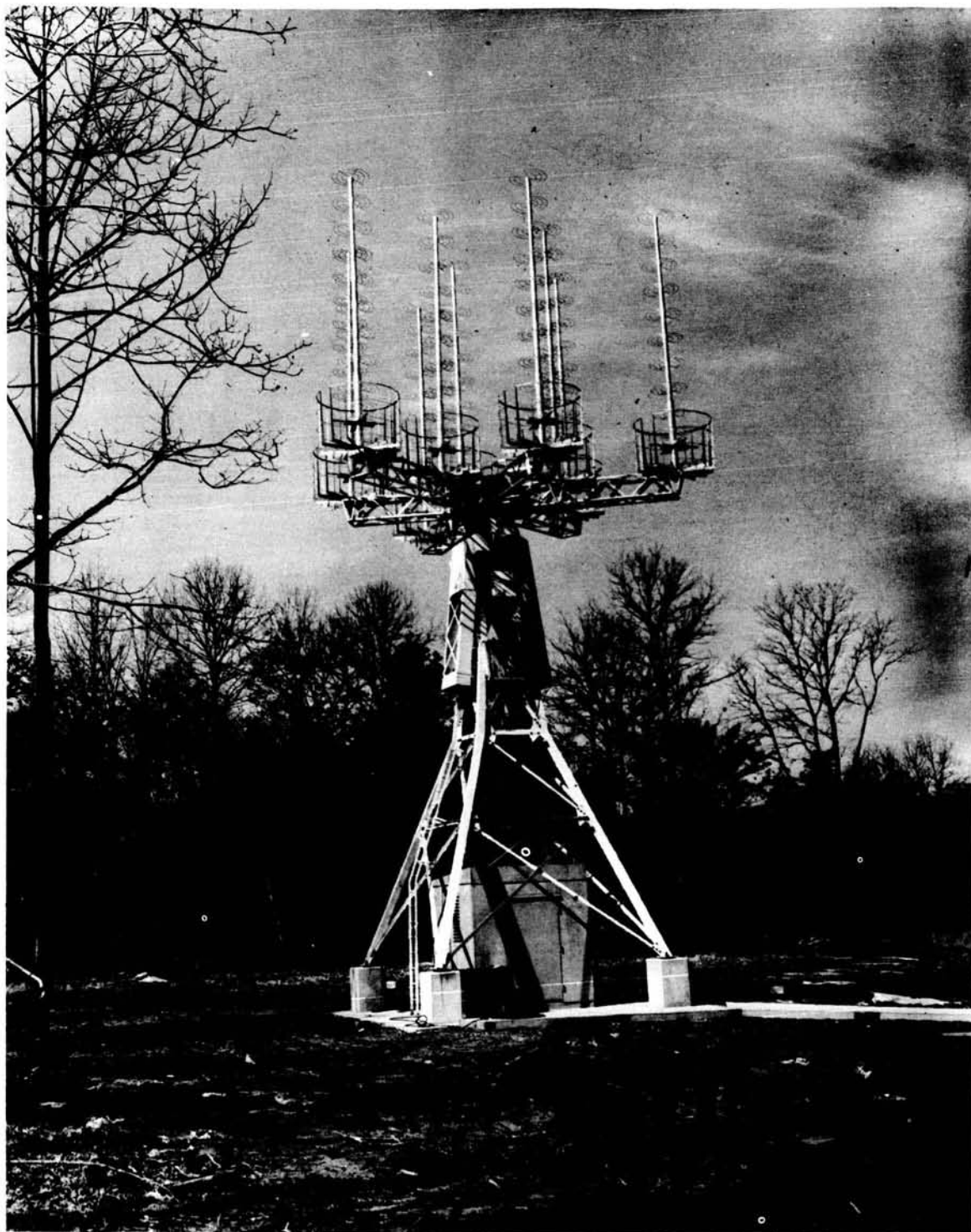


FIGURE 43.—Broadband high-power command array.

TABLE 11.—*Pattern Performance Broadband High-Power Command Antenna*

Frequency (Mc)	Beamwidth (°)			Sidelobe (db)		
	E-Plane	H-Plane	RHC	E-Plane	H-Plane	RHC
115.....	15.5	15.2	-----	16.5	15.1	-----
123.....	15.2	15.3	15.0	18.5	18.3	18.0
130.....	15.0	14.0	-----	16.9	18.0	-----
140.....	12.5	12.4	-----	15.8	18.4	-----
148.....	11.0	11.1	11.1	19.1	16.3	15.1
155.....	11.1	11.0	-----	14.6	13.2	-----

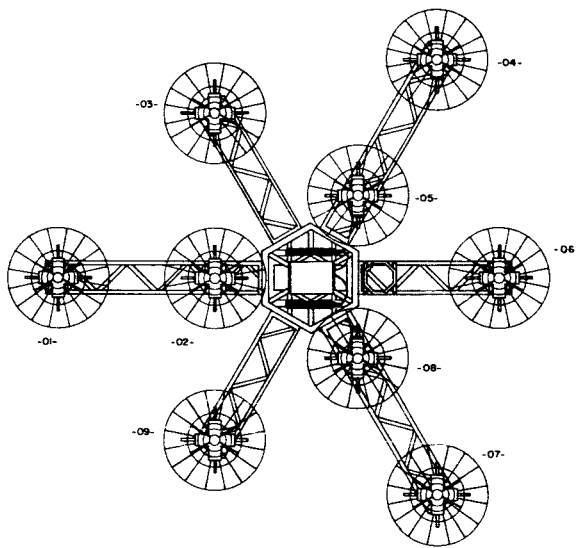


FIGURE 44.—Plan view, broadband high-power command array.

Pattern Performance

The radiation patterns of the antenna were measured by mounting on an elevation-over-azimuth pedestal and rotating about the (effective) Y-axis only. In this manner, the E- and H-Plane patterns and the circularly polarized patterns were measured. Representative patterns at 123 Mc and 148 Mc are plotted in Figure 45. It can be seen that the pattern is relatively unaffected by polarization change. The beamwidth falls off with frequency as expected and is 15 degrees wide at 123 Mc, decreasing to 12 degrees at 148 Mc. Sidelobes are well suppressed and are 18.0 db down at 123 Mc with a slight in-

crease at 148 Mc. Table 11 summarizes the measured pattern performance, and shows that the array has been adjusted for peak performance at 123 and 148 Mc.

Gain

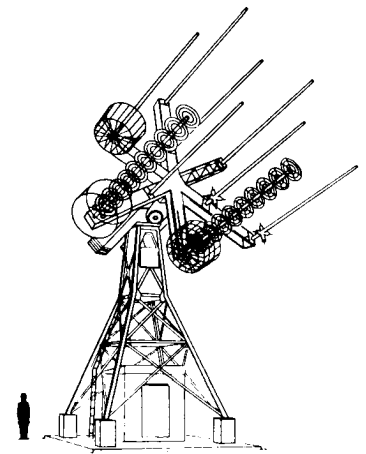
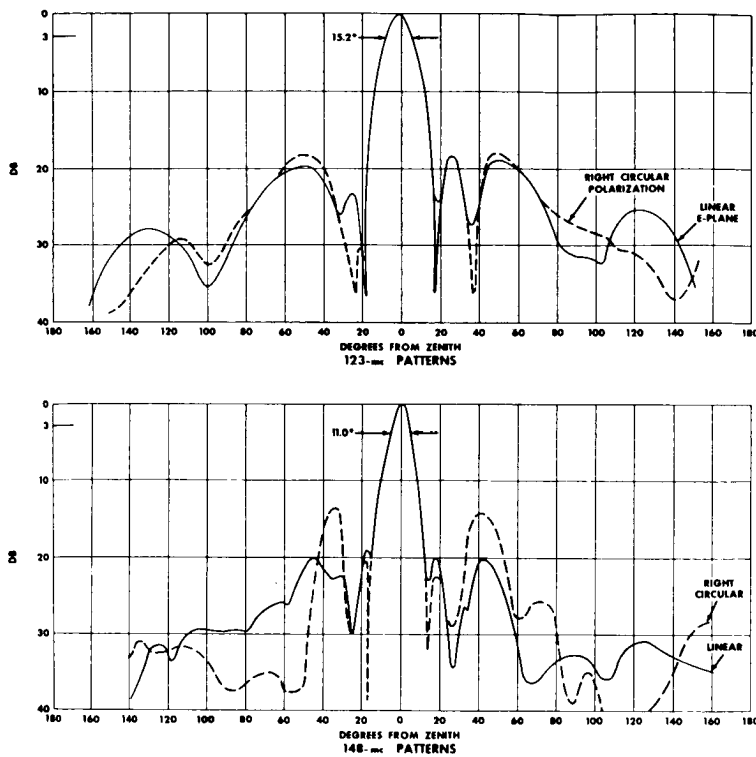
The gain of the array was measured by comparison with a calibrated standard and by integration of the volume under a pattern of radiated power. This was done at 123 Mc as well as at 148 Mc. The computed and measured values agree to within 0.5 db and the average of these values are listed in Table 12 as the gain above an isotropic source.

TABLE 12.—*Measured Gain Broadband High-Power Command Antenna*

Frequency (Mc)	Gain (db)
123.....	22.5
148.....	23.7

Voltage-Standing Wave Ratio (VSWR)

The VSWR seen looking into the polarization switch box was measured over the 120- to 155-Mc frequency band. Measurements were made looking into both linear and both circular polarization ports. Little change was observed in VSWR with polarization. Table 13 shows the values measured for one of the linear polarizations. The array is well matched at the 123- and 148-Mc operational frequencies.



NOTE:
GAIN: 22.5 db ABOVE AN ISOTROPE AT 123 mc
23.7 db ABOVE AN ISOTROPE AT 148 mc

FIGURE 45.—Radiation patterns, broadband high-power command antenna.

TABLE 13.—VSWR Broadband High-Power Command Array

Frequency (Mc)	VSWR
120	1.06
123	1.18
130	1.10
140	1.70
148	1.01
149	1.22
155	1.60

IV. GODDARD RANGE AND RANGE-RATE ANTENNAS

The Goddard range and range-rate system is a precision tracking-antenna system, nominally capable of measuring satellite range to an accuracy of 15 meters rms, rate of range change to an accuracy of 0.1 meter per second rms, and direction to an accuracy of 0.1 degree. The ground antenna transmits a signal which is translated in

frequency in the satellite and returned to the antenna (or alternatively to three antennas for trilateration). The phase delay in sidelobes is proportional to the two-way range. The ultimate system operates at S-band frequencies. Three antenna systems are in existence: one at Rosman, N.C. and one at Carnarvon, Australia. A location for the third system has not been assigned.

Very High Frequency (VHF) Antenna Array

Motorola, Inc.
Scottsdale, Arizona
NAS5-1962

General Description

The VHF antenna (Figure 46) is a 28- \times 28-foot monopulse-tracking phased array of cavity-backed slots. It functions as a transmitting antenna as well as a receiving antenna. The antenna was developed for use with the existing satellite frequencies, transmitting at 148 Mc and 136 Mc. It is a 6- \times 6-element planar array with the 36 slots (in each polarization plane) arranged

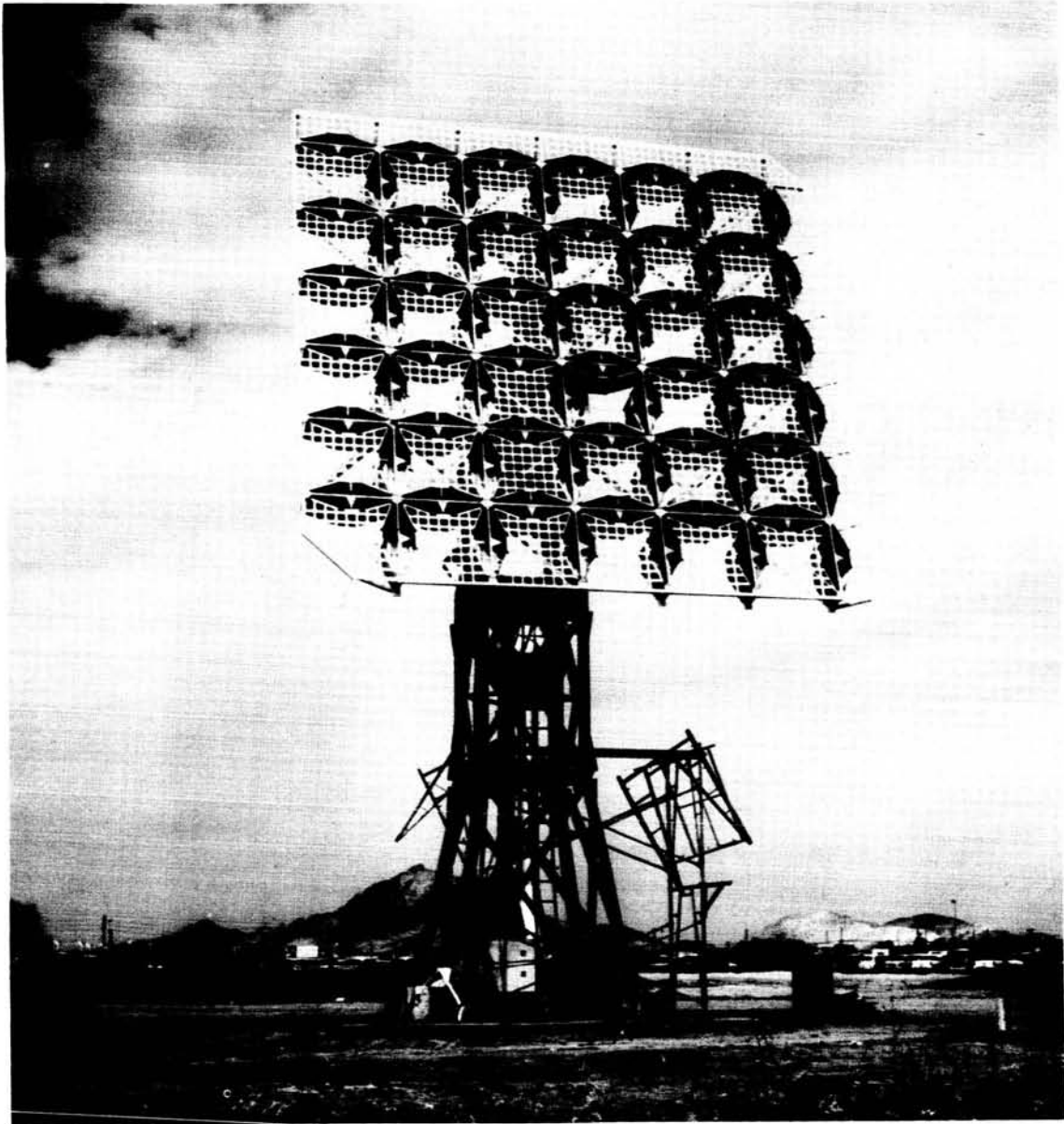


FIGURE 46.—Goddard range and range-rate antenna (VHF).

in four quadrants for phase monopulse tracking. Effective phase center spacing in the difference mode is 2.51λ in the E-plane and 2.56 in the H-plane, to yield the difference-pattern peak separation shown in Figure 47. The antenna provides for manual selection of the following polarizations for transmit or receive: right-hand circular, left-hand circular, parallel to the Y-axis linear, and parallel to the X-axis linear. It may be manually

positioned or sector scanned, one axis at a time. The relatively broad beamwidth is sector scanned until it locks onto the 136-Mc beacon in the satellite. The antenna is capable of supporting 12.0-kw peak power and is carried on a hydraulically driven X-Y pedestal fitted with coaxial rotary joints. Both the antenna array and pedestal were built by Antenna Systems, Inc., on a sub-contract from Motorola, Inc.

Slot center-to-center spacing is 53 inches (0.613λ at 136 Mc, and 0.669λ at 148 Mc). The slots are 48.1 inches long and 17.9 inches wide, tapering to 4 inches at the ends. The cavities behind each slot are 20 inches deep and are covered with 0.3-inch-thick sheets of Fiberglas polyester. The characteristic impedance of each slot is individually adjusted by means of tuning devices located inside the cavity, and by means of a loop across the slot outside the Fiberglas cover.

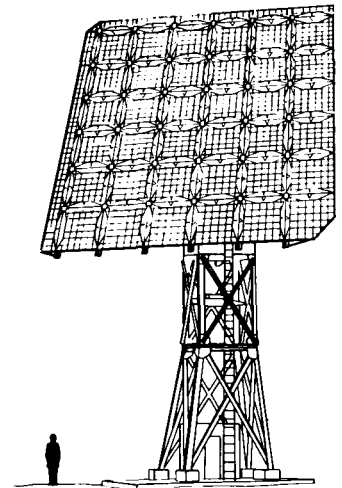
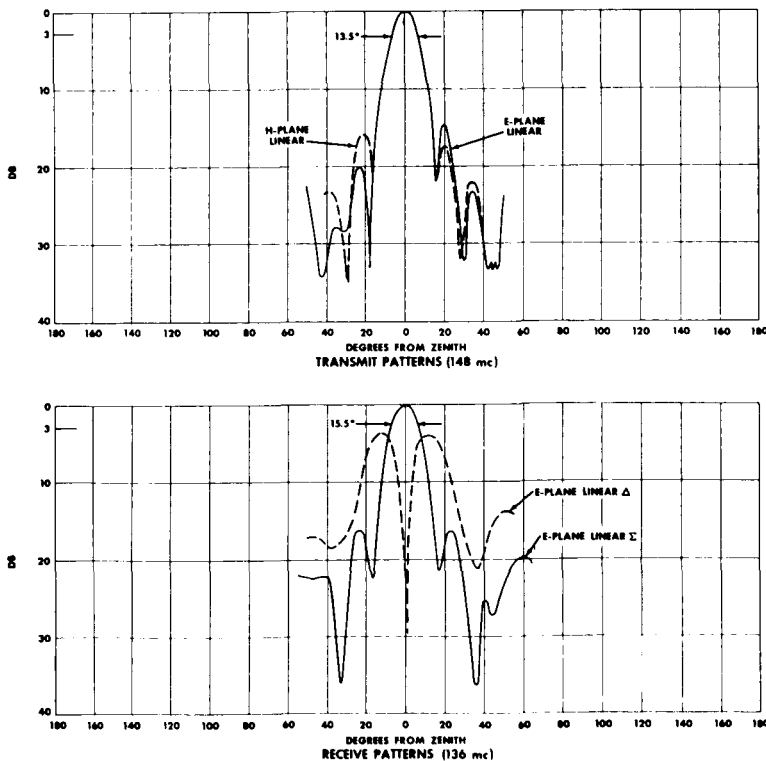
The current amplitudes across the array are tapered with the Tschebyscheff distribution for 17-db sidelobes. Although the coefficients for such a taper are 0.7, 0.9, 1.0, 1.0, 0.9, and 0.7, the coefficients for edge slots were tapered so that the corner slot coefficients were 0.49 to suppress sidelobes in the 45-degree planes. The feed system is composed of $\frac{7}{8}$ -inch and $\frac{1}{2}$ -inch 50-ohm Styroflex and RG-9/U "Aljak" (Amphenol type 21-5920). The polarization switching assembly feeds four 3-way power dividers each of which feed three 3-way power dividers. In this manner,

36 slots are fed in each orthogonal plane. RF rotary joints are not used.

The array is constructed of aluminum alloy. The total weight carried by the pedestal is 38,000 pounds, including the counterweight.

Radiation Performance

Radiation pattern characteristics were measured for the transmitting function over the 147- to 151-Mc band, in both the E- and H-planes. It was observed that the patterns were quite similar in both the E- and H-planes at all frequencies. Beamwidth decreased approximately 1.2 degrees per megacycle. In addition, pattern performance was measured for the receiving function, over the 134.5- to 138.5-Mc band and again the pattern was essentially unchanged over the band. Figure 47 is a plot of patterns measured at the operational frequencies 136 and 148 Mc. The pattern at 136 Mc was exactly symmetrical in the E- and H-planes down to the 16-db level. Beamwidth of the sum pattern can be seen to be 15.5 degrees and sidelobes are approximately 14.0 db down. The



NOTE:
GAIN ABOVE AN ISOTROPE
21.25 db AT 148.3 mc
20.50 db AT 136.5 mc

FIGURE 47.—Patterns, range and range-rate antenna (VHF).

difference-pattern slope is 3.5v ratio/degree (11.5 db). At the transmit frequency, the beam-width can be seen to be reduced to 13.5 degrees and sidelobes are 15.0 db down.

Gain

The gain of the array was measured by comparison with a calibrated standard (a 60-×85-inch corner reflector) in both the E- and H-planes at 136 and 148 Mc. This measurement was made independently at two locations and the measured findings were in very close agreement. Table 14 lists the measured gain relative to an

TABLE 14.—Measured Gain Goddard Range and Range-Rate Antenna (VHF)

Frequency (Mc)	Gain (db)	
	E-Plane	H-Plane
136.....	20.50	20.64
148.....	21.25	21.45

isotropic source. These measured values were compared to the gain computed from an integration of the area under a power plot of the patterns and agreement to within 0.2 db was obtained. The RF coaxial feed system insertion loss was measured to be 0.46 db. Measured gain was corrected in this amount for comparison with computed gain.

VSWR

The individual slots of this array are surprisingly broadband, presenting a VSWR less than 1.20 over a 15 percent frequency band. Mutual coupling effects, however, degrade this performance. Extensive efforts were made to reduce coupling and this characteristic was exhaustively measured. Maximum coupling is typically 17.0 db. This corresponds to a reflection coefficient of 0.14 and will change a perfect match to a VSWR of 1.3. Hence, it was necessary to utilize matching devices in the form of loops across each slot. Table 15 lists measured VSWR. These values include polarizer, diplexer, and, in the case of the transmitting frequencies, RF rotary joints.

Miscellaneous Characteristics

The boresight shift with frequency, over the tracking band, was measured and found to be somewhat erratic with frequency. With the beam position at 136 Mc as a reference, a shift of 3 minutes to the left was observed at 137 Mc and 17 seconds to the left at 138.5 Mc; the beam shifted 10 seconds to the right at 135 Mc and did not shift at 134.5 Mc.

Pointing accuracy of the antenna system is 0.5 degree, and the slew rate capability of the prime mover is 15 degrees per second. The prime driver is capable of acceleration or deceleration at rates up to 5 degrees per second.

S-Band Antenna

Motorola, Inc.
Scottsdale, Arizona
NAS5-1926

TABLE 15.—VSWR Goddard Range and Range-Rate Antenna (VHF)

Frequency (Mc)	VSWR			
	Y-Axis	X-Axis	Right Circular	Left Circular
147.....	1.18	1.16	1.09	1.10
148.....	1.16	1.06	1.07	1.11
149.....	1.20	1.15	1.06	1.16
150.....	1.18	1.12	1.06	1.06
151.....	1.20	1.15	1.17	1.02
136 (sum).....	1.09	1.10	1.05	1.08

General Description

The Goddard S-band range and range-rate antenna (Figure 48) consists of two Cassegrainian-fed 14-foot-diameter paraboloidal reflectors ($f/D = 0.469$) mounted on a hydraulically driven X-Y pedestal identical to that used for the VHF antenna. The parabolas are spaced 15 feet apart on the Y-axis (upper axis) resulting in a 12-inch clearance between reflector edges. One parabola transmits at 2271 Mc and the other receives at 1705 Mc. The hyperbolic subreflector, for both systems, is 36.5 inches in diameter with a focal length of 15.55 inches and a focal depth of 72.02 inches. This subreflector is located 55.16 inches away from the aperture of the primary illumination horn which is set 6.7 inches inside the vertex of the parabola. Consequently, the angle on the primary horn pattern subtended by the subreflec-

tor is 16.96 degrees. The aperture blockage ratio is 47 percent and rather high. The effective focal length of a Cassegrainian antenna system is the product of the magnification factor and focal length of the prime reflector. For this antenna $MF = 23.47$ feet.

The prime feed for the transmitting antenna is a pyramidal horn with square aperture 19.5 inches in inside dimension. It is mounted on a section of round waveguide attached to a section of circular waveguide containing 12 pins for use in changing polarization. This feed is covered with a flat sheet of Fiberglas and is pressurized to 2 psig providing a capability for supporting 12-kw peak power. Circular waveguide RF rotary joints are used on both axes of motion. The prime feed for the receiving (tracking) antenna consists of four phase-monopulse pyramidal horns with square

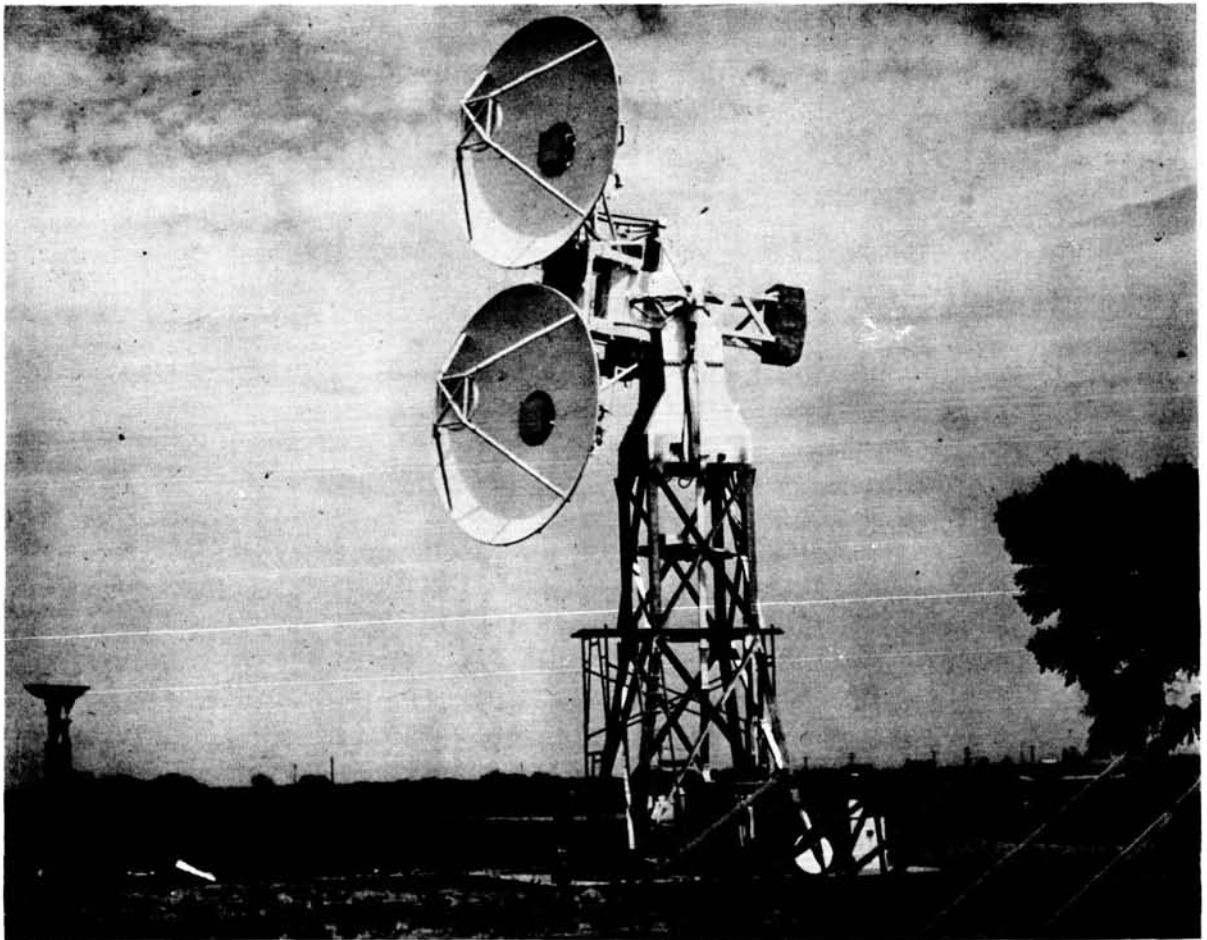


FIGURE 48.—Range and range-rate antenna (S-Band).

aperture 9.75 inches in inside dimension and a flare angle of 18.06 degrees. Each horn is mounted on a 17-inch length of round waveguide. Polarization selection capability is provided, but rotary joints are not used. The comparator circuitry is constructed of RG-104/U waveguide. The total weight carried by the pedestal is 21,600 pounds, including the counterweight.

Pattern Performance

Radiation pattern performance was measured at the transmitting and receiving frequency for all combinations of polarization adjustment. It was noted that pattern characteristics were unchanged by polarization change. Figure 49 is a plot of the primary and secondary patterns at 2270 Mc. The angle subtended by the hyperboloidal subreflector (16.96 degrees) is indicated on the primary horn patterns. When space attenuation is added, subreflector edge illumination is 7.7 db in the E-plane and 9.2 db in the H-plane. The high shoulders in the H-plane are characteristic of pyramidal horns. Secondary patterns are shown in Figure

49 for linear polarization parallel to the X-axis and linear polarization parallel to the Y-axis. The beamwidth can be seen to be 1.8 degrees and sidelobes are below 16.0 db. Figure 50 is a plot of the circularity and secondary patterns at 1705 Mc. Primary patterns were not measured. The beamwidth of the X-axis sum pattern is 2.6 degrees and sidelobes are below 16.0 db down. The difference-pattern null is 40.0 db down, resulting in a steep taper. Pattern characteristics for other polarizations were the same as shown in Figure 50. Transmit circularity was measured to be 0.2 db for right-circular polarization, an exceedingly symmetrical condition, and receive circularity was measured to be 2.6 db.

Gain

The gain of the two paraboloidal reflector antennas was measured by comparison with a calibrated standard and found to be 35.0 db at 1705 Mc and 37.0 db at 2271 Mc. These values are referred to the antenna feed input and do not include comparator loss. The measured gain

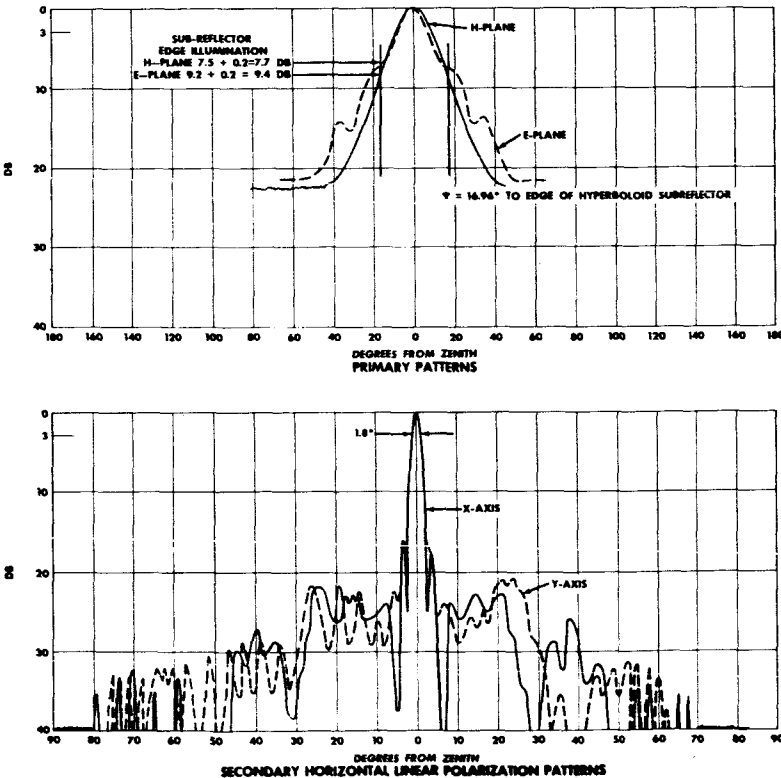
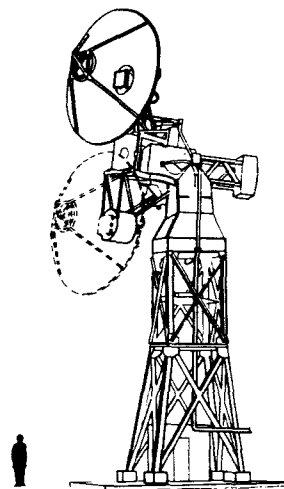


FIGURE 49.—Transmit patterns, range and range-rate antenna (S-band).



NOTE
GAIN ABOVE AN ISOTROPE
37.0 DB AT 2271 MC

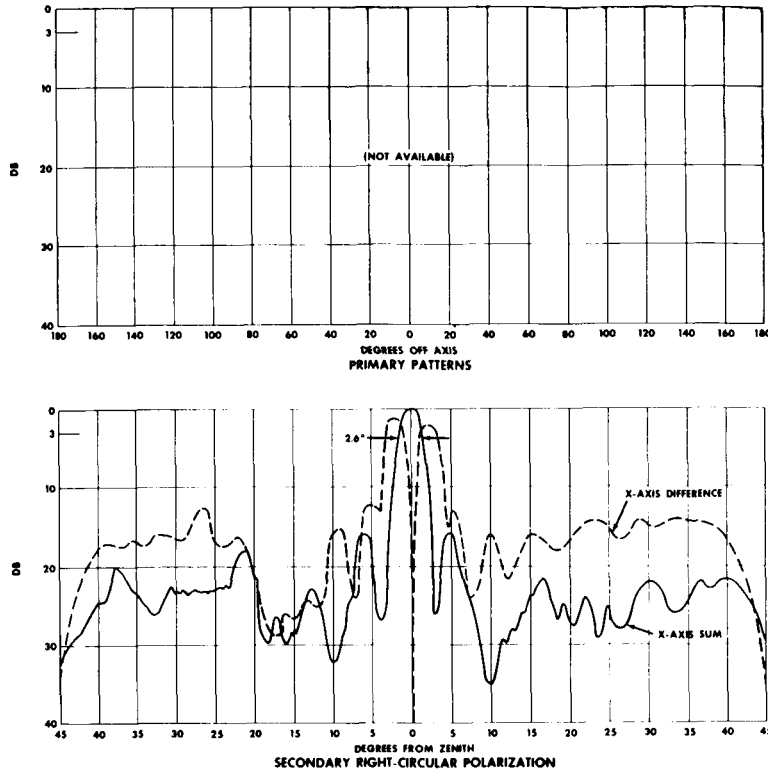


FIGURE 50.—Receive patterns, range and range-rate antenna (S-band).

TABLE 16.—VSWR Goddard Range and Range-Rate Antenna S-Band Receiving System

Frequency (Mc)	Polarization	Reference Terminal		
		Sum	X-Difference	Y-Difference
1700-----	Vertical	1.18	1.20	1.17
1705-----		1.10	1.17	1.15
1710-----		1.16	1.17	1.14
1700-----	Horizontal	1.15	1.19	1.19
1705-----		1.17	1.20	1.14
1710-----		1.10	1.18	1.16
1700-----	Right-Circular	1.09	1.17	1.14
1705-----		1.14	1.20	1.13
1710-----		1.18	1.16	1.19
1700-----	Left-Circular	1.13	1.14	1.13
1705-----		1.09	1.16	1.16
1710-----		1.18	1.18	1.19

values represent 50 percent efficiency at 1705 Mc and 48 percent at 2271 Mc.

VSWR

The VSWR, looking into the feed of the transmitting antenna as well as into the receiving antenna, was measured. A good match was observed, and measured VSWR is given in Table 16 for the receiving antenna and in Table 17 for the transmitting antenna.

TABLE 17.—VSWR Goddard Range and Range-Rate Antenna S-Band Transmitting System

Frequency (Mc)	Polarization	
	Right Circular	Left Circular
2225.....	1.18	1.14
2240.....	1.17	1.12
2255.....	1.14	1.17
2270.....	1.15	1.09
2285.....	1.18	1.13
2300.....	1.18	1.13
2315.....	1.17	1.12

Miscellaneous Characteristics

The RF coupling between paraboloids was measured at input for all polarizations. Coupling was found to be down 70.0 db at 2270 Mc and 72.0 db at 1705 Mc. Filters were inserted in the receiving feed line to reduce the coupling with the transmitting frequency to an additional 65.0 db, making the total system isolation 135.0 db at 2270 Mc. The pedestal is designed for a maximum velocity of 15 degrees per second and an acceleration of 5 degrees per second.

V. MISCELLANEOUS ANTENNAS

Dual-Yagi Command Antenna

Bendix Field Engineering Corporation
Owings Mills, Maryland
NAS5-734

General Description

This antenna (Figure 51) is an interim command transmitting device provided to all Mini-track stations for operation with Relay and subsequent satellites launched prior to the installation

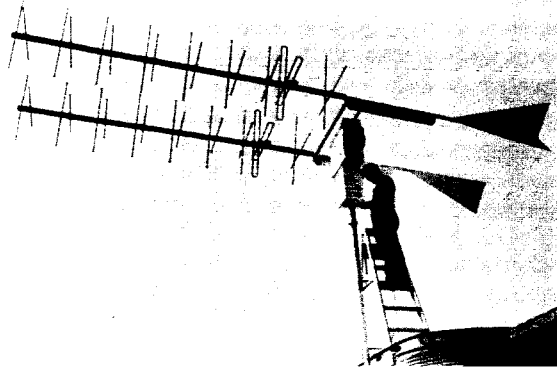


FIGURE 51.—Dual-yagi command antenna.

of the broadband, high-power command antenna described earlier in this handbook. It was built by Bendix Field Engineering Corporation, and consists of two crossed yagi radiators mounted 60 inches apart on a pole pedestal. The crossed yagi radiators were provided by Technical Appliance Corporation, and consist of seven directors above a folded dipole launcher, backed with a parasitic reflector element. One yagi structure is designed for operation at 123 Mc and the other is designed for operation at 148 Mc. The pair of yagi structures is mounted on two CDR HAM-M rotors (Cornell-Dubilier Electronics Model MRU-47). The upper rotor, together with a 5:1 ratio gear reduction box, provides elevation drive; the lower rotor provides azimuth drive. (These rotors are to be replaced with Telrex Laboratories Model 397-RIS heavy-duty rotors.) This dual assembly is mounted atop a 14-foot length of 2½-inch steel pipe. Windsails are provided to facilitate azimuth rotation and to act as elevation counterbalances.

The yagi structures are operated with a 200-watt Collins Radio Model 242G-2 transmitter located in the Minitrack operations building. Fifty-ohm Styroflex RF line ⅞-inch in diameter, connects the transmitter to the antenna. Polarization change is carried out manually by inserting tuned lengths of coaxial line in the RF feeding network. Lengths of extra flexible RG-281/U (FXR Amphenol Type 421-178) are used to bypass the axes of motion at the antenna. The antennas are capable of supporting 350-watts average power per yagi, or 700 watts for a crossed-

yagi assembly. They are constructed of aluminum. The 148-Mc yagi weighs 45 pounds and the 123-Mc yagi weighs 55 pounds. These are designated as Technical Appliance Corporation Models D-1444-8 and D-1444-3, respectively.

Pattern Performance

The radiation pattern performance was measured and is shown in Figure 52. Performance characteristics are summarized in Table 18.

Gain

The gain of each crossed yagi was measured by the identical antenna substitution method. It was found that the 123-Mc yagi has a gain of 12.7 db and that the 148-Mc yagi has a gain of 13.5 db.

Typical RF line loss between the transmitter and the antenna is 3.0 db.

VSWR

The input VSWR for one linear polarization was measured at the input to each yagi and was found to be 1.23 at 123 Mc and 1.06 at 148 Mc.

Log-Periodic Structure

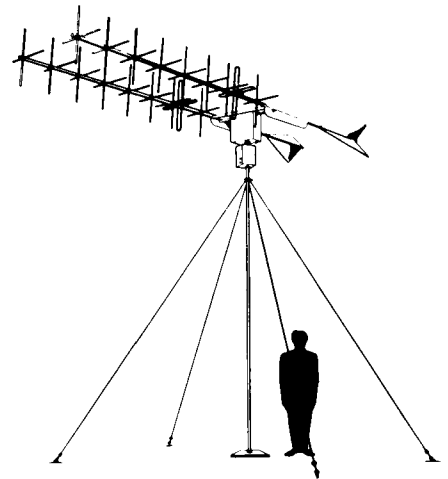
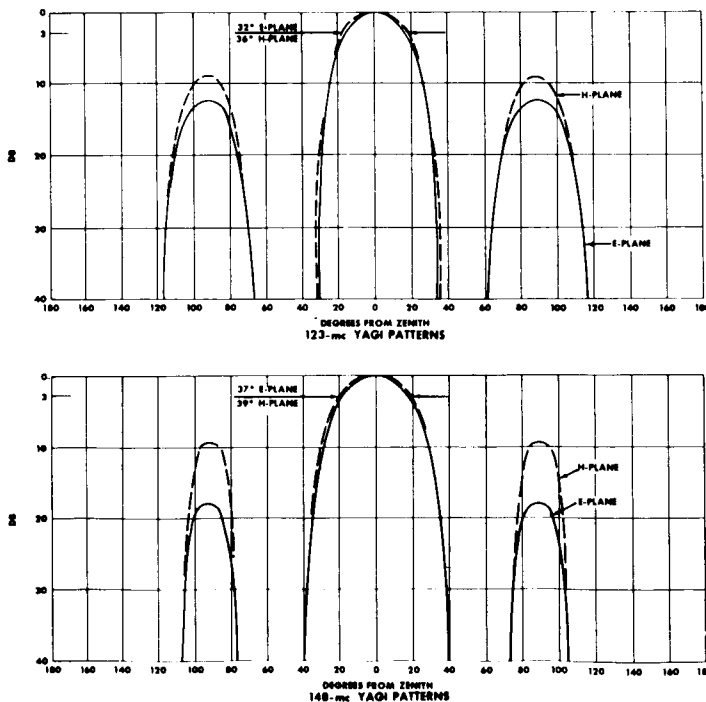
Collins Radio Company
Cedar Rapids, Iowa
(Naval Research Laboratory Contract)

General Description

The Collins Model 237A-2 log-periodic structure (Figure 53) is a two-bay antenna consisting

TABLE 18.—Dual-Yagi Command Antenna Performance

Frequency (Mc)	Beamwidth (°)		Sidelobe (db)	
	E-Plane	H-Plane	1st E-Plane	1st H-Plane
123	32	36	12.3	9.2
148	37	39	18.0	9.6



NOTE
GAIN ABOVE AN ISOTROPE
12.7 db AT 148 mc
12.3 db AT 123 mc

FIGURE 52.—Patterns of dual-yagi command antenna.

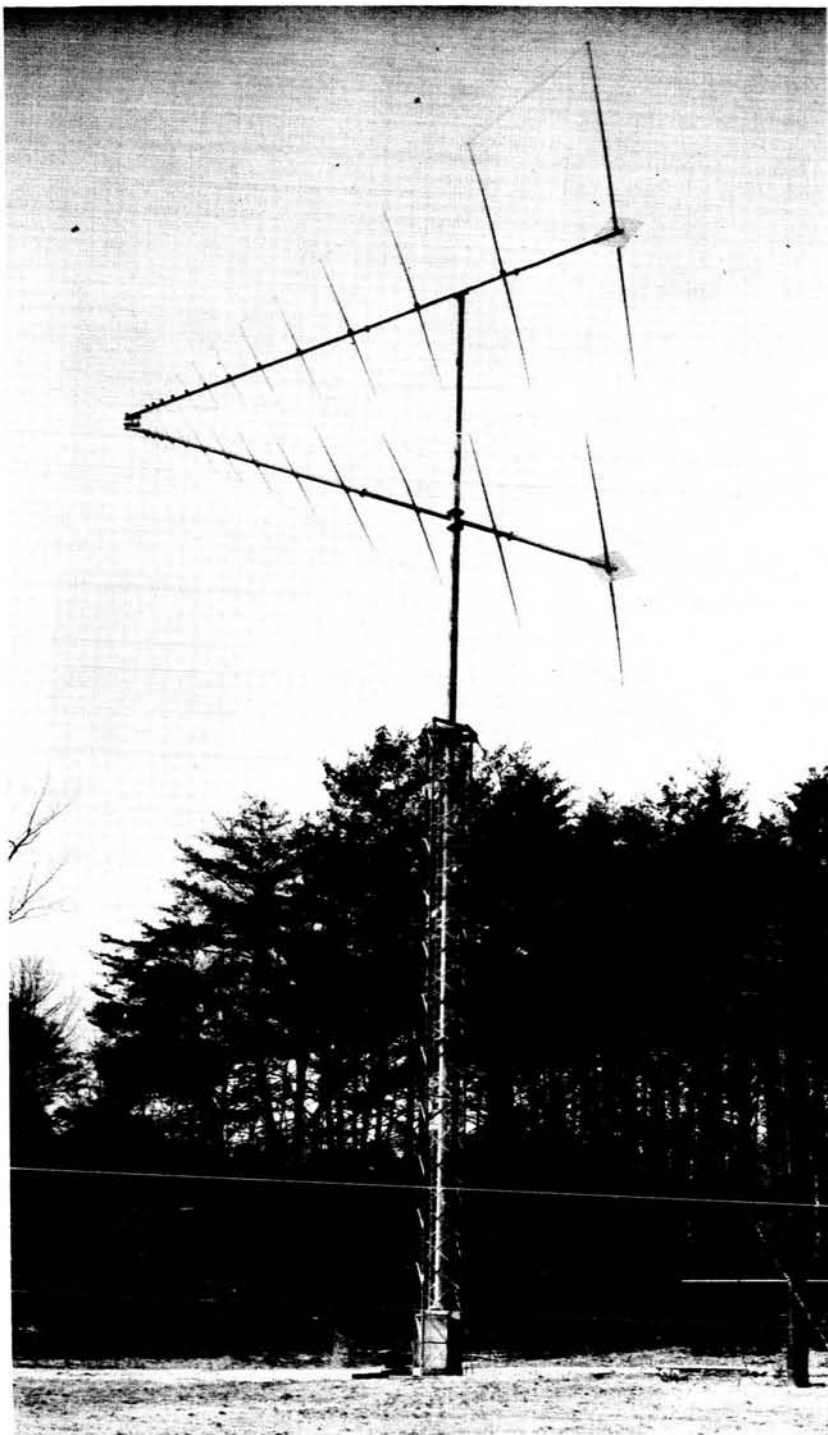


FIGURE 53.—11- to 60-Mc log-periodic structure, Blossom Point, Maryland.

of 13 elements in each bay. The bays are 41 feet long and the element lengths vary from 2 feet to 47 feet long, providing good pattern performance over the 11- to 60-Mc band. Unique characteristics of the antenna are low-input VSWR, high efficiency, and good pattern performance over a very broad frequency band. The antenna is horizontally polarized and has a free-space pencil beam off the apex. Because the antenna performs over a broad band, a balun cannot be used; consequently, the 150-ohm terminal-point impedance is matched to a 50-ohm $3\frac{1}{8}$ -inch coax through a tapered transmission line.

Pattern Performance

The effective elevation-plane (H-plane) pattern is the product of the free-space pattern and the ground effect factor, $2 \sin(2\pi h/\lambda \sin \theta)$. This pattern was computed for an 84-foot tower height. The findings were confirmed by measuring the elevation antenna pattern at the Blossom Point STADAN facility. Pattern measurements were made with a transmitter-equipped helicopter making vertical ascents. A dipole, on the ground,

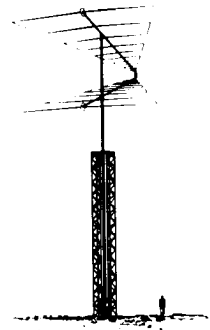
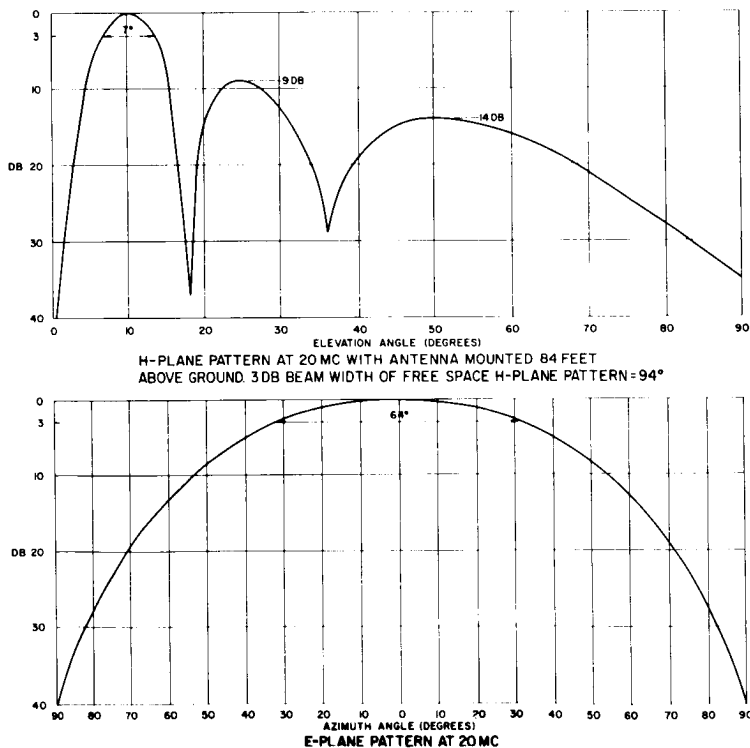
was used to monitor signal-strength variation with helicopter range. Figure 54, a plot of the measured pattern, shows an elevation beamwidth of 7 degrees, sidelobes 9-db and 14-db high, and nulls at 18 and 36 degrees. Because the azimuth-plane (E-plane) pattern is not modified by ground effect, it was not measured at Blossom Point. Figure 54 shows an E-plane beamwidth of 64 degrees for the free-space pattern, measured at Collins.

Antenna Gain

The gain of the antenna installed at Blossom Point was determined, by comparison with a dipole, to be 17.7 db corresponding to 64 percent efficiency. This gain, although higher than that claimed by the manufacturer, seems reasonable in comparison with a theoretical gain value of 18.2 db, determined by pattern integration.

Input Impedance

The input impedance measured at the base of the 84-foot Blossom Point tower showed the antenna was poorly matched, but suitable for



GAIN: (MEASURED)
17.7 DB ABOVE AN ISOTROPIC SOURCE AT 20 MC

FIGURE 54.—E- and H-plane patterns at 20-Mc, Collins log-periodic structure, Blossom Point, Maryland.

operation as a receiving antenna, since only 1.2-db transmission line loss is incurred for a VSWR of 3:1. The antenna should be used for transmitting only if it is matched to the specific transmitting frequency. Table 19 shows the data observed at Blossom Point.

Telemetry Test Antenna

Litton Industries
College Park, Maryland
NAS5-1555

General Description

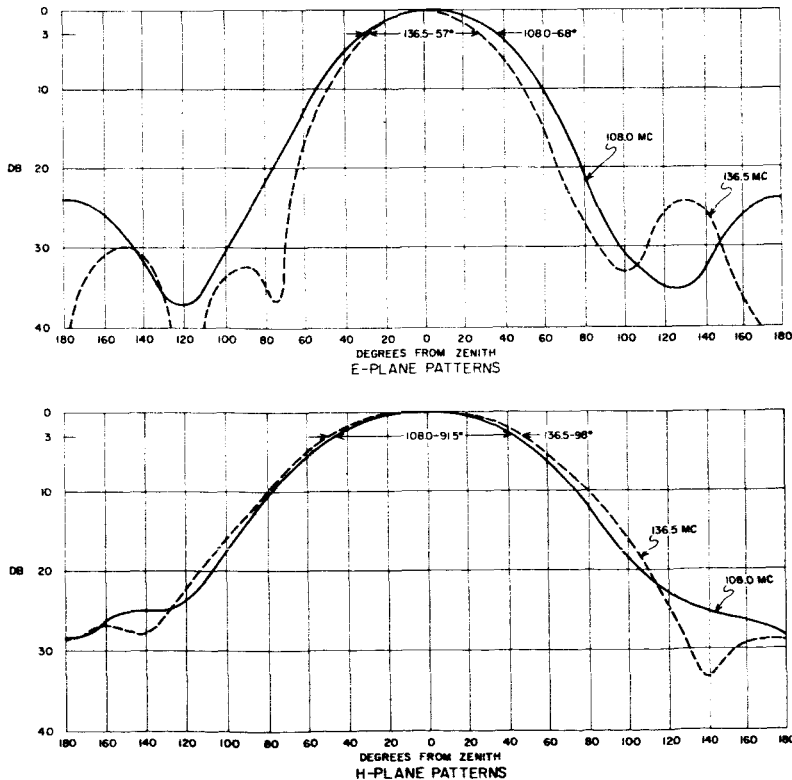
This antenna is provided to all STADAN facilities for calibrating the command and data-acquisition antennas, and associated system components. It consists of a Rohn Manufacturing Company (Peoria, Illinois) fold-over tower model 25 and two antenna standards. The tower is 69 feet tall, and can be extended an additional 8 feet by means of a top-fitted mast to which the antenna standards are attached.

One of the antenna calibration standards (Figure 55) is a planar logperiodic dipole array of

TABLE 19.—Input Impedance and VSWR 11- to 60-Mc Log-Periodic Structure

Frequency (Mc)	Impedance (ohms)	VSWR
11-----	60.0-j1.2	1.6
30-----	89.4-j0.8	1.7
50-----	68.7-j0.4	2.2
60-----	80.0-j0.8	3.0

eight elements, tuned to cover the 100- to 150-Mc frequency band. The antenna is fed by a coaxial line entering, and passing through and inside one of the balanced two-wire lines to the feed point. This antenna is mounted on a Cornell-Dubilier-Radiart model HAM-M rotator in order that it can be rotated about a horizontal axis in checking the circularity of antennas. The gain and pattern characteristics (Figure 55) of this standard antenna were precisely measured for use as a gain standard. (See Table 20.)



GAIN ABOVE AN ISOTROPIC SOURCE:
AT 108 MC - 7.6 DB
AT 120 MC - 8.0 DB
AT 136 MC - 8.1 DB
AT 150 MC - 8.2 DB

FIGURE 55.—Patterns of telemetry calibration and gain standard antenna.

TABLE 20.—Characteristics of Standard Log-Periodic Antennas

Frequency (Mc)	Gain* (db)	VSWR	Beamwidth	
			E-Plane	H-Plane
100.....	7.8	1.15	71.7	89.8
108.....	7.6	1.22	68.0	91.5
123.....	8.0	1.40	62.3	95.0
136.5.....	8.1	1.38	57.0	98.0
138.....	8.1	1.42	56.5	98.5
148.....	8.2	1.60	52.6	100.6
160.....	8.3	1.90	47.9	103.4

*Above a linear isotropic source.

Another antenna provided with this tower is a crossed-sleeve dipole which has been opened out into a flat sheet and mounted $\lambda/4$ above a cross-shaped screen composed of four aluminum slats.

The antenna covers the 90- to 160-Mc frequency band with a VSWR of 1.5 (3.5 db) or less. This antenna cannot be rotated; however, linear or circular polarization can be achieved by manually adjusting the feed cables. The purpose of the antenna is to provide a circularly polarized gain standard.

REFERENCES

1. RIBLET, H. H., "Slotted Dipole Impedance Theory," MIT Rad Lab Report 772, 1945.
2. BROWN, JACOB S., "Determination Antenna Gain from Measured Radiation Patterns," Syracuse University final report on contract AF-30(602)-300, Task EE-17, 30 September 1953.
3. SILVER, S., "Microwave Antenna Theory and Design," McGraw-Hill, 1949, p. 416.
4. JASIK, H., "Antenna Engineering Handbook," McGraw-Hill, 1961, p. 12-16.
5. CROMPTON, J. W., "On Optimum Illumination Taper for the Objective of a Microwave Aerial," JIEE 1954, p. 371.
6. ROBERTS, W. K., "A New Wide-Band Balun," *Proc. I.R.E.*, Vol. 45, No. 12, December 1957.

FLIGHT VIBRATION DATA OF THE AEROBEE 150A SOUNDING ROCKET*

JAMES A. NAGY AND GOMER L. COBLE, JR.

Two channels of vibration data were obtained from each of two Aerobees launched from Wallops Island. Vibration data for two vehicle axes were obtained during the powered phase of flight from accelerometers mounted on the attitude control system (ACS). The highest vibration levels occurred while the rocket was rail guided. During this period the average levels were approximately 7 g's RMS longitudinally and 7-8 g's RMS laterally. After tower exit vibration levels dropped to low values and built up slightly at maximum q and at rocket motor burnout. Previous flights of three-fin Aerobees indicated vibration levels, during the tower portion of flight, which were three times greater than those of the four-fin Aerobees.

INTRODUCTION

The primary objective of two Aerobee 150A rocket flights (NASA 4.20 launched on June 26, 1961 and NASA 4.68 launched on January 13, 1962) was to flight test the maneuvering capability, holding, and acquisition accuracy of a three axis, gyro referenced and controlled, rocket attitude control system. As a secondary effort, both rockets carried two scientific experiments and a vibration experiment. The objective of the latter is to provide launch and flight vibration data. Rocket borne payloads are required to demonstrate their ability to withstand the exposure to vibration encountered during flight. Laboratory vibration tests are conducted to provide an added safeguard against vibration damage.

The vibration environment must be adequately defined if laboratory test criteria are to be meaningful. Thus, the value of the test program is largely a function of the available flight data. In the case of Aerobee, some data were available from a Fort Churchill launch of a three-fin Aerobee.† These data indicated high levels and were questioned as not being applicable to four-fin Aerobees launched from Wallops Island.

The purpose of this report is to present the data measured during two flights of the four-fin

Aerobees. The vehicle and launch tower configurations are described and time histories of flight performance are presented.

THE FOUR-RAIL TOWER

The Wallops Island Aerobee four-rail tower (Figure 1) is 160 feet high and is supported on a gimbal-mounting 80 feet from the base. The tower is enclosed for all-weather operation. Figure 2 shows the pre-launch vehicle-tower arrangement. Tower azimuth and tilt are controlled from the blockhouse so that rockets can be fired in a direction to compensate for wind effects on the vehicle, thus minimizing dispersion of the impact point. Two blast doors at the bottom of the building are opened to relieve smoke and blast pressure (Figure 1). Figure 3 shows the four-rail tower with the rail designations referred to throughout this report.

The Aerobee 150A sounding rocket is an expendable, boosted, fin-stabilized, free-flight liquid-propellant rocket-powered vehicle. It has an extended cylindrical configuration with a 31-caliber, ogival nose (Figure 4). Payloads are mounted inside the nose structure and, if desired, a payload extension and a camera extension are utilized. Four fixed-fins are spaced 90 degrees apart around the aft end of the rocket to provide aerodynamic stability. Figure 5 presents the general arrangement of the Aerobee 150A. These

*Published as NASA Technical Note D-2314, June 1964.

†Roth, C. E., Jr., and Swanson, R. S., "Environmental Measurements in Sounding Rockets," IES Proceedings, 1961.



FIGURE 1.—Aerobee launch facility, Wallops Island, Virginia.

finns are adjustable and may be canted from $0^{\circ}0'$ to $0^{\circ}20'$ for a desired roll-rate. The rocket is boosted from a four-rail tower by a solid-propellant booster. Initial guidance is provided by riding lugs on the rocket (fore and aft), booster riding lugs, and a booster guide lug (Figures 6 and 7). Replicas of the fore and aft rocket riding lugs of NASA 4.48, which were recovered after the operation, are shown in Figures 8 and 9, respectively. The booster is a high-thrust, short duration rocket motor. The booster fins are installed at a preset angle of $2^{\circ}30'$ to impart initially a high-roll rate to the vehicle as it leaves the launching tower. The booster and vehicle are position-mated by a thrust structure and remain together by gravity and/or the overriding booster thrust until booster burnout.



FIGURE 2.—Aerobee launch facility, interior view.

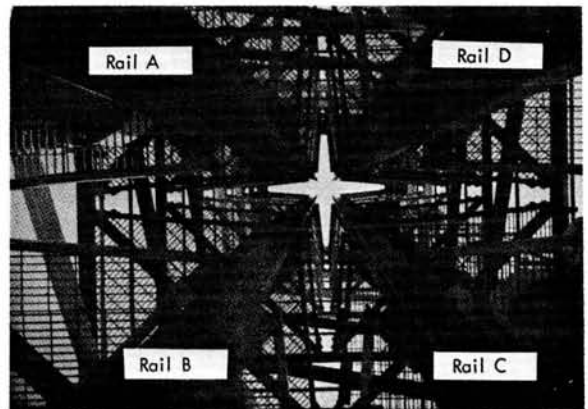


FIGURE 3.—Aerobee launch tower, looking up.

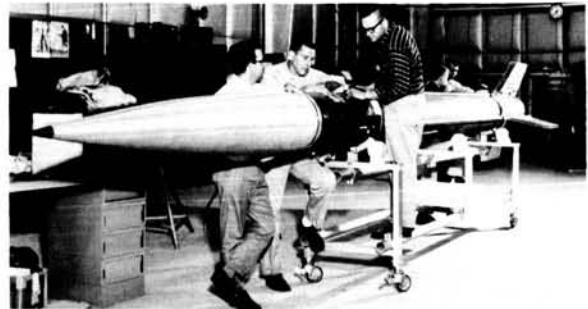


FIGURE 4.—Aerobee 150 A.

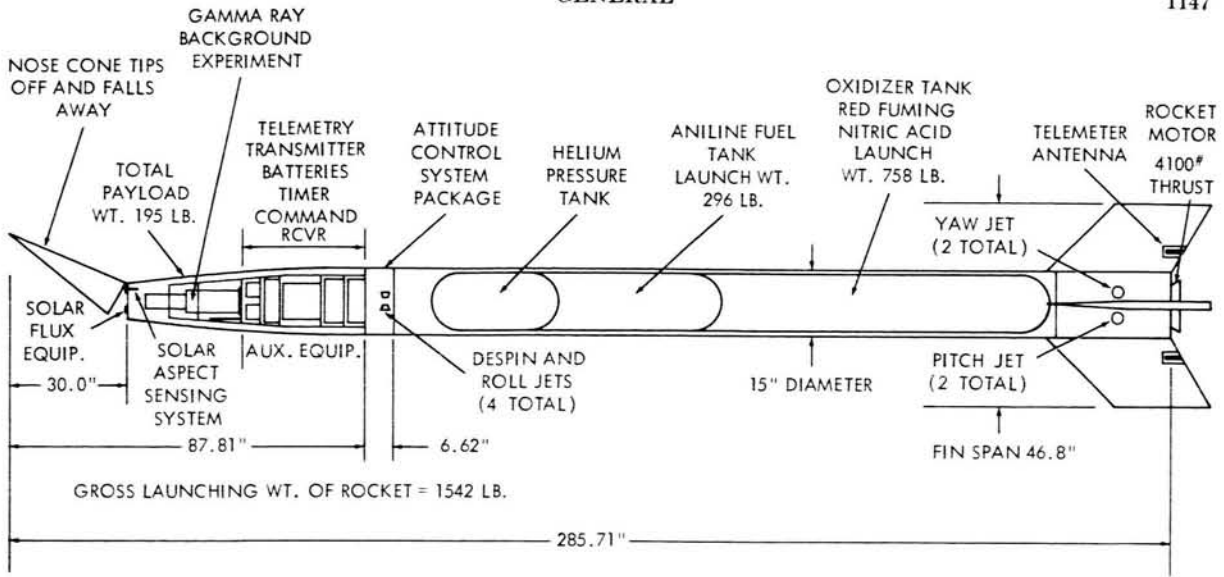


FIGURE 5.—Aerobee 150A, general arrangement.

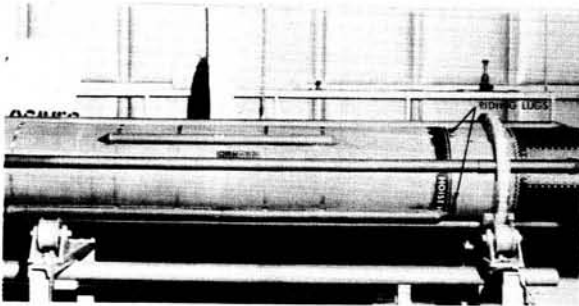


FIGURE 6.—Portion of Aerobee rocket showing forward riding lugs.

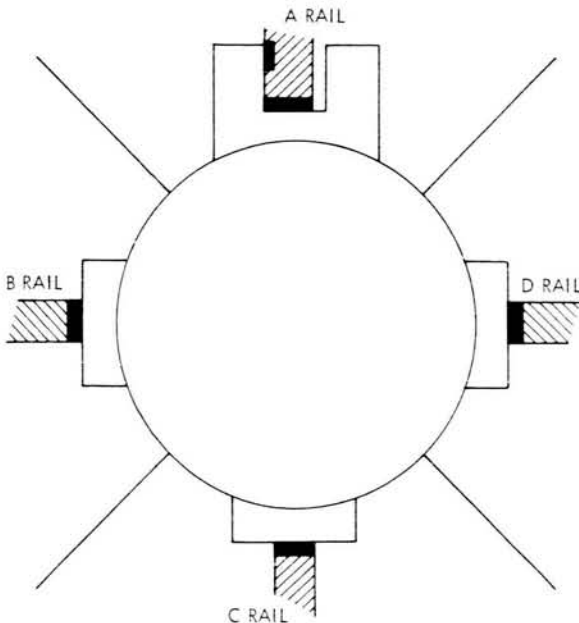


FIGURE 7.—Aerobee 150A orientation in tower (looking up) and view of rail-guide lug contact points (shaded).

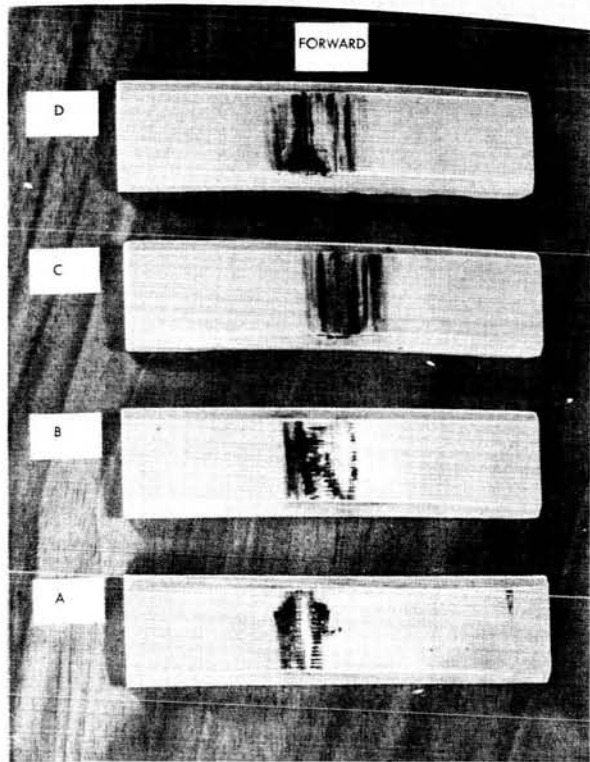


FIGURE 8.—Replicas of NASA forward rocket riding lugs.

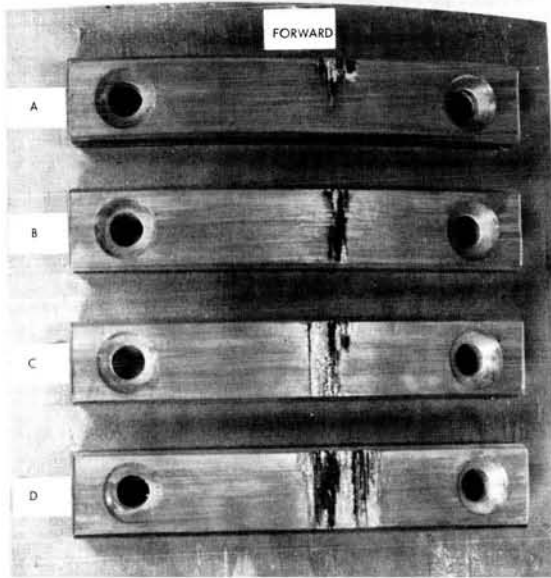


FIGURE 9.—Replicas of NASA aft rocket riding lugs.

FLIGHT TIME HISTORY

The major vehicle flight parameters and event times for NASA 4.20 and NASA 4.68 tests are presented in Table 1. It is appropriate to briefly discuss the flight time history of Aerobee as a reference for the flight vibration data. For high initial thrust, the booster is fired first by remote

Table 1
Vehicle Flight Parameters and Event Times;
Liftoff; $t = 0$ sec.

Event	Time (Sec)	
	NASA 4.20	NASA 4.68
Squib Firing	-0.3	-0.3
Liftoff	0.0	0.0
Booster Explosion	2.0	-
Hanger Exit	0.75*	0.78
Tower Exit	1.1*	1.1*
Booster Burnout	-	2.5*
Maximum Q	18.0*	19.0*
Rocket Burnout	52.0*	52.0*

*Approximate times.

control. The firing signal causes the overboard dump valve, incorporated in the pressure regulator valve, to be closed by a squib-activated piston located just aft of the attitude control system. This is the first event sensed by the vibration pickups. A microswitch senses this arming and completes circuitry to ignite the booster. As the vehicle moves vertically by initial booster thrust, it pulls away from a trip wire anchored to the launching tower. Resultant pulling of the trip wire initiates the sequence of events for sustainer ignition.

INSTRUMENTATION

NASA 4.20 Aerobee Rocket

Two piezoelectric transducers were installed on the aft side of the attitude control system (ACS) mounting plate at missile station 93 (Figures 10 and 11). One accelerometer was oriented longitudinally, along the thrust axis, the other, laterally along the B-D rail plane of the launch tower. Full deviation for the longitudinal channel was ± 201 g's; for the lateral channel ± 182 g's.

The accelerometer signals were conditioned utilizing "charge" amplifiers and filter-clipper-bias units. The conditioned signals were then fed into an FM/FM telemeter and transmitted to Wallop's ground station via a slot antenna in one of the rocket fins (Figure 5). Standard IRIG FM/FM telemetry was used with the longitudinal accelerations transmitted on channel E (70 kc, ± 15 percent) and the lateral accelerations transmitted on channel 16 (40 kc, $\pm 7\frac{1}{2}$ percent). The frequency response, typical for both channels, is shown in Figure 12.

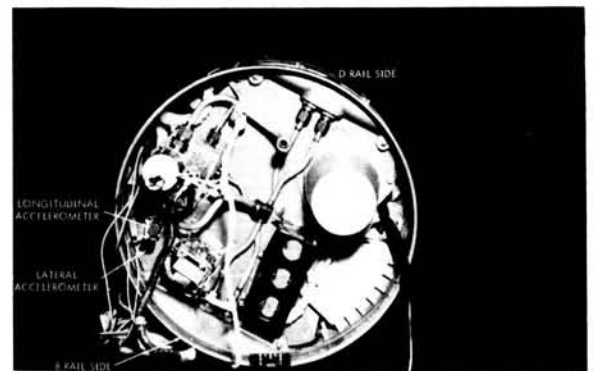


FIGURE 10.—NASA 4.20 ACS, rear view.

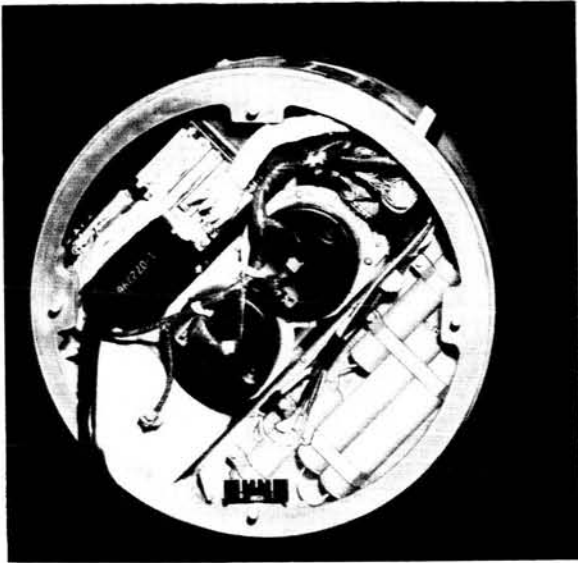


FIGURE 11.—NASA 4.20 ACS, front view.

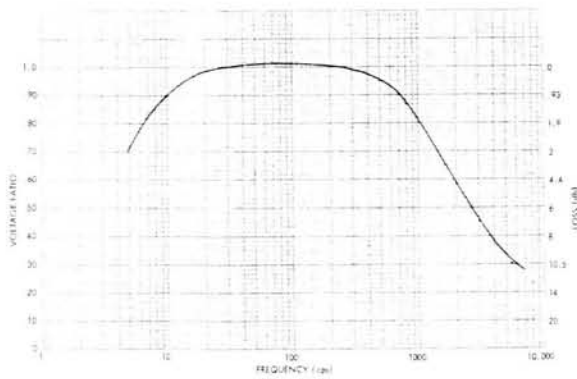


FIGURE 12.—Typical system frequency response for all channels, NASA 4.20, NASA 4.68.

NASA 4.68 Aerobee Rocket

Two piezoelectric transducers were installed on the aft side of the ACS mounting plate at MS 93. The accelerometer mounting block plate is shown in Figure 13. The orientation of the transducers was approximately the same as for NASA 4.20; however, modifications to the ACS obviated exact replication of the mounting block location. A front view of the NASA 4.68 ACS, (Figure 14), is included for comparison of the modifications. In addition, a third transducer was mounted on the base of the instrumentation rack at MS 87 and oriented laterally along the A-C rail plane.



FIGURE 13.—NASA 4.68, ACS rear view.

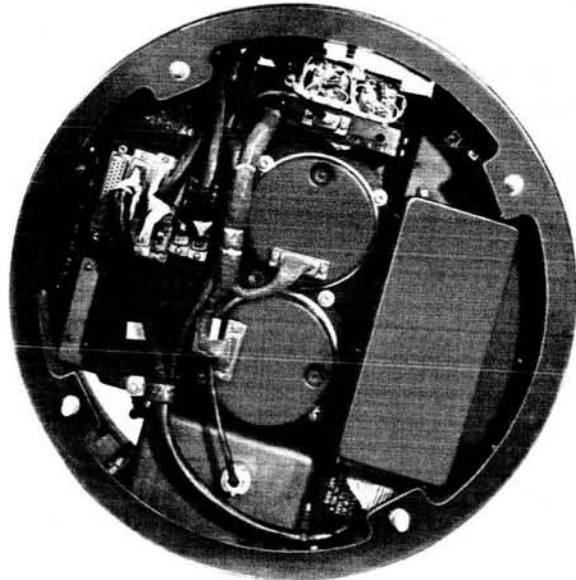


FIGURE 14.—NASA 4.68, ACS, front view.

Full deviation for all three channels was ± 40 g's. IRIG letter channels E, C, and A (± 15 percent deviation) were used in the FM/FM system. Longitudinal axis vibrations (MS 93) were measured on channel C, the lateral axis vibrations (MS 93) were measured on channel E, and the lateral axis vibrations (MS 87) were measured on channel A. The frequency response, typical for these channels is shown in Figure 12.

In addition to the in-flight instrumentation, the tower was instrumented with a microswitch

at the 85-foot level to determine booster exit time from the building; a microphone to determine the airborne noise level inside the building; and accelerometers on rails A (tangential), B and C (radial) at the 85-foot level to determine rail shock and vibration levels. Full deviation for the microphone channel was 167 db; for the accelerometer channels, ± 50 g's. The frequency response of the tower instrumentation was essentially flat to 10 kc. Just prior to the launch, a blue dye was sprayed on the rails for visual indication of lug contact.

DATA REDUCTION

A block diagram of the data reduction system is shown in Figure 15. Although not included in the system, additional equipment was used to inspect the subcarriers for relative amplitude, intermodulation noise level, and frequency drift to insure the validity of the data. Data reduction consisted of rms vibration levels versus time plots, instantaneous acceleration versus time plots and frequency versus time plots. The flight system frequency response was equalized from 0.2 kc to 3.0 kc prior to the data reduction. The low-pass output filters (LPOF) used are noted on the individual displays.

TEST RESULTS

NASA 4.20 Aerobee Rocket

An irregularity occurred when the booster exploded approximately 2.0 seconds after lift-off. The rocket survived the explosion and continued to function except that the roll rate was 3.0 rps instead of the expected 2.0 rps, and that the rocket evidently experienced a resonance due to the explosion.

Because the flight was not considered representative, data analyses were restricted to the time period from lift-off through booster explosion with the exception of the RMS vibration time histories. These are shown for the longitudinal and lateral axes in Figures 16 and 17 respectively. The abscissa (time) is linear; each division on the ordinate (amplitude) represents one db. Vibration transients are indicated at squib firing, lift-off, booster explosion, and rocket motor burnout in both axes. The highest vibration levels occurred between lift-off and hangar exit. During this period the average levels were approximately 7 g's RMS in both axes.

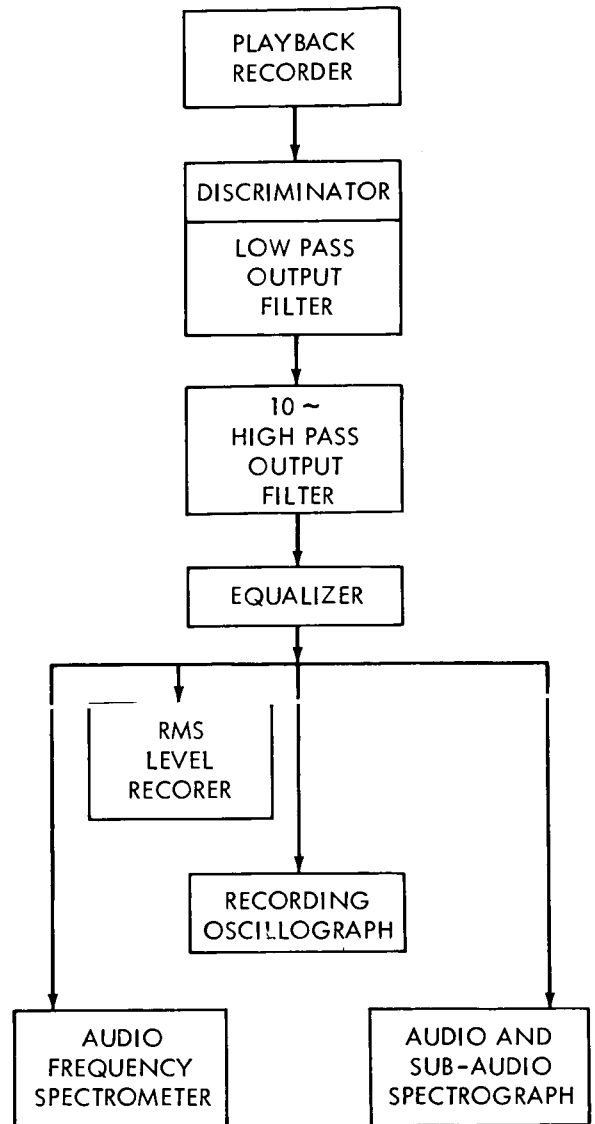


FIGURE 15.—Data reduction system.

Instantaneous acceleration versus time records for the period from T-0.4 seconds to T+2.2 seconds are shown in Figure 18. The display is linear with respect to both acceleration level and time. Maximum amplitudes measured during lift-off and booster explosion were 22 g-pk and 29 g-pk respectively in the longitudinal axis; 24 g-pk and 44 g-pk in the lateral axis.

In view of the short time duration (0.7 seconds) of the high level vibration, quantitative spectral analyses were considered unwarranted. However,

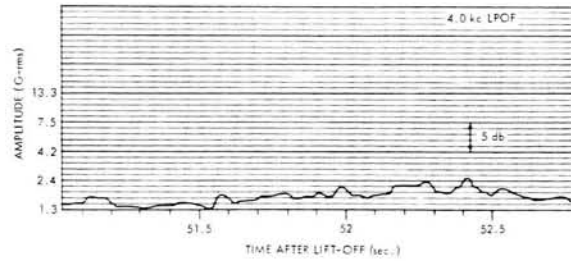
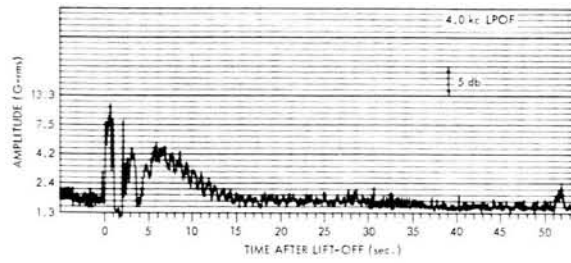
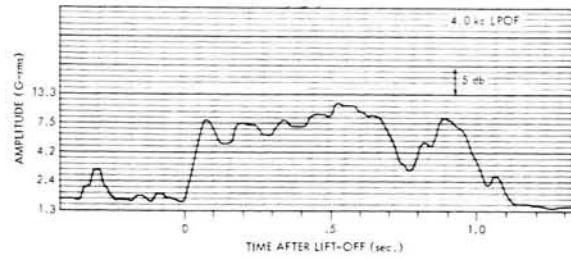
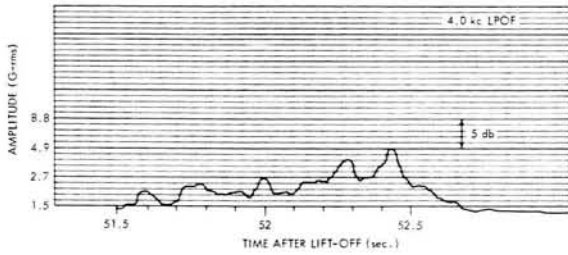
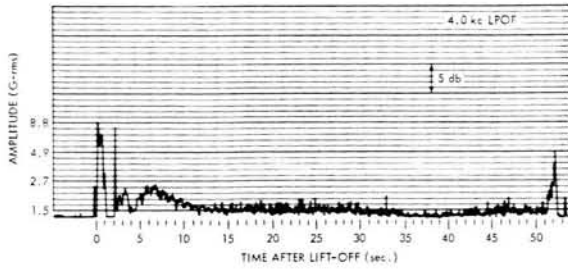
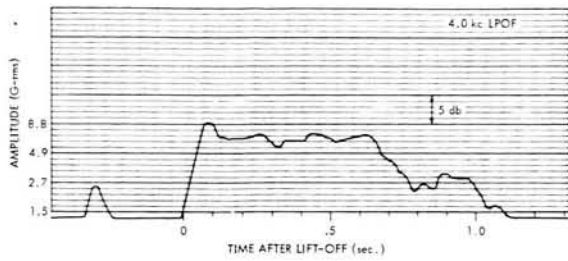


FIGURE 16.—Vibration time history NASA 4.20, ACS longitudinal.

FIGURE 17.—Vibration time history NASA 4.20, ACS lateral.

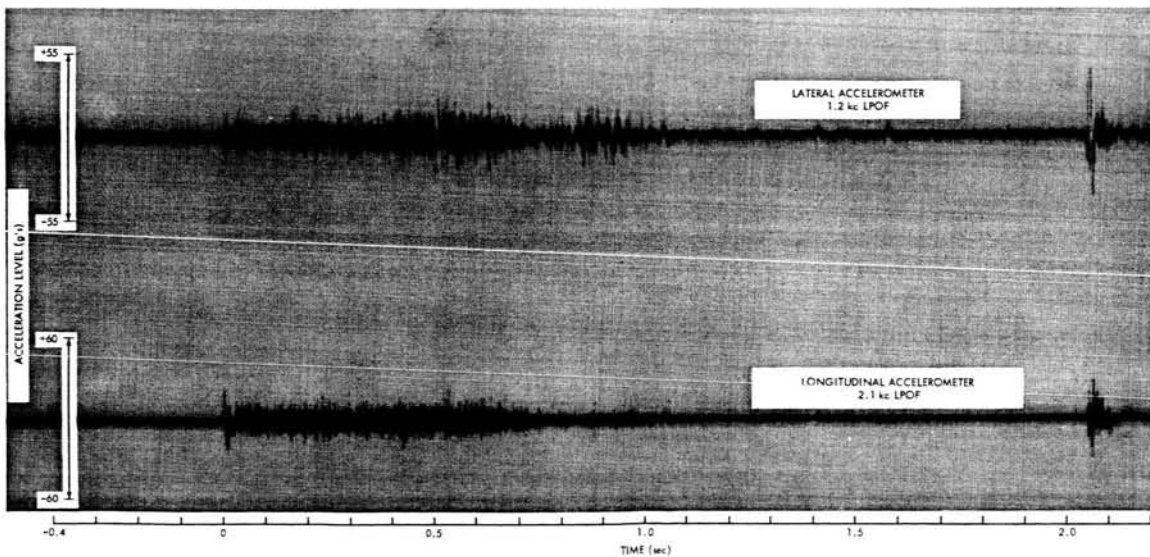


FIGURE 18.—NASA 4.20 vibration at liftoff.

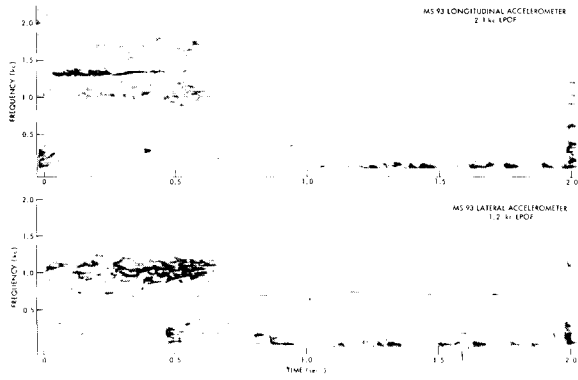


FIGURE 19.—NASA 4.20 launch history.

for indications of the spectra, an audio spectrograph having a 20 cps band pass filter is displayed (Figure 19). The display is a frequency versus time characteristic and is linear with respect to both parameters. The relative vibratory level is indicated by the shading between grey and black; black indicating the higher level. In the longitudinal axis, the bursts during the closed tower portion of flight and during booster explosion appears to be random. The predominate frequency was approximately 1350 cps and occurred during the former. In the lateral axis, the most significant frequencies noted are in the 800 cps—1200 cps range; the upper limit being restricted by the low-pass output filter.

NASA 4.68 Aerobee Rocket

The trajectory and vehicle performance of this flight is considered to be typical for the Aerobee 150A rocket. The RMS vibration time histories for the longitudinal and lateral axes at MS 93 are shown in Figures 20 and 21 respectively. The first recorded event is the closing, by squib firing, of the overboard dump valve. Lift-off follows in approximately 0.3 seconds. During the period between lift-off and hangar exit, the average levels were approximately 7 g's RMS in both axes. The next significant event was a drop in RMS level at approximately +0.7 seconds. The micro-switch installed at the 85 foot level indicated that the exit time was +0.78 seconds. From +0.7 seconds to tower exit at +1.1 seconds the levels drop, with the exception of peaks at hard impact, to very low levels. Comparison of Figures 16 and 17 with Figures 20 and 21 indicate similarity

of vibration levels during the tower portion of both flights.

Proceeding on, the rocket enters the region of maximum dynamic pressure at approximately +19 seconds and the resultant slight increase in level is attributed to buffeting. Levels drop again until motor turbulence is experienced near burnout.

The third accelerometer located at MS 87 malfunctioned at launch. Investigation revealed that it was probably bi-stable and either became inoperative or experienced a drastic change in sensitivity due to shock. The trouble was attributed to an improper manufacturing process.

The overall sound pressure level (uncorrected) versus time measured by the tower microphone is shown in Figure 22. It appears (from comparison of Figures 20 and 21 with Figure 22) that some amount of the rocket vibration (to +0.7 seconds)

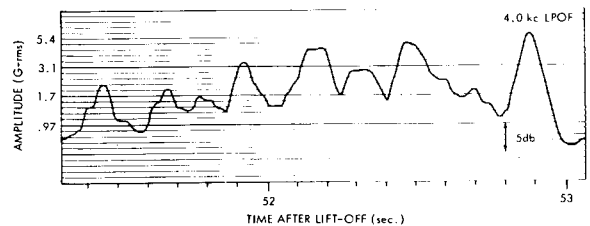
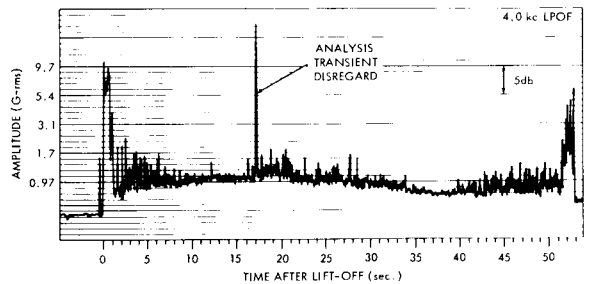
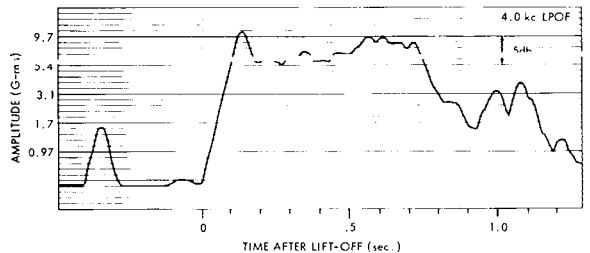


FIGURE 20.—Vibration time history, NASA 4.68, ACS longitudinal.

is due to acoustic coupling. The maximum sound pressure level measured during this time was approximately 146 db.

Instantaneous acceleration versus time plots for the period from -0.6 seconds to +2.2 seconds are presented in Figure 23 for the longitudinal and lateral axes at MS 93 and for the B rail radial accelerometer which was located at the 85-foot level. The sensitive axes of the lateral and B rail accelerometers were approximately parallel. Maximum values measured were 29 g-pk and 24 g-pk for the longitudinal and lateral accelerometers respectively. The B rail experienced accelerations greater than 50 g-pk for approximately 0.3 seconds; the measuring system was overloaded during this time period.

A plot of rail-lug contact as well as hard hits and joint hits is presented in Figure 24 (see also Table 2).

Rail inspection indicates that the rocket rides the rails rather smoothly during the first 88 feet of travel. This light and relatively smooth contact induces a vibratory motion, rather than shock impulse, to the rocket (Figure 23). Rail contacts

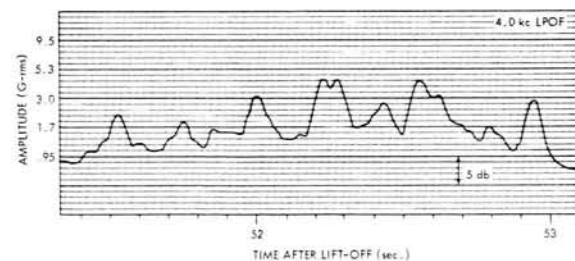
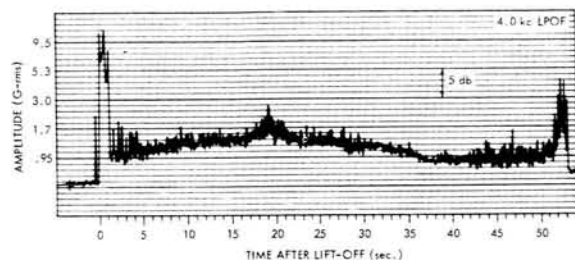
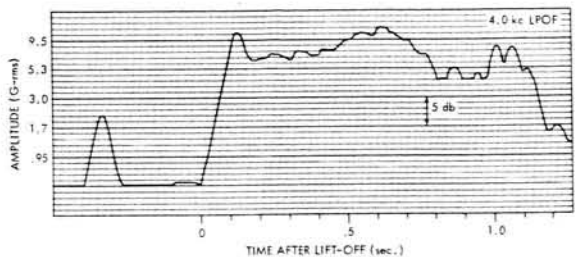


FIGURE 21.—Vibration time history, NASA 4.68, ACS lateral.

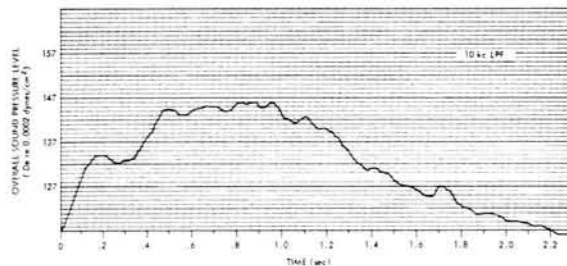


FIGURE 22.—NASA 4.68 airborne sound pressure level versus time; tower microphone 10 kc.

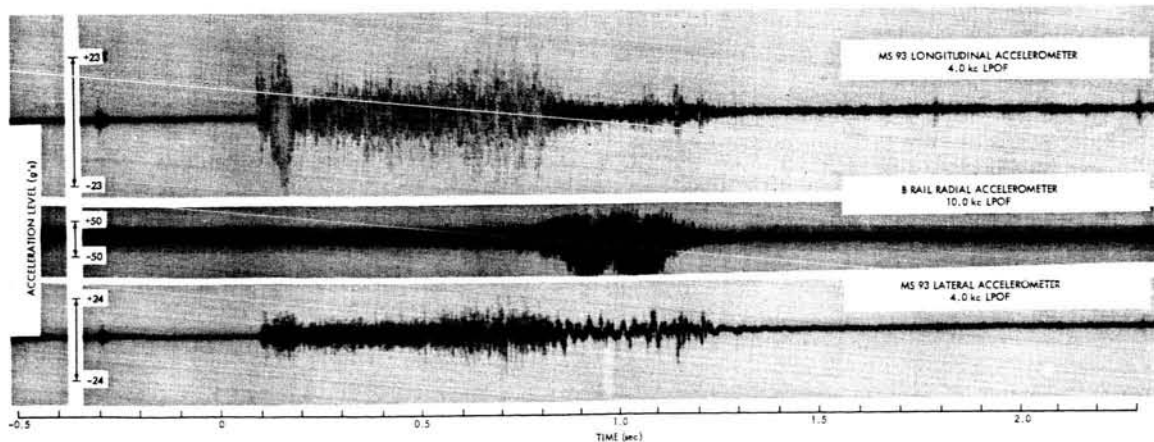


FIGURE 23.—NASA 4.68 vibration at liftoff.

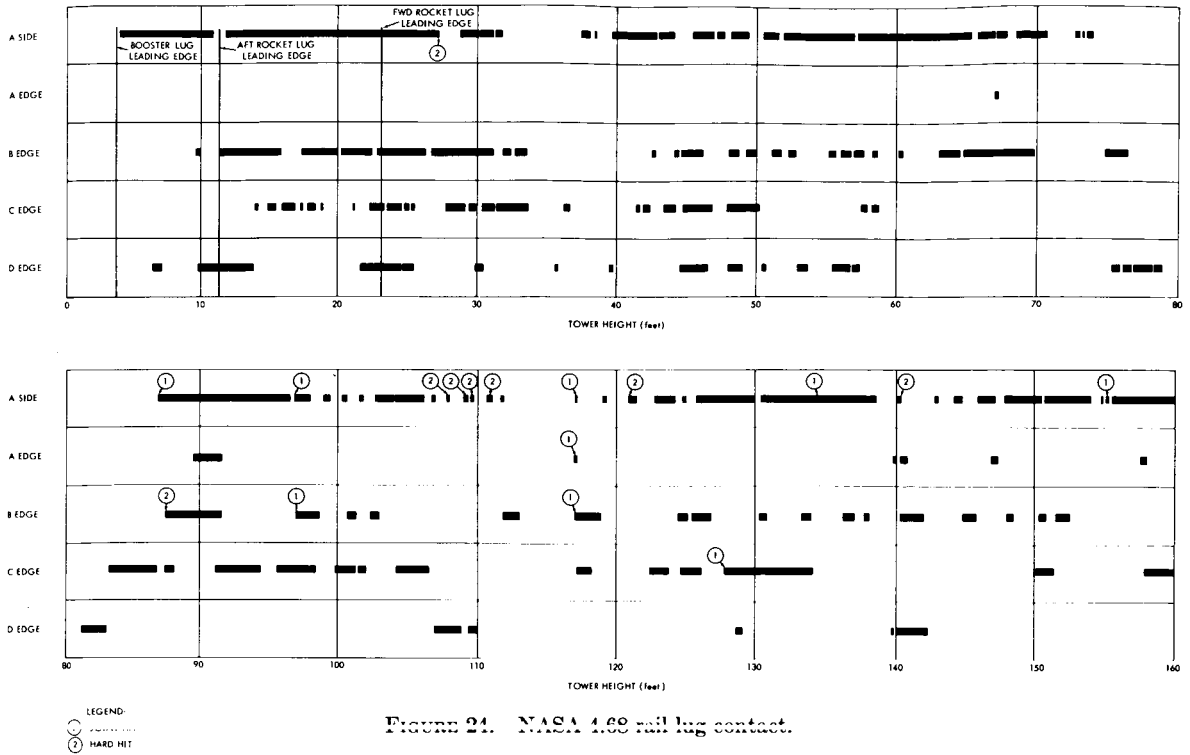


Figure 24. NASA 4.69 rail lug contact.

Table 2
Percentage of Rail-Lug Contact.

Contact	Rail				
	A Side	A Edge	B Edge	C Edge	D Edge
In-hanger	36.9%	0 %	22.6%	13.6%	13.6%
Out-of-hanger	29.8%	2.4%	12.6%	17.8%	3.1%
Total	66.7%	2.4%	35.2%	31.4%	16.7%

inside the hangar were hard to distinguish while contacts outside the hangar were very noticeable and at some points the rails were gouged. As the rocket leaves the building, it appears that its velocity is such that it breaks free of the rails and starts to impact during the open tower portion of flight. The low frequency lateral oscillations during the open tower portion of flight are attributed to the vehicle's response to the rail impacts.

The A rail side contact is due to the booster guide lug restraining spin. Note the small amount of A rail edge contact. At two points on the side of the A rail, welds were not ground flat, thus obstructing the booster guide lug, and were severely gouged. Two rail joints, located at tower heights of 27 feet, 4 inches and 87 feet,

11 inches, were out of line and gouged by the side of the booster guide lug.

Qualitative spectral analyses of the vibration data were performed using an audio spectrograph with a 20 cps bandwidth filter. The gain setting for each channel was held constant to permit comparison of amplitude variation with time. Lift-off, max Q, and rocket-burnout spectra for both MS 93 axes are shown in Figures 25, 26, and 27, respectively.

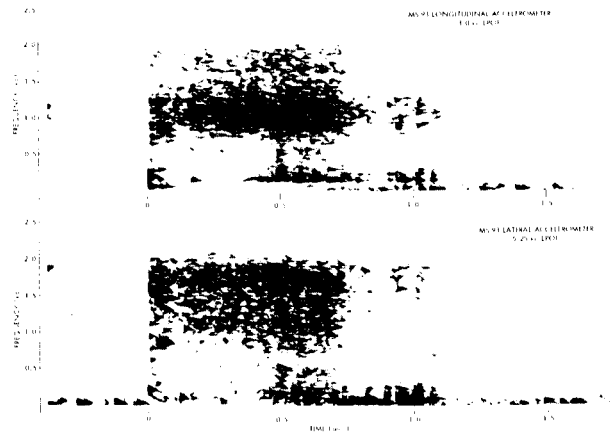


Figure 25.—NASA 4.68 launch history.

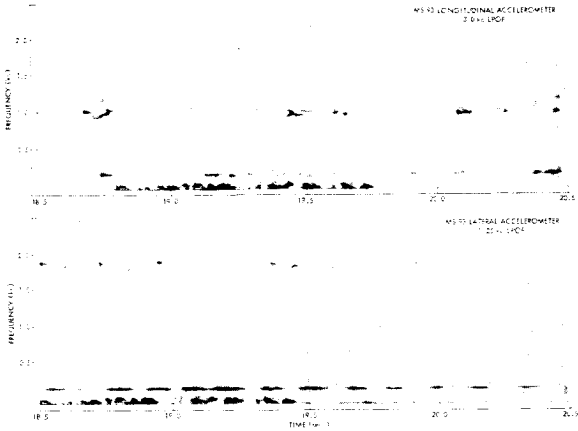


FIGURE 26.—NASA 4.68 max Q history.

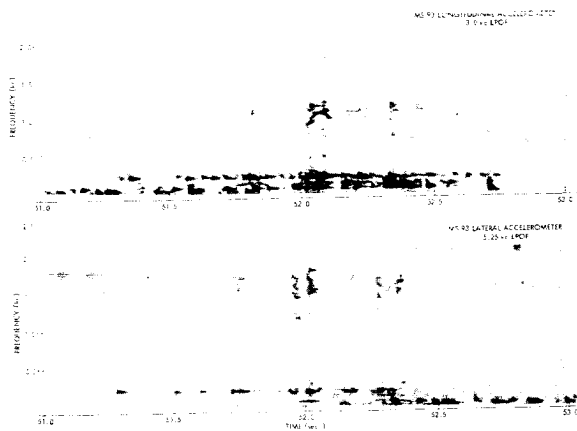


FIGURE 27.—NASA 4.68 rocket burnout history.

In both axes, the vibratory spectrum is relatively broad and intense during the closed-tower portion of flight, followed by a series of shock-type pulsations during the open-tower portion. During the time that the vehicle passes through

the region of max Q, the greatest amplitudes occur in the frequency bands of 175–200 cps and 900–1100 cps in the longitudinal axis; and in the frequency bands of 175–200 cps and 1500–1800 cps in the lateral axis. During rocket motor

Table 3
NASA 4.68 Launch Vibratory Acceleration Longitudinal Axis.
One-Third Octave Band Analysis, 4.0 kc LPOF.

Center Frequency (cps)	Maximum Acceleration (σ -rms) from t to t + 0.1 Seconds																		
	t = 0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	
40	0.6	0.3							0.3										
50	0.8	0.4																	
63	1.1	1.1										0.4							
80	1.6	1.6	0.3				0.6	0.5	0.3			0.5	0.4						
100	0.7	0.4			0.4	0.5	0.5	0.5	0.4	0.5	0.5								
125	0.4	0.4				0.7	0.6	0.5	0.5	0.6	1.0	0.6							
160	1.1	0.9	1.3	0.6	1.6	2.3	1.4	1.4	1.1	1.1	2.2	2.2	1.2						
200	1.6	1.6	1.0	0.4	2.4	2.5	1.9	1.9	1.2	0.8	1.5	1.1	0.5	0.3	0.3		0.4	0.4	
250	1.3	1.1	0.4	0.6	0.6	2.3	1.7	1.0	1.4	1.2	1.2	1.6	0.4	0.5	0.7	0.7	0.7	0.6	
315	1.5	1.1	0.7	0.8	1.2	2.6	1.3	1.2	0.5	0.7	1.1	1.2	0.5	0.4	0.4	0.5	0.4	0.5	
400	0.8	0.6	0.6	0.5	0.8	1.4	0.8	0.7	0.6	0.4	0.4	0.5							
500	0.9	0.8	1.0	0.8	1.0	1.6	1.8	0.8	0.6	0.4	0.4								
630	1.5	1.5	1.1	1.1	1.3	1.8	1.4	1.1	0.8	0.5	0.7	0.7							
800	1.8	5.0	2.1	3.3	2.7	5.4	4.6	3.5	1.2	1.0	1.5	1.1	0.4	0.4	0.4	0.4	0.4	0.5	
1000	1.8	5.2	7.1	7.1	6.9	8.7	9.4	8.4	2.4	2.2	2.2	2.1	0.6	0.4	0.4	0.4	0.6	0.6	
1250	1.8	4.4	4.5	5.0	4.3	6.8	6.2	6.6	2.1	1.8	2.0	2.0	0.6	0.4	0.4	0.4	0.4	0.5	
1600	1.9	2.8	3.2	2.8	3.5	3.5	4.4	3.2	1.2	1.0	1.0	1.0	0.3					0.6	
2000	2.2	2.5	3.1	2.7	3.0	2.8	3.5	2.9	1.3	1.1	1.3	0.9	0.6						
2500	1.4	1.4	1.9	1.7	2.1	1.6	2.2	1.2	0.8	0.6	0.5	0.4	0.5						
3150	1.1	1.1	1.5	1.4	1.8	1.1	1.5	0.9	0.6	0.6	0.6	0.6	0.8	0.3			0.4	1.0	
4000	0.6	0.6	0.8	0.7	1.1	0.7	0.8	0.5	0.4	0.4	0.4	0.5	0.6					0.7	

burnout, the vehicle again experiences a series of shock-type pulsations with the predominant frequencies less than 300 cps in both axes.

One-third octave analyses of the MS 93 accelerometer data and the tower microphone data were then performed. The maximum levels occur-

ring within 0.1 second intervals were measured; the values which were above the ambient noise level are presented in Tables 3, 4, and 5. Photographs of the three-dimensional models constructed from the resulting records are shown in Figures 28 through 33.

Table 4
NASA 4. 68 Launch Vibratory Acceleration Lateral Axis.
One-Third Octave Band Analysis, 4.0 kc LPOF.

Center Frequency (cps)	Maximum Acceleration (g-rms) from t to t + 0.1 Seconds																	
	t = 0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7
40	0.7	0.4	1.2	1.0	1.1	1.1	1.7	4.7	3.8	4.8	4.8	0.7	0.5	0.4				
50	0.6	0.5	0.8	0.6	1.0	1.4	2.1	1.6	1.2	1.7	1.9	0.9						
63	0.5	0.6	0.5	0.5	0.6	2.0	2.8	1.0	2.3	1.2	4.5	3.8	0.7	0.6	0.5	0.5	0.4	0.3
80	0.9	1.1	1.1	1.2	0.7	1.9	4.7	1.7	1.7	1.5	4.6	3.4	0.7	0.4	0.5	0.6	0.6	
100	1.1	1.5	0.8	0.8	1.0	2.5	2.6	1.3	2.9	2.2	1.5	3.4	1.1	0.4	0.3	0.3	0.5	0.8
125	0.5	1.5	1.0	0.6	1.0	1.7	1.5	1.3	1.5	2.2	2.6	1.7	1.1					
160	1.0	1.2	1.1	0.9	1.2	3.0	3.8	7.4	1.0	1.7	3.2	3.2	0.4					
200	1.2	1.0	0.7	0.9	2.3	2.7	2.8	1.7	1.7	1.5	3.4	3.4						
250	1.0	0.7	0.5	0.9	1.0	2.4	2.4	1.5	1.3	2.0	2.7	3.0	0.3	0.4			0.5	0.5
315	1.2	1.0	1.0	0.9	1.3	2.1	1.7	1.8	0.8	1.2	1.2	1.0	0.3	0.7	0.5	0.5	0.6	0.6
400	1.0	1.0	1.3	1.0	1.0	2.6	2.0	1.3	1.4	1.3	1.0	0.6						
500	1.1	1.0	1.0	1.2	1.2	3.4	3.1	1.7	0.7	0.7	0.9	0.5						
630	1.2	1.2	1.2	1.3	1.5	1.8	2.1	1.1	0.8	0.4	1.1	1.4	0.4					
800	1.9	4.8	1.7	2.0	3.0	3.0	3.8	2.2	1.0	0.8	1.0	0.8	0.3					
1000	3.0	3.3	3.0	3.2	3.4	5.3	4.8	4.2	1.1	1.2	1.2	0.8						
1250	1.7	4.2	4.2	4.8	4.8	7.2	6.5	6.5	2.1	1.7	1.7	1.4	0.4	0.4	0.3	0.3	0.3	0.4
1600	3.0	6.8	6.6	7.8	9.1	9.1	11.7	7.9	3.3	3.5	3.6	2.5	0.6	0.6	0.7	0.6	0.6	0.7
2000	2.1	5.0	5.4	5.4	6.5	7.2	7.2	5.1	2.9	1.7	2.6	2.0	0.5	0.4	0.4	0.4	0.6	0.7
2500	1.6	1.6	1.1	1.3	1.3	1.7	2.2	1.3	1.0	0.8	0.7	0.7	0.6	0.4	0.4	0.3	0.6	0.9
3150	1.2	1.2	1.2	1.3	1.6	1.2	1.8	1.2	1.1	0.9	0.9	0.6	0.6	0.6	0.4	0.4	0.6	1.0
4000	0.6	0.7	0.6	0.6	1.3	0.7	0.9	0.5	0.8	0.6	0.5	0.7	0.6	0.4	0.4	0.4	0.5	0.9

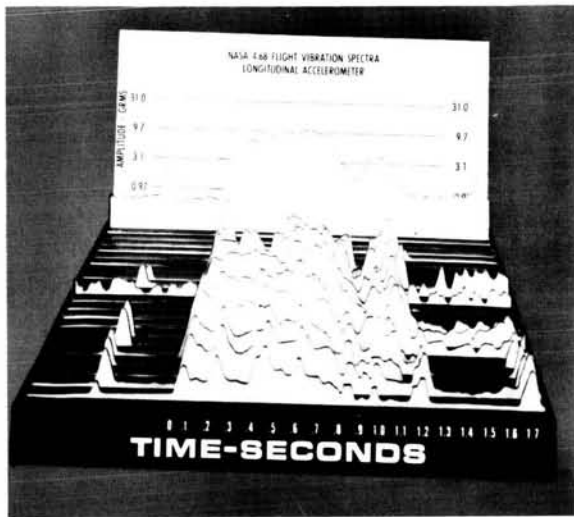


FIGURE 28.—Front view of the three-dimensional plot of the NASA 4.68 overall acceleration and one-third octave analysis—longitudinal accelerometer 4.0 kc LPOF.

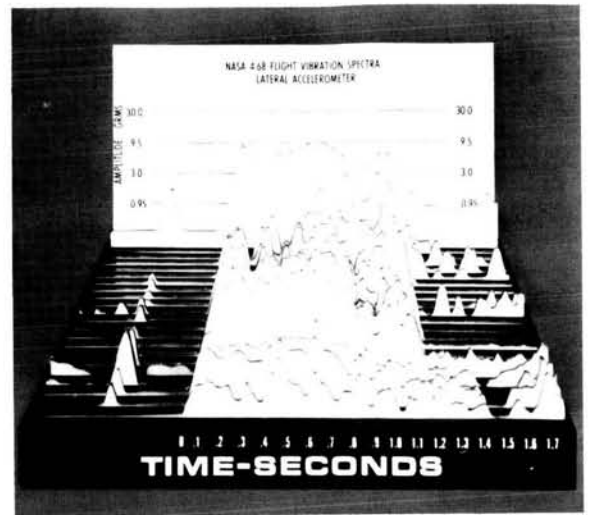


FIGURE 30.—Front view of the three-dimensional plot of the NASA 4.68 overall acceleration and one-third octave analysis—lateral accelerometer 4.0 kc LPOF.

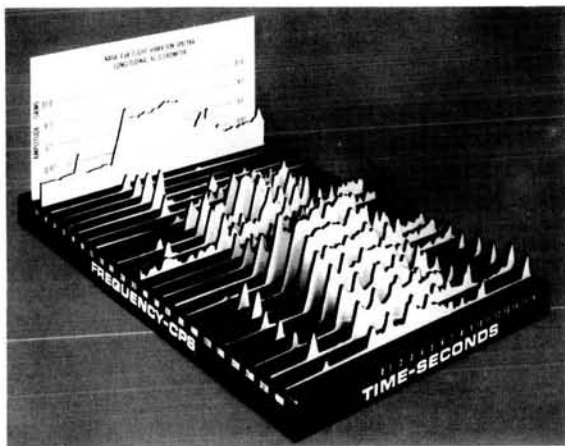


FIGURE 29.—Left oblique view of the three-dimensional plot of NASA 4.68 overall acceleration and one-third octave analysis—longitudinal accelerometer 4.0 kc LPOF.

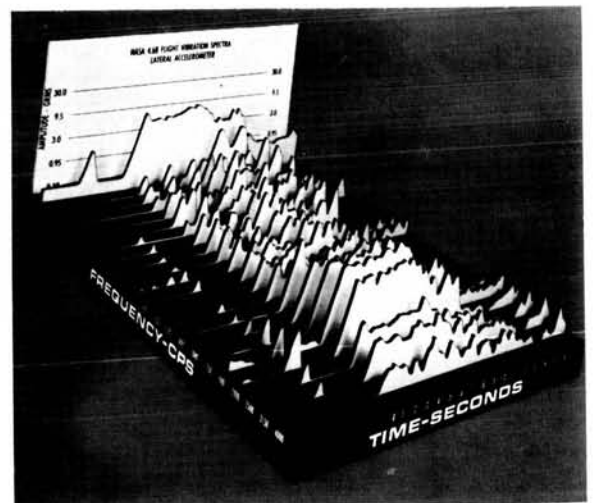


FIGURE 31.—Left oblique view of the three-dimensional plot of NASA 4.68 overall acceleration level and one-third octave analysis—lateral accelerometer 4.0 kc LPOF.



FIGURE 32.—Front view of the three-dimensional plot of NASA 4.68 overall noise level and one-third octave analysis—tower microphone 10.0 kc LPOF.

CONCLUSIONS AND RECOMMENDATIONS

Similar vibration data have been obtained from two flights of Aerobee 150A 4-fin rockets. The maximum levels occurred during the in-tower portion of flight and are considerably lower than those reported for the 3-fin Aerobeets.

The following recommendations are made, based on visual examination of the tower rails:

1. All rail joints should be inspected for misalignment prior to flight to avoid joint interference.
2. If it is desirable to further control the vibration during tower flight or if newer vehicles are designed for this tower, it is suggested

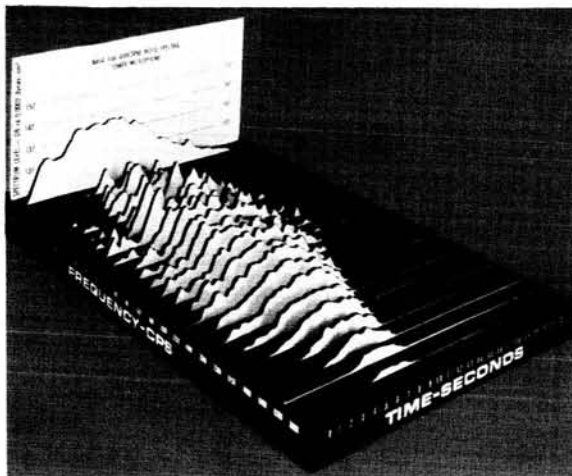


FIGURE 33.—Left oblique view of the three-dimensional plot of NASA 4.68 overall noise level and one-third octave analysis—tower microphone.

that a thorough analysis of launch tower dynamics be made. In view of the deep gouges in the rails, some attention should be given to the design of the riding lugs and their supporting structure and the fore and aft location and number of lugs.

BIBLIOGRAPHY

- "Aerobee—150A Sounding Rocket Model AJ 60-13," Aerojet-General Corporation, Technical Manual Service Instructions, Publication No. HS-520104, June 1960.
- NEWELL, HOMER E., JR., "Sounding Rockets," New York, McGraw-Hill, 1959, p. 89.
- RUSSELL, W. A., JR., "NASA 4.20 Post Flight Summary," Goddard Space Flight Center.
- RUSSELL, W. A., JR., "NASA 4.68 Post Flight Summary," Goddard Space Flight Center, January 26, 1962.

PROPOSED MECHANISM FOR THERMOPHOTOTROPIC BEHAVIOR IN PEROVSKITE-STRUCTURED TITANATES*

STANLEY M. NEUDER

Many compounds in the solid or liquid state have been known to change color when exposed to electromagnetic radiation. Thermophototropism is the term used to describe the phenomena of radiation-induced coloration and thermally stimulated reversion to the original state. As a result of experimental investigations by the author, a mechanism is proposed to account for thermophototropic behavior in perovskite-structured alkaline-earth titanates. A brief review of these experimental results and of several theories reported in the literature is included.

INTRODUCTION

The fundamental optical absorption edges for electronic transitions in the titanates of barium, calcium, and strontium at room temperature occur at about 3950Å, 3650Å, and 3850Å, respectively (References 1-3). Figure 1 illustrates the

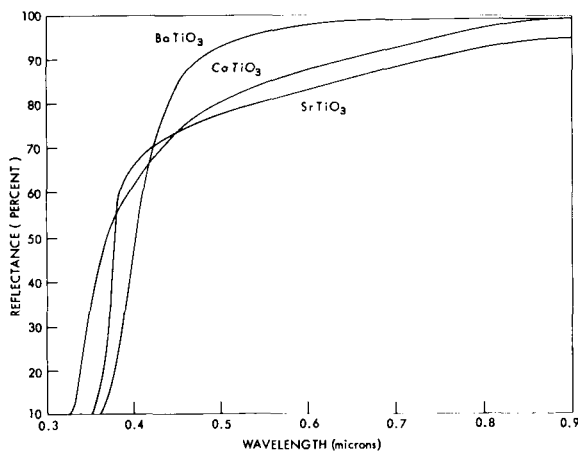


FIGURE 1.—Reflectance as a function of wavelength for BaTiO₃, CaTiO₃, and SrTiO₃ in arbitrary units.

room temperature spectral reflectance in arbitrary units measured with a spectrophotometer by using pressed powder ceramic titanate samples fired in air. The fundamental absorption of titanium dioxide occurs in the same region of

wavelengths (References 4 and 5). Figure 2 is the spectral reflectance of titanium dioxide measured under the same conditions. The photon energies corresponding to the electronic energy gap of these materials lie near the value of $h\nu$ of the order of 3.3 eV.

It has been shown that when pure ceramic BaTiO₃, SrTiO₃, CaTiO₃, and TiO₂ are exposed to ultraviolet radiation at room temperature, no phototropic reaction occurs (References 5-13). When impure (or doped) specimens are irradiated under the same conditions, a noticeable color change ensues. However, irradiation of doped samples at room temperature with filtered light

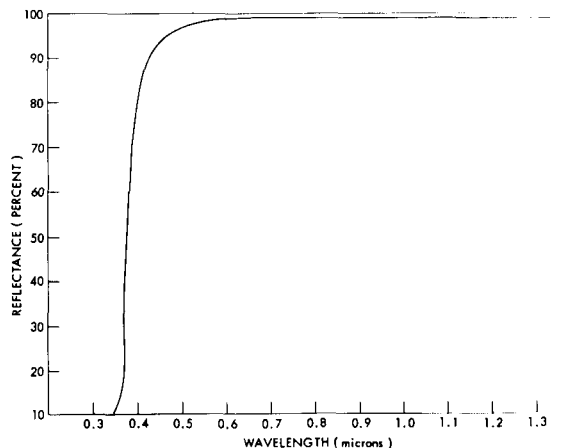


FIGURE 2.—Reflectance as a function of wavelength for TiO₂ in arbitrary units.

*Published as NASA Technical Note D-2258, March 1964.

of $\lambda \geq 5300\text{\AA}$ does not induce phototropy. Figures 3-6 are typical results obtained under these conditions. The upper curves represent the spectral reflectance of the bleached specimens. (A bleached sample is one which has been heated to about 250°C in air or vacuum and is white in appearance.) The lower curves represent the reflectance obtained after these samples were irradiated by a mercury light source. It can be seen that the relatively pure material shows slight phototropy due to the residual amounts of impurities present. The reflectance loss in the visible region of the spectrum after irradiation indicates the degree of coloration.

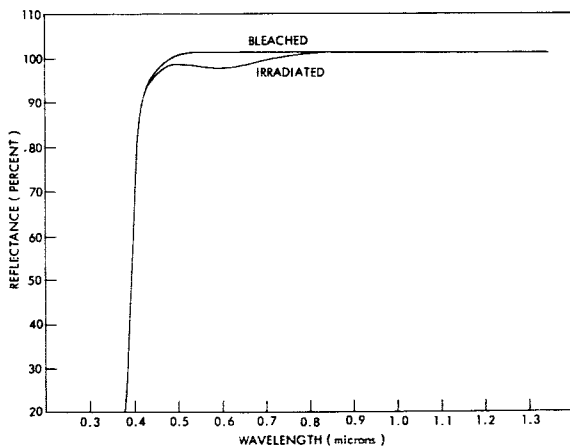


FIGURE 3.—Reflectance as a function of wavelength for bleached and irradiated, chemically pure BaTiO_3 .

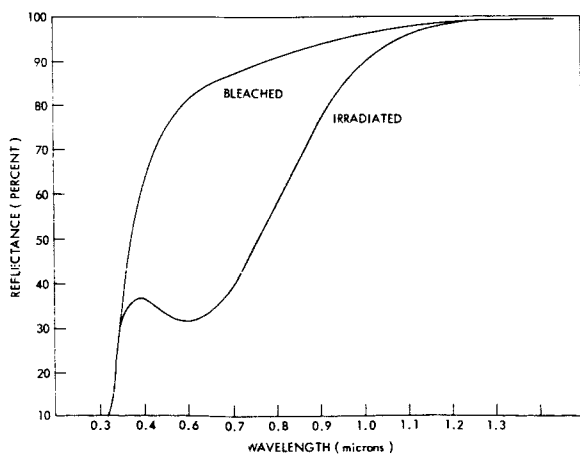


FIGURE 4.—Reflectance as a function of wavelength for bleached and irradiated, relatively impure CaTiO_3 , in arbitrary units.

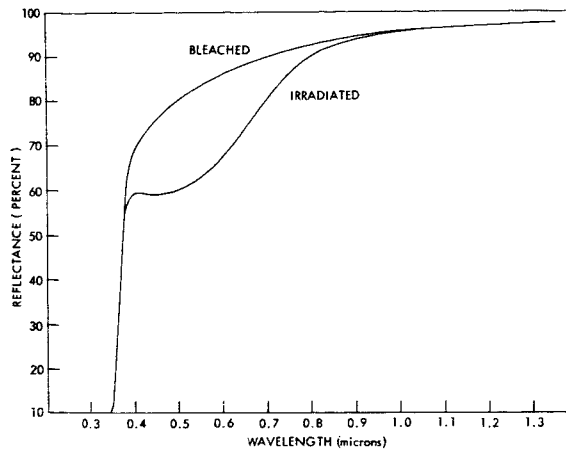


FIGURE 5.—Reflectance as a function of wavelength for bleached and irradiated relatively impure SrTiO_3 , in arbitrary units.

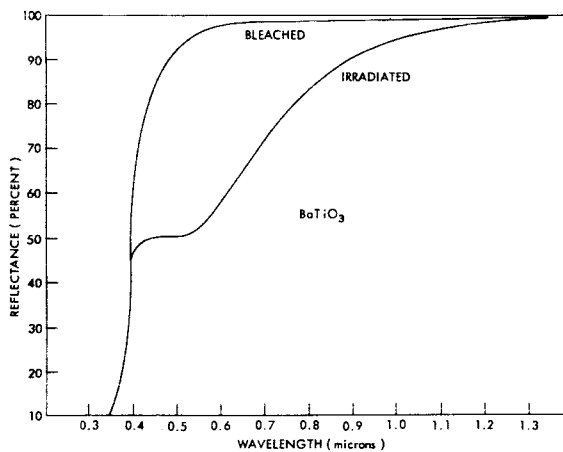


FIGURE 6.—Reflectance as a function of wavelength for bleached and irradiated, relatively impure BaTiO_3 , in arbitrary units.

Typical impurities and their concentrations are given in Table 1 for both pure and impure barium titanate. The impurities may enter into the lattice substantially. Ti^{4+} can be replaced (References 14-16) by the transition elements (characterized by an incomplete d shell) such as Zr^{4+} and Nb^{5+} , or by ions of similar size, such as Al^{3+} . Other impurities such as Sr^{2+} or Ca^{2+} can be substituted for Ba^{2+} (References 17 and 18).

REVIEW OF EXPERIMENTAL RESULTS

The crystalline structures of barium, calcium, and strontium titanate are of the perovskite type,

Table 1

Impurities and Their Concentrations
(by Weight Percentage).*

Constituents	Chemically Pure BaTiO ₃ (lot 302)	BaTiO ₃ (lot 58C)
B ₂ O ₃	< 0.005	< 0.01
SiO ₂	0.005	0.45
P ₂ O ₅	< 0.3	0.17
ZrO ₂	0.03	0.05
Sb ₂ O ₃	< 0.01	0.01
HfO ₂	< 0.02	< 0.02
Al ₂ O ₃	0.03	0.27
Fe ₂ O ₃	0.002	0.01
MnO ₂	< 0.0005	0.005
PbO	< 0.005	< 0.005
MgO	0.005	0.01
SnO ₂	< 0.005	0.005
Nb ₂ O ₅	< 0.01	< 0.01
WO ₃	< 0.05	< 0.05
Cr ₂ O ₃	< 0.002	< 0.002
NiO	< 0.005	< 0.005
TiO ₂	34.42	33.55
MoO ₃	< 0.01	< 0.01
CaO	0.03	0.09
V ₂ O ₅	< 0.003	< 0.01
CuO	0.0001	0.0005
ZnO	< 0.2	< 0.2
BaO	65.35	63.94
SrO	0.05	0.42
Na ₂ O	0.005	0.22
Li ₂ O	< 0.005	< 0.005
K ₂ O	< 0.01	0.03
SO ₃	-	0.21
CO ₂	-	0.47

*Material and chemical analysis supplied by the National Lead Company, Titanium Alloy Manufacturing Division, New York City.

oxygen forming regular octahedra about the titanium ion (Reference 19). These titanates have the formula A²⁺ B⁴⁺O₃, where the quadrivalent cation is surrounded by the octahedra and the divalent cation occupies the large holes between the octahedra. A somewhat similar octahedral structure exists in rutile (TiO₂) (References 5 and 20). These compounds exhibit thermophototropic behavior. In contrast, neither pure nor doped magnesium titanate which crystallizes in the cubic close-packed ilmenite structure (Reference 21) changes color upon irradiation (Reference 7). The ilmenite structure does not have the Ti-O chains present in the perovskite form. Although each Ti atom is surrounded by six O atoms, only the oxygens form a close-packed configuration, the Mg and Ti ions being in the interstices. CaTiO₃ has a structure in which Ca and O form a close-packed cubic lattice.

In previous experiments at the Goddard Space Flight Center the behavior of these phototropes have been investigated for various values of temperature, vacuum and gaseous ambients, and selective spectral irradiation. The results are reviewed briefly here. The material is obtained in powered form die-pressed at about 10,000 psi into disc-shaped pellets, air-fired at 1100°C for about 1 hour and allowed to cool slowly for several hours. The samples are exposed to the radiation of a mercury lamp or carbon arc. Spectroreflectance measurements are made before and after irradiation. The reflectometer range extends from 0.2 to 2.4 microns.* The pure and impure samples behave as previously discussed when irradiated in air. After firing or bleaching, a specimen appears white. Following irradiation in air it appears pink-violet or red-brown. At room temperature this darkened state persists after irradiation, essentially unchanged in color. The process of coloration is thermally reversible in that heating the darkened sample will induce a reversion to the original bleached state. The color change saturates after approximately 30 minutes of irradiation and no coloration penetrates below the surface, even after many hours of irradiation. Irradiation with photons of energy

*Spectroreflectance data have not been reproduced here beyond 13,500Å, since the phototropic reaction does not alter the characteristic reflectivity in the longer wavelength region at room temperature.

appreciably less than that of the band gap will not induce phototropy.

Irradiating an air-bleached sample in a vacuum of the order of 10^{-5} mm of Hg causes a 55 percent reduction in the phototropic response at 5000Å compared with that exhibited in air. Irradiating a vacuum-bleached or air-bleached sample in air or oxygen at 1 atm yields a relatively large or the maximum color change (depicted in Figures 4-6). Vacuum heating prior to room temperature vacuum irradiation (with no intermediate exposure to air) almost, if not completely, eliminates the phototropic response. The temperature at which this occurs is greatest for CaTiO_3 . Irradiating a vacuum-bleached sample in 1 atm of nitrogen at room temperature yields a 60 percent reduction in the phototropic response at 5000Å compared with that in air. Under similar conditions in an argon atmosphere the phototropic response at the same wavelength is reduced by 77 percent. (Both N_2 and A contain residual O_2 gas.) Irradiating a vacuum-bleached sample in oxygen at 10^{-2} mm of Hg (equivalent to a partial pressure of oxygen between that in nitrogen and that in argon at 1 atm) results in a response value intermediate between that for nitrogen and argon (see Table 2).

Specimens irradiated at progressively elevated temperatures either in vacuum or in air exhibit progressively reduced phototropic responses. An upper temperature exists for all the compounds, above which darkening cannot occur. Irradiating a specimen in air at lowered temperatures, down to that of liquid nitrogen, will cause rapid darkening much like that at room temperature. However, the color quickly bleaches out as the sample returns to room temperature. Cooling a darkened sample does not alter its state. Table 2 summarizes these results.

X-ray analysis of BaTiO_3 at room temperature showed no detectable phase change between the darkened and bleached states. No phosphorescence radiation was detectable from a darkened sample at room temperature.

DISCUSSION OF THE PROPOSED MECHANISM

It is apparent that thermophototropism (TPT) is manifested in the alkaline-earth titanates when three major conditions prevail: (1) a perovskite-type structure exists, oxygen forming octahedral

bonds about the titanium atom. The ionic radii of Sr, Ca, and Ba differ considerably (Reference 22), but this is relatively insignificant in regard to the existence of TPT in a perovskite-structured environment.* (2) Appreciable concentrations of certain types of impurities are present, as is readily seen from a comparison of Figures 3 and 4-6. (3) An oxygen-containing gaseous ambient is present. The responses observed from thermal-vacuum studies, together with the different gaseous ambients (and associated oxygen and water vapor content) strongly indicate that TPT is primarily an oxygen-dependent surface phenomenon.

The suggested mechanism is as follows: Ceramic oxides are known to be oxygen deficient (References 5 and 23-25) with a high density of surface vacancies. Further, surface sorption and desorption is readily accomplished (Reference 26). When the material is exposed to an air ambient, oxygen atoms adhere to the surface and form Ti-O bonds. Subsequent ultraviolet irradiation of photon energy approximately equal to or greater than the energy gap photoionizes electrons from the Ti-O bonds. Except for compounds with high concentrations of trivalent impurities, non-stoichiometric material with an excess of titanium behaves as an n-type semiconductor, the impurities acting as trapping centers for the conduction electrons from the Ti-O source at the surface. The impurity levels lie below and close to the conduction band in the forbidden energy gap. The reduction of impurities to less stable valence states causes coloration and a non-discrete or broad-band absorption is exhibited, characteristic of interacting wave functions in solids (Reference 23). Reduction of impurities increases the electrical conductivity appreciably (References 27-30), and increases over several orders of magnitude have been reported with sample darkening (Reference 31). It is likely that the conductivity is essentially electronic in nature (References 27-31)† the oxygen ions remaining at the surface in the role of recombination centers. This point will be discussed subsequently in conjunction with thermoluminescent behavior. It follows from the results of thermal-

*A large variation of ionic size will degrade the symmetry, producing non-perovskite structure. This case is excluded from the discussion.

†Also see the abstracts by M. Mashkovitch from the All-Union Conference on the Physics of Dielectrics, 1956.

Table 2

Summary of Experimental Results

Experimental Conditions for BaTiO ₃ ^(a)	Reflectance R or Change of Reflectance ΔR at 5000A
Impure or pure sample fired or bleached in air.	Maximum R
Impure or pure sample bleached in vacuum of 9×10^{-5} mm Hg at approximately 180°C.	Maximum R
Impure sample bleached in air, then exposed ^(b) in air at room temperature (r. t.).	$\Delta R \sim - 43$ percent ^{(c) (d)}
Impure sample bleached in air, then exposed in vacuum at r. t.	$\Delta R \sim - 19$ percent
Impure sample bleached in vacuum, then exposed in vacuum at r. t.	$\Delta R \sim$ negligible
Impure sample bleached in vacuum, then exposed in 1 atm O ₂ or air r. t.	$\Delta R \sim - 39$ percent
Impure sample bleached in vacuum, then exposed in 1 atm N ₂ at r. t.	$\Delta R \sim - 17$ percent
Impure sample bleached in vacuum, then exposed in 1 atm argon at r. t.	$\Delta R \sim - 9$ percent
Impure sample bleached in vacuum, then exposed in O ₂ at 0.02 mm Hg at r. t.	$\Delta R \sim - 12$ percent
Darkened sample bleached in air or vacuum.	$\Delta R \sim + 41$ percent ^(e)
Darkened sample left unexposed for several days at r. t.	$\Delta R \sim + 1$ percent
Bleached sample exposed to radiation of $\lambda \gtrsim 5200A$.	$\Delta R = 0$
Bleached impure sample exposed to radiation of 3000-4200A.	$\Delta R \sim - 30$ percent ^(f)
Impure sample bleached, then exposed in air at 80°C. ^(g)	$\Delta R \sim - 41$ percent
Impure sample bleached, then exposed in air at 140°C. ^(g)	$\Delta R \sim - 36$ percent
Impure sample bleached, then exposed in air at 250°C. ^(g)	$\Delta R \sim - 24$ percent
Impure sample bleached, then exposed near liquid nitrogen temperature.	$\Delta R \gtrsim - 30$ percent
Relatively pure sample bleached, then exposed in air at r. t.	$\Delta R \sim - 3$ percent
Relatively pure sample bleached, then exposed in air near liquid N ₂ temperature.	$\Delta R \gtrsim - 15$ percent

(a) Similar results are true for CaTiO₃ and SrTiO₃, differing only in degree.

(b) All exposures are for 30 minutes for a low pressure mercury source.

(c) Forty-three percent is about the maximum change found for BaTiO₃ samples with the previously mentioned impurities.

(d) The negative sign means that the initial reflectance $>$ final reflectance and the sample becomes darkened (colored). Thus in item 3 above R_i is of the order of 0.93 and R_f is of the order of 0.50, so that $\Delta R = R_f - R_i$ is of the order of -43 percent. A positive ΔR means that the sample bleached.

(e) Vacuum-bleached samples are sometimes a bit whiter than air-bleached samples. No distinction is made in this report.

(f) The reduced value of ΔR , compared with item 3, is due to the lower intensity of irradiation caused by filter attenuation.

(g) Similar results are true for exposures in vacuum at elevated temperatures, differing only in degree.

vacuum studies that the source of the migrating charge must lie at the surface. An air-bleached or vacuum-bleached sample irradiated in air has a maximum amount of ionizable surface bonds available and, therefore, exhibits the maximum response. Placing an air-bleached sample in

vacuum will desorb a substantial* amount of sur-

*Strictly speaking, this is true primarily for BaTiO₃ and SrTiO₃. Because of the greater energy gap associated with CaTiO₃, the thermophototropic response in vacuum at room temperature is only slightly altered. Thus, as expected, a higher temperature was necessary, in both vacuum and air, to fully bleach CaTiO₃. The larger gap energy in CaTiO₃ is presumably due to the shorter Ti-O bond lengths in the octahedra, the larger force constants yielding stronger electronic bonds.

face oxygen, with the result that vacuum irradiation yields a reduced phototropic response. Vacuum bleaching will accelerate the surface desorption to the extent that subsequent vacuum irradiation will not produce coloration. The entire process can be represented by an energy band scheme. Depending on the particular divalent cation, the fundamental energy gap of these titanates at room temperature is about 3.3 eV. Since TiO_2 exhibits the same absorption limit it seems plausible that the gap energy corresponds to the first electronic ionization potential of the Ti-O bond. Except for impurities, trapping mechanisms such as structural defects, including oxygen vacancies, play a relatively insignificant role in the coloration process itself, at least at or above room temperature. It has been found that abrading the surface reduces the phototropic response by introducing a greater density of surface traps. Bleaching a darkened sample thermally ionizes the trap levels, causing the electrons to fall back to the valence band predominantly by means of recombination at the surface.

From this model it follows that radiative electronic transitions must occur and recombination radiation should be observed. Coloration produced at room temperature remains relatively stable as long as the sample remains at that temperature. It is only during a rapid bleaching process that recombination radiation can be detected.

Theoretically no low temperature bound for phototropic activity is expected. Thus samples were darkened at liquid nitrogen temperatures and bleached by being allowed to return rapidly to room temperature. During this process thermoluminescence was observed, corresponding to the emptying of shallow trapping levels. These shallow energy levels are effective at low temperatures only and cause darkening comparable with that at room temperature. Further, relatively pure BaTiO_3 samples were exposed at liquid nitrogen temperatures and the darkening was considerably greater than for exposure at room temperature. However, the change was only about half that found in relatively impure BaTiO_3 at low temperature.

These results tend to indicate that the thermophototropic response is strongly influenced by the type and concentration of traps. Increasing the

type and concentration of impurities having trap energies effective at room temperature will produce a greater color change at that temperature. Similarly, bleaching temperatures may be lowered by excluding impurities with deeply lying energy states. Experiments with gaseous ambients did indicate, as expected, a strong correlation between TPT and the presence of oxygen. It is likely that the oxygen ions remain at the surface after ionization radiation darkens the specimen. The presence of these ions maintains the electrical neutrality of the crystal and these ions are the recombination centers for thermoluminescent behavior.

CURRENT THEORIES

Several qualitative theories of TPT in the compounds discussed have appeared in the literature. MacNevin and Ogle conclude that it is structure-dependent, because of the nonresponsive nature of MgTiO_3 (Reference 7). Further, the presence of certain specific impurities is essential, those having valences other than +4 are needed for electron transfer to occur and ionic radii near that of Ti^{4+} are necessary to fit the lattice with only moderate distortion. MacNevin and Ogle also found that phototropic response increased with impurity concentration but that neither moisture nor oxygen influenced the color change. They attributed some of the darkening in CaTiO_3 , to the reduction of Ti^{4+} to colored Ti^{3+} . Since there was no mention of preoxidation heating and since the alkaline-earth titanates do not accept more oxygen than the stoichiometric composition calls for (Reference 28) it is readily understood why the oxygen dependence was not found.

Kosman and Bursian attribute the darkening of polycrystalline BaTiO_3 to the partial reduction and precipitation of the metal (Reference 29). However, they do not consider the role of impurities nor surface ambients.

Tanaka (Reference 8) reported surface coloration in CaTiO_3 , due to daylight, x-ray, and ultraviolet radiation, which did not fade in the dark, was dependent on the presence of impurities, and was thermoreversible. He inferred that TPT is moisture dependent since it did not occur in the presence of phosphorous pentoxide. He concluded that some impurity decomposes water photochemically, and that the hydrogen formed

thereby acts to reduce TiO_2 to a lower oxide, causing color. It is difficult to reconcile Tanaka's theory with the fact that increased concentrations of oxygen increase the color change.

Weyl and his co-workers (References 32-34) conclude that when light strikes an impurity ion, say Fe^{3+} , an electron of the foreign ion is excited. It moves either into an oxygen vacancy of the defective structure, thereby producing Fe^{4+} , or attaches itself to a Ti^{4+} ion to give colored Ti^{3+} . This mechanism cannot be entirely accepted since it permits coloration independent of an ambient. Further, different impurities having different activation energies will negate the critical dependence of TPT on wavelength.

Bear and McTaggart (References 4 and 6), working with various mixtures of anatase and rutile and a number of other host oxides, conclude that impurities are absorbed at the surface or at interfaces. They found that samples sealed in vacuum with moisture did not develop color when irradiated; thus an oxygen-containing atmosphere was necessary. The valence state of the impurities is changed by the oxygen when the sample is exposed to light. The valence state of the impurity atoms corresponding to the exposed or unexposed state was not determined conclusively. They believe that the impurity is oxidized by light causing the sample to darken, that the reduced form of the impurity is the stable form, and that heating reduces the impurity causing bleaching. This is contrary to the valence state found in iron-doped samples by Sellwood and his co-workers (Reference 35). The reversion of valence is accelerated by heating at relatively low temperatures, causing a rapid reversal of color. They feel that the presence of water molecules catalyzes or promotes the development of color by action with oxygen or by entering into the bonding between the impurity and the host oxide. The host oxide converts the incident light energy into a form that activates the valence change of the impurity. Further, they find that the bleaching temperature is strongly dependent on the particular doping substance (iron-doped oxides bleach at 100°C , chromium-doped oxides at 190°C , and vanadium-doped oxides at about 175°C). Although they do not specify how light energy activates a valence change, nor the exact roles played by oxygen and water molecules, their conclusions

generally coincide with those presented here for the perovskite titanates.

CONCLUSION

A mechanism has been proposed which accounts for thermophototropic behavior in BaTiO_3 and related material. A high density of surface vacancies results in oxygen adsorption. Radiation with photon energies approximately equal to or greater than the gap energy frees electrons from the Ti-O bonds formed at the surface. The conduction electrons reduce the impurities present in the material causing coloration. Reversion to the original color is achieved by thermal ionization of the impurity levels accompanied by recombination radiation.

ACKNOWLEDGMENT

The author wishes to express his appreciation to Dr. Louis Walter for his continued interest and stimulating discussions throughout the course of this work.

REFERENCES

1. HORIE, T., KAWABE, K., and SAWADA, S., "Optical Behaviour of Multi-Domain Single Crystal of BaTiO_3 in Dependence on Temperature," *Phys. Soc. Japan J.* 9(5):823-825, September-October 1954.
2. LINZ, A., and HERRINGTON, K., "Electrical and Optical Properties of Synthetic Calcium Titanate Crystal," *J. Chem. Phys.* 28(5):824-825, May 1958.
3. NOLAND, J. A., "Optical Absorption of Single-Crystal Strontium Titanate," *Phys. Rev.* 94:724, May 1, 1954.
4. BEAR, J., and McTAGGART, F. K., "Phototropic Effects in Oxides. II. White Oxides in General," *J. Appl. Chem.* 8(1):72-76, 1958.
5. BRECKENRIDGE, R. G., and HOSLER, W. R., "Electrical Properties of Titanium Dioxide Semi-conductors," *Phys. Rev.* 91:793-802, August 15, 1953.
6. McTAGGART, F. K., and BEAR, J., "Phototropic Effects in Oxides. I. Titanium Dioxide," *J. Appl. Chem.* 5:643-653, 1955.
7. MACNEVIN, W. M., and OGLE, P. R., "Phototropy of the Alkaline Earth Titanates," *Amer. Chem. Soc. J.* 76:3846-3848, 1954.
8. TANAKA, Y., "Studies on the Reactions between Oxides in Solid State at Higher Temperatures. II. The Reaction between Calcium Oxide and Titanium Oxide and the Photo-Sensitivity of Calcium Titanate," *Chem. Soc. Japan Bull.* 16(12):455-463, December 1941.
9. LEVIN, S. B., FIELD, N. J., PLOCK, F. M., and MERKER, L., "Some Optical Properties of Strontium Titanate Crystal," *Opt. Soc. Amer. J.* 45(9):737-739, September 1955.

10. WILLIAMSON, W. O., "Darkening of Some Commercial Titanium Dioxide in Daylight," *Nature* 140:238-239, 1937.
11. WILLIAMSON, W. O., "Photosensitive Titanium Dioxide," *Nature* 143:279, 1939.
12. WILLIAMSON, W. O., "Reversible Photosensitivity in Some Artificial Materials Containing Rutile," *Mineralog. Mag.* 25:513-528, 1940.
13. WILLIAMSON, W. O., "Reversible Darkening in Daylight of Some Glazes Containing Titanium," *Brit. Ceram. Soc. Trans.* 39:345-365, 1940.
14. MCQUARRIE, M., and BEHNKE, F. W., "Structural and Dielectric Studies in the System (Ba, Ca)-(Ti, Zr)-O₃," *Amer. Ceram. Soc. J.* 37:539-543, 1954.
15. COFFEEN, W. W., "Dielectric Bodies in Metal Stannate-Barium Titanate Binary Systems," *Amer. Ceram. Soc. J.* 37(10):480-489, October 1, 1954.
16. WAINER, E., and WENTWORTH, C., "Niobate and Tantalate Dielectrics," *Amer. Ceram. Soc. J.* 35:207-214, 1952.
17. JONKER, G. H., "Keramische materialen voor de electrotechniek," *Chem. Weekblad* 49:932-936, December 5, 1953.
18. MEGAW, H. D., "Crystal Structure of Double Oxides of the Perovskite Type," *Phys. Soc. Proc.* 58:133-152, March 1946.
19. KAY, H. F., and BAILEY, P. C., "Structure and Properties of CaTiO₃," *Acta. Crystallog.* 10(3):219-266, March 10, 1957.
20. WYCKOFF, R. W. G., "Crystal Structures," New York: Interscience Publishers, 1948.
21. WELLS, A. F., "Structural Inorganic Chemistry," Oxford, England: The Clarendon Press, 1945.
22. RANKAMA, K., and SAHAMA, T. G., "Geochemistry," Chicago: University of Chicago Press, 1950.
23. KINGERY, W. D., "Introduction to Ceramics," New York: John Wiley, 1960.
24. ANLIKER, M., BRUGGER, H. R., and KANZIG, W., "Das Verhalten von kolloidalen Seignetteelektrika III, Bariumtitanat BaTiO₃," *Helv. Phys. Acta* 27(2):99-124, 1954.
25. KANZIG, W., "Space Charge Layer Near the Surface of a Ferroelectric," *Phys. Rev.* 98(2):549-550, April 15, 1955.
26. ADAMSON, A. W., "Physical Chemistry of Surfaces," New York: Interscience Publishers, 1960.
27. CRONEMEYER, D. C., "Electrical and Optical Properties of Rutile Single Crystals," *Phys. Rev.* 87:876-886, September 1, 1952.
28. WEISE, E. K., and LESK, I. A., "On The Electrical Conductivity of Some Alkaline Earth Titanates," *J. Chem. Phys.* 21:801-806, May 1953.
29. KOSMAN, M. S., and BURSIAK, E. V., "Coloration of BaTiO₃ Crystals," *Dokl. Akad. Nauk SSSR* 115(3):483-485, 1957 (In Russian).
30. SALMANG, H., "Ceramics: Physical and Chemical Fundamentals," (Translated by M. Francis), London: Butterworths Publications, 1961.
31. NOVOSILTSEV, N. S., KHODAKOV, A. L., and SHULMAN, M. C., "Metastabilnie sostoyaniya BaTiO₃," *Dokl. Akad. Nauk SSSR* 83(6):829-831, April 1952.
32. WEYL, W. A., and FÖRLAND, T., "Photochemistry of Rutile," *Indus. Eng. Chem.* 42:257-263, 1950.
33. JOHNSON, G., and WEYL, W. A., "Influence of Minor Additions on Color and Electrical Properties of Rutile," *Amer. Ceram. Soc. J.* 32(12):398-401, December 1, 1949.
34. WEYL, W. A., and TERHUNE, N., "Crystal Chemistry Applied to Foreign Atoms in Titanate Ceramics," *Ceram. Age* 61:22, August 1953.
35. SELWOOD, P. W., ELLIS, M., and WETHINGTON, K., "Supported Oxides of Iron," *Amer. Chem. Soc. J.* 71:2181-2184, 1949.

A PHENOMENOLOGICAL MODEL OF THERMOPHOTOTROPISM*

JOHN B. SCHUTT

An equation is developed which describes the phenomenology of thermophototropism. The equation could have application to thermal analysis and, hence, be used in a feasibility study on thermophototropes for temperature control in the space environment. A literature survey is included in the appendix.

INTRODUCTION

Thermophototropism, or temperature-dependent variable absorptivity, is the explicit designation of a phenomenon known variously as phototropism, thermotropism, or both. The term "thermophototropism," although a bit unwieldy, cannot be misinterpreted as can one of the latter, "phototropism," which has a biological connotation. The importance of formulating a phenomenological model of thermophototropic materials was twofold. First, it would give a physical basis on which various experimental observations could be correlated and, second, it would be of use in thermal design.

The materials have been found as members of both major classes of chemical compounds, inorganic and organic. Inorganic thermophototropes can be classified with respect to their anion constituent, oxygen or sulfur. With oxygen present, the matrix and doping cations are members of group II, III, V, or VI in the periodic table; whereas, with sulfur present, the cation is mercury (IIa). The oxide compounds (most of which are reversible) derive their behavior from:

1. Surface oxidation in the presence of oxygen, moisture, or both;
2. Internal impurity adsorption on the surfaces of the individual microcrystalline hosts;
3. Vacancy migration;
4. Internal oxidation.

The mercuric compounds (most of which are irreversible) derive their behavior from:

1. Phase changes;
2. Diffusion mechanisms;
3. Oxidation and reduction reactions (in the presence of light, oxygen, or both).

Because reversibility is necessary for thermal coating applications, this phenomenological development of a model of thermophototropism is applicable only to reversible materials. The next section, which is concerned with the phenomenology of thermophototropism, is developed from certain qualitative experimental observations on these materials.† In the succeeding section, a partial differential equation (to which a solution is found) is developed from the phenomenology. The final section is devoted to a discussion of the equation and to the possible application of optical measurements in the evaluation of certain parameters in the solution. The inorganic materials mentioned in the body of the report as well as additional ones are discussed in Appendix A. Also, organic substances are briefly considered.

PHENOMENOLOGY OF THERMOPHOTOTROPISM

Since a phenomenological model of the temperature-dependent variable absorptivity property possessed by certain materials must be capable of reproducing certain gross experimental observations,† it is the purpose of this section to present and elucidate these observations, in preparation

*Published as *NASA Technical Report R-184*, January 1964.

†NASA contract NAS 5-582, Seventh Quarterly Progress Report, Lexington Laboratories, Cambridge, Massachusetts, p. 14a.

for the mathematical development of the model.

As the name "thermophototropism" implies, the energy absorbed by a thermophototrope is dependent upon its photoresponse, i.e., its change in absorptivity with time at a particular temperature and illumination level. For example, should a given sample happen to be in its minimum absorptivity state at a given temperature, upon illumination we would expect its absorptivity in the visible range of the spectrum to increase to a maximum value governed by the temperature of the sample. Likewise, with the removal of illumination, the absorptivity would gradually return to its minimum value, the rate of return being dependent upon temperature.

The thermal dependence of the absorptivity of these materials may be divided into four regions. In two of these regions thermophototropes cannot respond temporally; i.e., if a sample is at either too low or too high a temperature its absorptivity cannot be changed from its minimum value by solar radiation. Consequently, whenever the temperature is less than some critical level T_0 , absorptivity changes are, at best, imperceptively slow. Further, by adiabatically reducing a darkened sample to T_0 , its absorptivity can be frozen in. Likewise, whenever the temperature is equal to or greater than a certain level, T_u , darkening becomes impossible. Therefore, as the temperature is raised from the minimum critical value T_0 a third temperature region is encountered, in which positive increments in temperature exert a synergistic effect with photon flux in producing an enhancement in the darkening ability of a given sample; i.e., $\partial\alpha/\partial T$ must be positive, where α and T represent absorptivity and temperature, respectively. Consequently, a temperature T_c , lying between T_0 and T_u , will be reached at which $\partial\alpha/\partial T$ is zero. The fourth temperature region is that in which further increments in temperature result in a negative value for $\partial\alpha/\partial T$. This division into four regions also applies to any sublayer of a sample.

Finally, the overall behavior of a material will necessarily be dependent upon the illumination intensity and spectral distribution. It is clear then that the energy absorbed, W , and α can be represented in the functional form:

$$W = W(\alpha) .$$

where

$$\alpha = \alpha\{t[I(\lambda)], x[I(\lambda)], T[I(\lambda)]\}.$$

with

t = time exposed to illumination,

x = depth measured from the surface (a positive value),

T = temperature,

$I(\lambda)$ = spectral intensity.

The effects of illumination intensity and spectral distribution can be assessed only from optical measurements and microscopic theory. Consequently, the $I(\lambda)$ dependence will be implicitly understood in the next section in connection with the determination of proportionality and integration constants. Therefore $\alpha = \alpha[t, x, T; I(\lambda)]$.

MATHEMATICAL DEVELOPMENT

Now a partial differential equation will be derived which, when solved, can be used in describing, qualitatively, the temporal, thermal, and depth dependences of the materials. Since the change in the energy absorbed relative to the absorptivity minimum is time-dependent,

$$\left. \frac{\partial W}{\partial (t - t_M)} \right|_{x, T}$$

is the required first order contribution at an arbitrary depth x and temperature T . The derivative is written in terms of a relative time because the response depends upon how much time has passed since the material was in its maximum absorptivity state located in time at t_M . This derivative has a positive sign because ΔW and $t - t_M$ are both negative. Similarly, the inclusion of temperature to the first order gives the expression

$$-\tau \left. \frac{\partial W}{\partial (T - T_c)} \right|_{t, x}$$

where again the independent variable has been written in relative form since the temperature gradient of the absorptivity may be positive or negative, depending on whether T is greater or

less than T_c . A thermal response τ must be included because darkening is dependent upon the response of the material to any temperature change which occurs as a result of illumination.

The energy penetration term is

$$u \frac{\partial W}{\partial (x - x_0)} \Big|_{t, \tau} + u\eta W \Big|_{t, \tau}$$

where u and η are the penetration velocity and

$$\frac{dW}{dt} = \frac{\partial W}{\partial (t - t_M)} - \frac{\partial W}{\partial (T - T_c)} \tau + u \frac{\partial W}{\partial (x - x_0)} + u\eta W. \quad (1)$$

Before writing Equation 1 in terms of absorptivity, two regions for the spectral distribution must be defined. The first region contains wavelengths up to a critical magnitude; the second region contains the remainder. It has been experimentally observed that wavelengths greater than the critical one may cause bleaching (a decrease in absorptivity), whereas those which are shorter induce darkening.* Therefore,

$$W = A_p I \int_0^t (a - a_0) d\zeta. \quad (2)$$

absorption index, respectively; the derivative is taken relative to an arbitrary reference point x_0 in order to identify penetration from any chosen layer. This relation is clearly the case because the gradient across any infinitesimal layer is dependent not only upon the rate at which the darkened layer progresses into the material from a chosen layer, but also upon the rate of energy absorption in that layer. Finally, the total change in the relative energy absorbed is

where

A_p = projection of sample area normal to the incident illumination,

I = illumination intensity,

a = absorptivity at wavelengths less than the critical value,

a_0 = minimum absorptivity,

and W is (as mentioned in the previous section) the incremental energy absorbed relative to that absorbed at a_0 . Substituting Equation 2 into Equation 1 yields.

$$\frac{d(a - a_0)}{dt} = \frac{\partial (a - a_0)}{\partial (t - t_M)} - \frac{\partial (a - a_0)}{\partial (T - T_c)} \tau + u \frac{\partial (a - a_0)}{\partial (x - x_0)} + u\eta (a - a_0). \quad (3)$$

since the integral must be uniformly convergent. Expressing $d(a - a_0)/dt$ as $(a - a_0)/t$ and rewriting Equation 3 gives

$$\frac{\partial (a - a_0)}{\partial (t - t_M)} - \frac{\partial (a - a_0)}{\partial (T - T_c)} \tau + u \frac{\partial (a - a_0)}{\partial (x - x_0)} + u\eta (a - a_0) - \frac{a - a_0}{t} = 0. \quad (4)$$

where the state a is now set at $t=0$.

We can make the substitution

$$a - a_0 = g(t) f(T) h(x), \quad (5)$$

to obtain the solution:

$$a = a_0 + (\text{const.}) t \exp - \left[S_1 (t - t_M) + \frac{S_2}{\tau} (T - T_c) + \frac{S_2 - S_1 + u\eta}{u} (x - x_0) \right]. \quad (6)$$

*NASA contract NAS 5-582, Seventh Quarterly Progress Report, Lexington Laboratories, Cambridge, Massachusetts, p. 14a.

There remains the problem of obtaining the integration constant and S_1 and S_2 , the separation constants. The integration constant may be found by observing that the maximum absorptivity, a_M , for a given illumination intensity and spectral distribution is defined when $t=t_M$, $T=T_c$, and

$x=x_0$, from which it follows that

$$a_M = a_0 + (\text{const.}) t_M$$

or

$$\text{const.} = \frac{a_M - a_0}{t_M}$$

and

$$a = a_0 + (a_M - a_0) \frac{t}{t_M} \exp - \left[S_1 (t - t_M) + \frac{S_2}{\tau} (T - T_c) + \frac{S_2 - S_1 + u\eta}{u} (x - x_0) \right] \quad (7)$$

The first separation constant S_1 may be determined by partially differentiating Equation 7 with respect to t :

$$\frac{\partial a}{\partial t} = (a_M - a_0) \left(1 - t S_1 \right) \frac{1}{t_M} \exp - \left[S_1 (t - t_M) + \frac{S_2}{\tau} (T - T_c) + \frac{S_2 - S_1 + u\eta}{u} (x - x_0) \right]$$

and since

$$\left. \frac{\partial a}{\partial t} \right|_{t=t_M, x=x_0, T=T_c} = 0$$

then

$$S_1 = \frac{1}{t_M}$$

Since Equation 4 is a first order partial differential equation, it amounts to a specified geometric relation between slopes parallel to the various three-dimensional hyperplanes in the four space defined by a , t , τ , and x . An immediate consequence of this relationship is the following system of ordinary differential equations:

$$d(t - t_M) = - \frac{d(T - T_c)}{\tau} = \frac{d(x - x_0)}{u} = \frac{d(a - a_0)}{\left(-u\eta + \frac{1}{t} \right) (a - a_0)} \quad (8)$$

where the number of independent variables has been increased by one by means of the transformation:

$$z = z[y_i, (a - a_0)] \quad (9)$$

By virtue of

$$\frac{\partial z}{\partial y_i} + \frac{\partial z}{\partial (a - a_0)} \frac{\partial (a - a_0)}{\partial y_i} = 0 \quad (10)$$

Equation 4 can now be written:

$$\frac{\partial z}{\partial (t - t_M)} - \frac{\partial z}{\partial (T - T_c)} \tau + u \frac{\partial z}{\partial (x - x_0)} - u\eta (a - a_0) \frac{\partial z}{\partial a} + \frac{a - a_0}{t} \frac{\partial z}{\partial a} = 0 \quad (11)$$

Equation 8 is an immediate consequence of Equation 11. Integration of the first part of Equation 8 gives

$$t - t_0 = - \frac{1}{\tau} (T - T_0) \quad (12)$$

Because letting $t=t_M$ is equivalent to setting $T=T_c$, it is clear from Equations 7 and 8 that

multiplying both sides of Equation 12 by S_1 will show S_2 equal to S_1 ,

$$S_1 = S_2 = \frac{1}{t_M}$$

Equation 11 also serves to relate t_M and T_c ,

$$t_M = - \frac{1}{\tau} (T_0 - T_c)$$

by virtue of Equation 8 where t or any reference time t_0 is equal to 0. T_0 has now been identified as the first critical temperature, described in the previous section.

Likewise, from the integration of

$$dt = \frac{d(x - x_0)}{u}$$

u , the velocity at which the lowest measurable darkening level progresses through the material, may be evaluated:

$$t_L = \frac{L}{u}$$

or

$$u = \frac{L}{t_L}$$

where L is the sample thickness and t_L is the time required for measurable darkening to take place on the unilluminated surface.

In order to evaluate τ , the thermal response, assume that the sample is at uniform temperature

apart from heating due to radiation absorption. Consequently, at the surface

$$mC_p\tau = \gamma I a_0 (1 - \beta_v) + I a_0 (1 - \gamma) - \epsilon \sigma T^4$$

or

$$\tau = \frac{\gamma I a_0 (1 - \beta_v)}{mC_p} + \frac{I a_0 (1 - \gamma)}{mC_p} - \frac{\epsilon \sigma T^4}{mC_p} \quad (13)$$

where

m = mass of the sample,

C_p = heat capacity at constant pressure,

γ = fraction of illumination responsible for phototropic change in the sample,

β_v = quantum efficiency of the wavelength band in the neighborhood of the critical wavelength,

ϵ = sample emittance,

σ = Stefan-Boltzmann constant. (14)

The final result is:

$$a = a_0 + (a_M - a_0) \frac{t}{t_M} \exp - \left[\frac{t - t_M}{t_M} + \frac{T - T_c}{\tau t_M} + \eta(x - x_0) \right] \quad (15)$$

After being illuminated to a given level of absorptivity where

$$a_0 \leq a \leq a_M,$$

a thermophototrope will bleach from a to a_0 . Consequently, in the absence of illumination,

$$\frac{dW}{dt} = 0;$$

and Equation 4 becomes:

$$\frac{\partial(a - a_0)}{\partial(t - t_\xi)} - \frac{\partial(a - a_0)}{\partial(T - T_c)} + u \frac{\partial(a - a_0)}{\partial(x - x_0)} + u\eta(a - a_0) = 0 \quad (16)$$

where t_ξ is the exposure time. From Equation 16

$$a = a_0 + (a_M - a_0) \exp - \left[\delta_1 (t - t_\xi) + \frac{\delta_1}{\tau} (T - T_c) + \eta (x - x_0) \right] \quad (17)$$

If δ_1 is defined as the relative time required for the absorptivity to decay to the e^{th} part at a given T and x , then

$$\delta_1 = \frac{1}{t_0 - t_\xi} \tag{18}$$

which gives

$$a = a_0 + (\text{const.}) \exp - \left[\frac{t - t_\xi}{t_0 - t_\xi} + \frac{T - T_c}{\tau_\xi(t_0 - t_\xi)} + \eta_\xi(x - x_0) \right] \tag{19}$$

where $t \geq t_\xi$; τ_ξ and η_ξ identify the decay of their respective parts of the exponential at a given temperature and depth and consequently are defined by the e^{th} part of their contributions:

$$\tau_\xi = \frac{T_\xi - T_c}{t_0 - t_\xi} \tag{20}$$

and

$$\eta_\xi = \frac{1}{x_\xi - x_0} \tag{21}$$

At time t_ξ , temperature T_ξ , and depth x_ξ

$$(a_\xi - a_0)_{t_\xi, \text{darkening}} = (a_\xi - a_0)_{t_\xi, \text{bleaching}} \tag{22}$$

From Equations 15 and 22 the constant in Equation 19 may be found:

$$\text{const.} = (a_m - a_0) \frac{t_\xi}{t_m} \exp - \left[\frac{t_\xi - t_m}{t_m} + \frac{T_\xi - T_c}{\tau t_m} - \frac{T_\xi - T_c}{\tau_\xi(t_0 - t_\xi)} + \eta(x_\xi - x_0) - \eta_\xi(x_\xi - x_0) \right] \tag{23}$$

The resultant decay is therefore, by virtue of Equations 19 and 23:

$$a = a_0 + (a_m - a_0) \left\{ \frac{t_\xi}{t_m} \exp - \left[\frac{t_\xi - t_m}{t_m} + \frac{T_\xi - T_c}{\tau t_m} - \frac{T_\xi - T_c}{\tau_\xi(t_0 - t_\xi)} + \eta(x_\xi - x_0) - \eta_\xi(x_\xi - x_0) \right] \right\} \cdot \exp - \left[\frac{t - t_\xi}{t_0 - t_\xi} + \frac{T - T_c}{\tau_\xi(t_0 - t_\xi)} + \eta_\xi(x - x_0) \right] \tag{24}$$

In order to set a bound on the temperature T_u (the maximum temperature, at which a_0 is the only possible absorptivity state) let $x_\xi = x_0$, $t = t_m$, and $T = T_u$ in Equation 15:

$$T_u = n\tau t_m + T_c \tag{25}$$

where n is the e^{th} part to be taken. Again T_u

has been defined in a manner similar to t_0 ; i.e., decay to the e^{th} part.

DISCUSSION

It is clear that Equation 15 qualitatively fits the phenomenology of thermophototropism presented in the first section. A reasonable temporal variation of absorptivity has been found and the syn-

ergy of time and temperature reproduced. Derivation relative to the parameters α_0 , t_M , and T_c , the magnitudes of which can be found by experimentation, facilitated the formulation of the first order differential. Had this not been done, it would have been necessary to extend the derivation in a manner analogous to derivations of the Navier-Stokes equations. Such an extension would have led to equations which are not amenable to solution. This fact justifies the approach, because an equation has been developed which, after "calibration," can reproduce the thermophototropic property.

Evaluation of the parameters α_0 , α_M , t_M , T_c , T_0 , τ , and η is a matter of experimentation. The α 's may be found spectrophotometrically, once the effect of beam intensity on a given material has been determined. The values of t_M and T_c must be found simultaneously, by trial, and may be checked for consistency with Equation 8. T_0 and T_u may be found simply from simultaneous illumination and heating experiments. The value for τ can be found from Equations 13, once β_v is found from optical measurements. And η is defined once β_v is known and $\bar{\eta}$ has been evaluated by means of optical transmission measurements. T_u and t_0 are defined from their respective e^{th} parts, since Equation 15 gives no upper bounds for T and t .

In an effort to demonstrate the potential usefulness of Equation 15, comparison plots of Equation 15 and data from NASA contract NAS 5-582* are given in Figure 1, where relative incremental

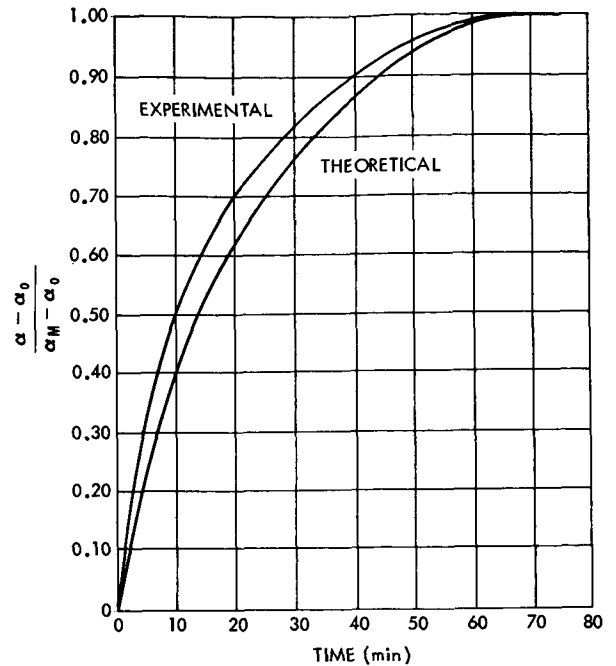


FIGURE 1—A comparison of theoretical and experimental incremental absorptivity curves for barium titanate at 550 m μ .

absorptivity is the ordinate and time the abscissa. The agreement is quite good; however, it must be remembered that the data for the experimental curve were not taken expressly for theoretical evaluation, but rather to give qualitative trends. It is certain that, when the samples were exposed, conditions were not isothermal; consequently, since T_c was probably above room temperature, the synergistic effect mentioned earlier probably was present.

*NASA contract NAS 5-582, Seventh Quarterly Progress Report, Lexington Laboratories, Cambridge, Massachusetts, p. 14a.

APPENDIX A

LITERATURE SURVEY ON PHOTOTROPISM

Takei (References A1-A4) has studied, extensively, mercuric compounds of the forms $\text{HgX}_2 \cdot 2\text{HgS}$ and $\text{HgXCNS} \cdot 2\text{HgS}$ (where X is a halogen atom). From thermophototropic studies of these substances, those which show the most promise for use in the space environment are $\text{HgI}_2 \cdot 2\text{HgS}$ and $\text{HgBr}_2 \cdot 2\text{HgS}$ (Table A1). $\text{HgI}_2 \cdot 2\text{HgS}$ has a decomposition temperature which is nearly twice its bleaching temperature (the favorable ratio of the group in Table A1). $\text{HgBr}_2 \cdot 2\text{HgS}$ does not have as favorable a decomposition to bleaching temperature ratio, but does possess a relative darkening rate nearly 3 times that of $\text{HgI}_2 \cdot 2\text{HgS}$. Mercuric compounds do not bleach completely at room temperature in the absence of light; for $\text{HgI}_2 \cdot 2\text{HgS}$, the bleaching process requires 20 minutes at 90–100°C.

Takei has taken photomicrographs of bleached $\text{HgI}_2 \cdot 2\text{HgS}$ that, according to his analysis, have shown colloidal mercury deposited on the surface. He attributes this to thermal decomposition. However, he did not report the evolution of sulfur

dioxide. It should be pointed out that HgI_2 sublimes at the bleaching temperature of $\text{HgI}_2 \cdot 2\text{HgS}$. Takei's HgS was prepared by the hydrogen sulfide process, which coprecipitates Hg along with HgS ; this could explain the appearance of surface deposits of colloidal mercury. The addition of 0.65–1.1 percent selenium to $\text{HgI}_2 \cdot 2\text{HgS}$ accelerated the bleaching process.

Magnetic susceptibility studies on the compounds listed in Table A1 show: (1) Each mixture changed from a diamagnetic to a paramagnetic state at the temperature of compound formation. (2) The $\text{HgX}_2 \cdot 2\text{HgS}$ compounds were stable enough to undergo repeated thermal phototropic cycles (a contradiction to Takei's oxidation hypothesis).

The thermophototropic behavior of mercury compounds is not understood. Takei reports that a phase change takes place, but that it is entered into by less than 20 mole percent of the molecules. If, for instance, sulfur has a preference for the surface layer in the presence of illumination (at

Table A1

Some Properties of $\text{HgX}_2 \cdot 2\text{HgS}$ and $\text{HgXCNS} \cdot 2\text{HgS}$ Thermal Phototropes.

Compound	Relative Darkening Rate	Decomposition Temperature (°C)	Bleaching Temperature (°C)	Color Change
$\text{HgI}_2 \cdot 2\text{HgS}$	1.0	185	90	Orange-yellow to black
$\text{HgBr}_2 \cdot 2\text{HgS}$	2.95	250	197	Pale yellow to black
$\text{HgCl}_2 \cdot 2\text{HgS}$	0.79	260	225	White to black
$\text{HgICNS} \cdot 2\text{HgS}$	0.56	213	136	Orange-yellow to black
$\text{HgBrCNS} \cdot 2\text{HgS}$	1.08	288	202	White-yellow to black
$\text{HgICNS} \cdot 2\text{HgS}$	0.026	290	260	White-yellow to black

room temperature), this could explain darkening; i.e., the surface concentration of sulfur would increase (HgS can exist in a black modification). On the other hand, if iodine has preference for the surface layer at elevated temperatures, the bleaching process is explained (HgI_2 can exist in a yellow modification). This also explains the tendency of HgI_2 to distill from the surface at the bleaching temperature of $\text{HgI}_2 \cdot 2\text{HgS}$. Consequently, the mechanism in $\text{HgI}_2 \cdot 2\text{HgS}$ appears to be a combination of molecular rotation (a phase change) and thermal diffusion.

Extensive investigations on thermal phototropic oxides have been carried out by McTaggart and Bear on titanium dioxide substrates (References A5 and A6). They investigated both anatase and rutile structures and found that only rutile gave combined thermal and optical activities. Some of their results are given in Table 2A; it can be

Table A2
Color Effects of Certain Metals in Rutile.

Element	Temp. (°C)	Fusion Time (hr.)	Unexposed Color	Exposed Color
Nd	830	1	White	Very pale pink-gray
	1020	$\frac{1}{2}$	White	Grayer
Pr	830	1	White	Very pale pink-gray
	1020	$\frac{1}{2}$	White	Grayer
Sm	830	1	White	Very pale pink-gray
	1020	$\frac{1}{2}$	White	Grayer
Co	840	1	Very pale yellow-green	Faint pink-gray
	1000	$\frac{1}{2}$	Very pale yellow-green	Pale lemon yellow
Mn	810	1	White	Pale pink-gray

observed that no member of the table is sufficiently thermophototropic to be of use in the space environment. Table A3 gives additional results of the investigations of McTaggart and Bear from a more recent paper (Reference A6). Their conclusions are:

1. Individual crystal adsorption of the doping agent is responsible for thermal phototropism in rutile.

2. Light quanta eliminate oxygen from the rutile structure. This produces a reduced impurity and results in a fatigue effect.
3. Removal of illumination allows the oxidized impurity to return to its normal state of oxidation.
4. Heat accelerates the reduction of the oxidized impurity.
5. The presence of water vapor is necessary to free the oxygen from the rutile and make it thermophototropic.

McTaggart and Bear emphasized the importance of fusion time in the formation of phototropic oxides. It was shown, for instance, that prolonged heating at 800°C yielded no phototropic oxides. This was observed because the impurity had time to diffuse into the rutile structure from its adsorbed state on a rutile crystal.

The maximum color change occurred in rutile doped with chromic oxide—from yellow to black. However, chromium-doped rutile would not darken in the absence of oxygen, and its phototropism was destroyed when it was heated in a vacuum. When it was heated in the presence of oxygen it followed the usual behavior.

MacNevin and Ogle, Tanaka, and Weyl and Förland have studied phototropism in titanates (References A7–A9). They found that calcium titanate doped with iron, zinc, antimony, or vanadium was phototropic; the corresponding doped barium titanate was less phototropic. Silver, copper, tin, and zirconium doping agents gave no effect with either calcium or barium titanate. In contrast to the properties of rutile phototropes, MacNevin and Ogle found that there was no fatigue effect, no oxygen effect, no water vapor effect, and no combined oxygen and water vapor effect on doped calcium and barium titanates. Their conclusions were:

1. The metallic ion in the impurity must have a slightly different ionic radius than Ti^{+4} in order to deform the potential field and produce phototropism.
2. The charge of the metallic ion in the impurity must not be $+4$.

Chalkley wrote an extensive review of phototropism, in which he discussed organic (solid and

Table A3
Compilation of Thermophototropes.

Host	Fe	Cr	Cu	Ni	V	Mn	Bleaching of Dye*
TiO ₂ (anatase)	x	+	+	+	+	+	Very fast
TiO ₂ (rutile)	x	x	x	+	x	x	Very fast
Nb ₂ O ₅	Pale cream to dark gray x	x	Off white to light brown x	+	x	x	Slow
Al ₂ O ₃	+	x	+	+	x	+	Fast
ZnO (carbonate)	+	+	x	+	x	+	
ZnO (nitrate)	+	+	x	+	x	-	
HfO ₂	+	x	-	-	-	-	
ThO ₂	+	+	+	+	x	+	Slow
SnO ₂	+	+	Light gray to brown x	-	-	+	Fairly fast
Ta ₂ O ₅	+	+	Brown to tan x	-	-	+	Fast
ZrO ₂	+	+	x	x	+	+	Very Slow
BeO	+	+	+	+	x	+	Slow
GeO ₂	+	+	+	+	+	+	Very slow
SiO ₂	+	+	+	+	+	+	Very slow
MgO	+	+	+	+	+	+	Very slow

*These compounds were mixed with chlorazol sky blue FF dye in order that bleaching rates could be checked more accurately.

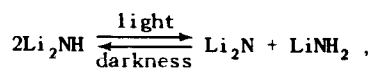
*Present.

+No observed effect.

liquid) and inorganic compounds (Reference A10). (In this review liquid phototropes will not be presented.) Chalkley found solid organic phototropes among the following classes of compounds: hydrazones, osazones, fulgides, stilbenes, substituted biphenyl disulfides, and substituted biphenyl disulfoxides. Unfortunately he presented no information on the absorption shifts expected from these compounds. Fatigue effects were present in organic phototropes, but without significant chemical change; whereas for inorganic phototropes fatigue effects were accompanied by drastic chemical changes. It is open to speculation, therefore, whether or not antifatigue compounds could be incorporated into certain

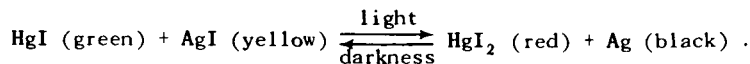
organic phototropes. Salicylidene aniline is a notable exception; i.e., it is phototropically stable.

Chalkley briefly discussed a few inorganic phototropes—lithium amide, zinc oxide with zinc sulfide, and mercurous iodide with silver iodide. Lithium amide, which derives its phototropism from the reaction,



was not discussed with regard to color change or phototropic stability. Mixtures containing zinc oxide and zinc sulfide change from white to black

in the presence of sunlight and moisture; the reverse reaction requires oxygen. The phototropic reaction of the mixture of mercurous iodide with silver iodide was written:



Unfortunately, information on possible fatigue effects and absorption shifts was not presented.

Phototropes were reported to have produced more thermoelectric power and exhibited the photoelectric effect to a greater extent than compounds of similar constitution which were not phototropic.

REFERENCES

- A1. TAKEI, K., "Photoreversible Reaction. I. Photoreversible Reaction of Mercuric Complex Salts," *J. Chem. Soc. Japan (Pure Chem. Sec.)* 73:202-204, 1952.
- A2. TAKEI, K., "Photoreversible Reaction. II. Kinetic Study of Photoreversible Reaction of Mercuric Complex Salts," *J. Chem. Soc. Japan (Pure Chem. Sec.)* 73:204-207, 1952.
- A3. TAKEI, K., "Study of Reversible Photo-Chemical Reactions. III. Mechanism of the Phototropy of Mercuric Complex Salts. I. Changes of the Physical Properties in the Reversible Color Change Reaction of $\text{HgI}_2 \cdot 2\text{HgS}$," *Bull. Chem. Soc. Japan* 28(6):403-408, August 1955.
- A4. TAKEI, K., "Study of Reversible Photo-Chemical Reactions. IV. Phototropy of $\text{HgX}_2 \cdot 2\text{HgS}$ and HgXCNS ," *Bull. Chem. Soc. Japan* 28(6):408-410, August 1955.
- A5. McTAGGART, F. K., and BEAR, J., "Phototropic Effects in Oxides. II. Titanium Dioxide," *J. Appl. Chem.* 5:643-653, 1955.
- A6. BEAR, J., and McTAGGART, F. K., "Phototropic Effects in Oxides. II. White Oxides in General," *J. Appl. Chem.* 8(1):72-76, 1958.
- A7. MACNEVIN, W. M., and OGLE, P. R., "Phototropy of the Alkaline Earth Titanates," *J. American Chem. Soc.* 76:3846-3848, 1954.
- A8. TANAKA, Y., "Studies on the Reactions Between Oxides in Solid State at Higher Temperatures. I. The Reaction Between Magnesium Oxide and Titanium Oxide," *Bull. Chem. Soc. Japan* 16(11):428-441, November 1941.
- A9. WEYL, W. A., and FÖRLAND, T., "Photochemistry of Rutile," *Ind. Eng. Chem.* 42:257-263, 1950.
- A10. CHALKLEY, L., JR., "Phototropy," *Chem. Rev.* 6:217-280, 1929.

THERMOPHOTOTROPISM IN CORTICATED ALKALINE EARTH TITANATES*

J. B. SCHUTT AND J. A. COLONY
D. R. LEPP†

A comparison of acidic and acidic-to-basic impurity introduction on the thermophototropic behavior of the alkaline earth titanates is presented. The latter is a new method of introducing impurities into alkaline earth titanates which will be described here. It involves the synthesis of a photoactive titanate from a previously synthesized titanate or titanyl oxalate. This method gave a considerably enhanced absorptivity change for barium titanate, but proved to be much less effective for the titanates of calcium and strontium. In all cases the slurries were made alkaline with alkali metal hydroxides. For barium titanate the optimum alkalinity was obtained with 10 mole percent sodium hydroxide, which led to the conclusion that the phenomenon is a surface one in which the alkali metal oxide corticates the surfaces of the host crystals.

INTRODUCTION

In the rather large group of photochemically active materials which have been categorized according to their optical properties, there exist many materials which change color when exposed to light of certain wavelengths. This phenomenon is referred to in its broadest sense as "phototropy."[‡] The additional property of immediate color reversal with the removal of illumination is usually implied, but is not satisfied, *ipso facto*, by all phototropes. The alkaline earth titanates (with the exception of magnesium) provide an example. The speed of their change in color depends on their temperature. At room temperature they revert to their original color imperceptibly slowly. For instance, a decrease of 2 to 3 percent in the absorptivity of barium titanate at room temperature requires approximately 3 weeks. However, an exposed sample heated to 300°C returns to its original color in about 30 min. Therefore, a graph for barium titanate of the functional dependence of coloration on temperature gives a broad curve which extends from about zero to several hundred degrees centigrade. Conse-

quently, the alkaline earth titanates are not phototropic in the same sense as organics and metal organics (which bleach at room temperature). Materials such as the alkaline earth titanates are called thermophototropes.

This paper reports a new procedure for the synthesis of a photoactive titanate from a previously synthesized titanate—a procedure which causes a considerably enhanced thermophototropic activity for the barium analog. It is appropriate to mention briefly other possible synthetic approaches. One is to dissolve a salt (impurity), usually a chloride, in distilled water and dilute this solution to such a volume that, when it is slurried with the titanate, a free flowing mixture is formed. Uniform distribution is then assured. A less convenient method involves wet or dry grinding of the impurity with the titanate. Identical procedures are suitable for use with the hydrated titanyl oxalates. Still another technique involves the coprecipitation of the impurity with a titanyl oxalate as part of the synthetic procedure.

McNevin and Ogle§ have reported qualitative results on the thermophototropic behavior of the alkaline earth titanates. The work was repeated here in order to establish quantitatively the extent of the photoresponse.

*Published as *NASA Technical Note D-2306*, June 1964.

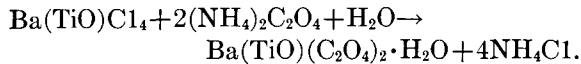
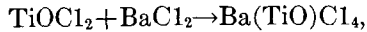
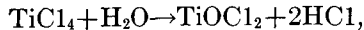
†TAAG Designs, Inc. Mr. Lepp did research work for this paper under contract NAS-52383 held by TAAG Designs, Inc., College Park, Maryland.

‡Chalkley, L., Jr., "Phototropy," *Chem. Rev.* 6:217-260, 1929.

§McNevin, W. M., and Ogle, P. R., "Phototropy of the Alkaline Earth Titanates," *J. Amer. Chem. Soc.* 76: 3846-3848, 1954.

THE ADDITION OF IMPURITIES

Initially, barium titanate was slurried with the chloride impurities. The resultant mixture was dried, pressed, and fired for 1 hr at 1100°C. The titanates were synthesized via the oxalate process:



The conversion to a titanate was accomplished at 1100°C in 4 hr. Previous to firing, the hydrated titanyl oxalate was slurried with VOCl_3 , CrCl_3 , FeCl_3 , NiCl_2 , CuCl_2 , and ZnCl_2 at a level of 0.02 mole percent. After firing, samples of the titanate were reslurried separately with solutions of elements in the first transition series. The exposures were carried out under a mercury lamp, rich in the violet and near ultraviolet wavelengths, for 24 hr.

ABSORPTIVITY CHANGES FOR VARIOUS IMPURITIES

The results of the spectrophotometric studies are given in Table 1 for selected wavelengths. In an effort to find a scheme for the incorporation of impurities into the titanates, it seemed appropriate to look for synergisms. $\text{V}^{+3, 5}$, Cr^{+3} , Fe^{+3} , and $\text{Mn}^{+3, +4}$ would be expected to predominantly replace Ti^{+4} , and Co^{+2} , Cu^{+2} , and Zn^{+2} the alkaline earth metal. Thus it appeared that synergistic effects could be produced through simultaneous replacement of titanium and barium with cations carrying different ionic charges than that on the cation to be replaced. To this end, the alkali metals were chosen. The results are given in Table 1 for barium titanate with sodium hydroxide at the 10 mole percent level. They are compared with the corresponding sample prepared without sodium hydroxide. The hydroxide was normally added by reslurring the doped sample; vanadium was added simultaneously with sodium by means of the pervanadate. Broadly, therefore, Table 1 is a comparison of impurity activity in an acidic titanate with that in a basic titanate.

In an effort to find out whether the enhanced photoactivity was synergistic, barium titanate was mixed with sodium hydroxide alone. Figure

Table 1
Comparison of the Absorptivity Changes of BaTiO_3 with Acidic and Acidic-Basic Impurity Introduction.

Impurity Addition	Absorptivity Change (percent)			
	$\lambda = 0.4\mu$	$\lambda = 0.5\mu$	$\lambda = 0.6\mu$	$\lambda = 0.7\mu$
$\text{V}^{+3, +5}$	0	2	1	1
$\text{V}^{+3, +5} + \text{NaOH}$	17	56	49	39
Cr^{+3}	0	2	1	1
$\text{Cr}^{+3} + \text{NaOH}$	21	55	54	41
$\text{Mn}^{+3, +4}$	0	1/2	1/2	1
$\text{Mn}^{+3, +4} + \text{NaOH}$	8	26	32	30
Fe^{+3}	1	6	4	3
$\text{Fe}^{+3} + \text{NaOH}$	10	46	53	48
Co^{+2}	6	4	4	3
$\text{Co}^{+2} + \text{NaOH}$	6	10	16	25
Ni^{+2}	0	2	2	1
$\text{Ni}^{+2} + \text{NaOH}$	23	57	57	50
Cu^{+2}	2	8	4	0
$\text{Cu}^{+2} + \text{NaOH}$	10	41	47	40
Zn^{+2}	0	6	5	4
$\text{Zn}^{+2} + \text{NaOH}$	11	44	45	33

1 gives the reflectance curves for lithium, potassium, and sodium added to this mixture, from 0.35 to 0.7 μ . In this and the following figures the upper set of curves represents bleached samples and the lower set darkened samples. (In all the figures and tables the alkali metals were added at the level of 10 mole percent unless otherwise stated).

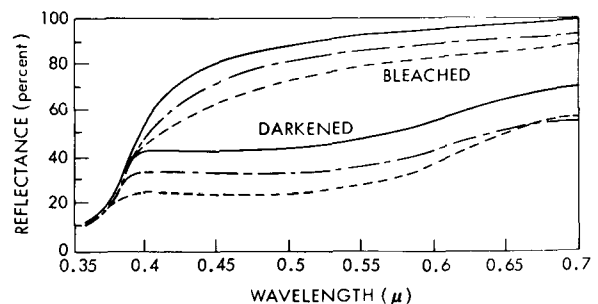


FIGURE 1.—A comparison of the colorations of impure barium titanate with 10 mole percent additions of lithium (unbroken curve), potassium (dashed curve), and sodium (dotted curve).

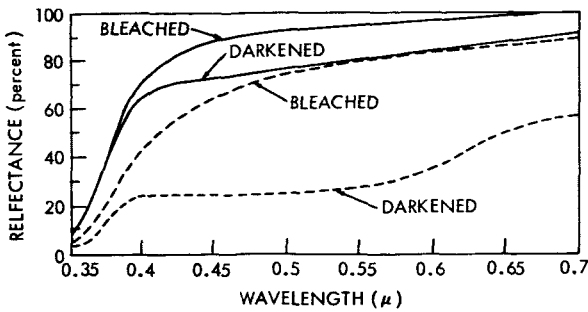


FIGURE 2.—A comparison of the coloration of barium titanate with 10 mole percent of sodium added (unbroken curve) and barium titanate with impurities and 10 mole percent of sodium added.

Figure 2 shows curve for barium titanate with the impurities V, Cr, Fe, and Mn and 10 mole percent sodium as before, and curves for the addition of sodium to the oxalate. Figure 3 is a comparison of the addition of sodium in hydroxide and carbonate forms. The enhanced photoactivity found for the carbonate cannot stem from the ease of calcination of sodium carbonate versus the ease of dehydration of sodium hydroxide, since the heat of formation of the former is about twice that of the latter. More likely, the enhancement can be attributed to a lack of decomposition, which gives a relatively dense electron environment.

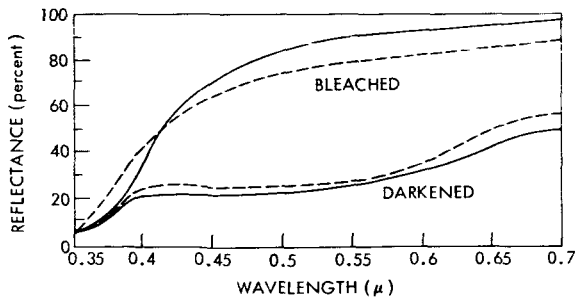


FIGURE 3.—A comparison of the colorations of impure barium titanate with 10 mole percent additions of sodium hydroxide (broken curve) and sodium carbonate.

In an effort to establish an approximate optimum concentration for sodium in barium titanate, the mole percent was varied from 2 to 30 while the samples were fired at 1100°C for 1 hr. The results are presented in Table 2, which shows that 10 mole percent is optimum. Samples with the optimum concentration for sodium were fired according to

Table 2
Absorptivity Increments of Impure BaTiO₃ for the Addition of Various Concentrations of Sodium.

Sodium Concentration (mole percent)	Absorptivity Change (percent)			
	$\lambda = 0.4\mu$	$\lambda = 0.5\mu$	$\lambda = 0.6\mu$	$\lambda = 0.7\mu$
2	6	30	26	17
5	9	50	47	36
10	15	50	49	37
15	9	55	52	37
20	9	52	51	37
30	8	48	47	35

Table 3
Absorptivity Changes of the BaTiO₃-Na₂O System under Various Firing Conditions.

Conditions	Absorptivity Change (percent)			
	$\lambda = 0.4\mu$	$\lambda = 0.5\mu$	$\lambda = 0.6\mu$	$\lambda = 0.7\mu$
820°C for 1 hr	4	43	37	27
820°C for 3 hr	3	48	44	33
950°C for 1 hr	6	46	43	32
1100°C for 1 hr	7	45	42	31

the schedules in Table 3. The optimum firing conditions turned out to be 1 hr and in the range 950–1100°C. Samples were made with lithium and potassium additions, under the optimum conditions for sodium addition to barium titanate. These were compared with the sodium addition and the results are presented in Figure 1. Lithium addition gave the least response. Additions of sodium and potassium gave similar responses, but initially the sample with potassium colored more rapidly than that with sodium.

Identical synthetic procedures were carried out for calcium and strontium titanates, i. e., V, Cr, Mn, and Fe were added to the respective oxalates and this was followed by repeated additions of sodium to the titanate. After drying, the alkali metals were added and the mixtures were fired. The results are given in Table 4. Sodium and potassium increased the overall photoactivity in calcium but did not increase the photoactivity when added to strontium. Lithium inhibited the

Table 4

Absorptivity Changes (percent) of Sodium-Doped CaTiO_3 and Sodium-Doped SrTiO_3 Systems at Various Wavelengths for the Introduction of Different Substances.

Addition	Absorptivity Change (percent)							
	CaTiO_3				SrTiO_3			
	$\lambda = 0.4\mu$	$\lambda = 0.5\mu$	$\lambda = 0.6\mu$	$\lambda = 0.7\mu$	$\lambda = 0.4\mu$	$\lambda = 0.5\mu$	$\lambda = 0.6\mu$	$\lambda = 0.7\mu$
Na	8	25	26	23	10	17	14	9
K	6	21	19	14	13	18	15	10
Li	-	-	-	-	9	13	11	8
Impure	6	13	13	9	19	17	14	8

overall response. The impure sample was a commercial alkaline earth titanate which was included for the sake of comparison.

CONCLUSION

The addition of sodium or potassium ions to impure barium titanate enhances the overall thermophototropic response. Since the level of addition was of such a magnitude that it precluded doping in the semiconductor sense, it seems reasonable to assume that the alkali metals are primarily surface active and do not predominately enter internal lattice positions. Consequently, they corticate the surfaces of the titanate host

microcrystals. Lithium gave the least optical response and was temporarily the slowest with respect to the rate of coloration. For potassium coloration was nearly as great as for sodium and darkening proceeded more rapidly than for sodium during the initial stages. Since the ionic radius of potassium, 1.33A, is nearly identical to that of barium, 1.35A, the replacement of barium by potassium is favored over substitution by sodium (0.95A). Since sodium can measurably increase the coloration, compared with potassium, it follows that Goldschmidt's Rules are only approximately obeyed for the surface-derived phenomenon of thermophototropism.

AUTHOR INDEX

- ALEXANDER, J. K., 1
ALLEN, W. K., 149
BACKUS, D. L., 153
BAILEY, J. W., 823
BAKER, J. L., 163
BALDERSTON, J. M., 185
BARTKO, F., 339
BAUERNSCHUB, J. P., JR., 25
BEHRING, W. E., 966
BERNIER, R. E., 567
BLAMONT, J. E., 1037
BOOTON, R. C., JR., 453
BOYLE, J. C., 577
BUSH, E. G., 193
CANTARANO, S., 463
CARR, F. A., 7
CARUSO, A. J., 366, 375
CATOE, C., 339
CHERRY, W. R., 835
CHI, A. R., 1041
CLEM, T. D., 200, 292
CLIFF, R. A., 854
COBLE, G. L., JR., 1145
COLONY, J. A., 1179
CONN, J. H., 31
COPE, D. S., 609
CROCKETT, W. R., 464
CUNNINGHAM, B. T., 618, 866
DARGO, D. R., 506
DARLING, E. M., JR., 625, 642
DELACHE, P., 1037
DEVINE, E. J., 36
DUNKELMAN, L., 377
ENG, T. L., 45
ENGELS, P., 473
ENGLES, P. D., 877
EVANS, D. S., 383
FEINBERG, P. M., 139
FITZGERALD, R. T., 473
FLATLEY, T. W., 652
FLOWERS, W. E., 823
FORSYTHE, R. W., 659, 669
GRANATA, R. L., 877
GRANT, M. M., 54
GREENSPON, J. E., 679
HABIB, E., 473
HALEV, M., 339
HALLAM, L., 394
HANLIN, K. L., 879
HARDGROVE, W. F., 725
HASSON, D. F., 1051
HEFFERNAN, P. J., 208
HENNES, J. P., 377
HOFFMAN, R. H., 567
HOLMES, D. W., 228
HORD, W. H., JR., 684
HORIUCHI, H. S., 234
HOROWITZ, R., 405
HOVEY, R. K., 823
HUNTER, C. M., 228
JOHN, J. E. A., 725
JOHNSON, A. C., 823
JOHNSON, C. C., 490, 499
JONES, G. H., 731
KAISER, J. A., 879, 884
KAISER, R. L., 242
KORVIN, R., 549
KUNDE, V., 339
KUTZ, R. L., 98
LANG, W. E., 742
LANTZ, P. A., 1071, 1095
LEAVITT, W. E., 242
LEIVO, W., 490
LEPP, D. R., 1179
LESKO, J. G., JR., 139
LITTLE, J. H., 884
LUNDE, B. K., 398
LYONS, J. C., 506
MACKEY, R. J., JR., 20
MANGUS, J., 394
MAURER, H., JR., 765
MAURY, J. L., JR., 889
MAXWELL, M. S., 98
MCCAUL, P. F., 877
MILLER, G. E., LT., USN, 405
MISH, W. H., 900
MITCHKO, M., 473
MONTES DE OCA, R. A., 777
MOSS, E., 618
NAGY, J. A., 1145
NEFF, W. J., 777

- NEUDER, S. M., 1160
NEUPERT, W. M., 366, 375
NEWTON, G. P., 405
NYBERG, W. C., 242
PAULKOVICH, J., 246
PELZ, D. T., 405
PITTMAN, M. E., 801
PLOTKIN, H. H., 926
POWERS, E. I., 567, 931
POYNTER, J. R., 412
RAUMANN, N. A., 945
RILEY, C. F., JR., 823
ROBINSON, G. B., 514
ROCHELLE, R. W., 252
SAITO, S., 412
SCHAFFERT, J. C., 292
SCHUTT, J. B., 1168, 1179
SCIULLI, J. A., 531
SCOTT, C. W., 823
SEMYAN, J., 979
SHAFFER, W., 473
SHARP, R. L., 866
SLIFER, L. W., 835, 866
SMITH, G. D., 542
SOAR, W. S., 988
SOLLARS, D. L., 995
SOMERLOCK, C. R., 298
STARBUCK, F., 823
STARK, L. W., 107, 119
STECKEL, J., 549
STEIN, H., 490
STEPHANIDES, C. C., 54
STEWART, W. N., 54
STOBER, A. K., 1037
STONE, R. G., 1
STUDER, P. A., 131
STYLES, F. J., 889
THORNWALL, J. C., 419
TIMMINS, A. R., 567
TOMBOULIAN, D. H., 430, 966
VEST, C. E., 812
WAGNER, C. A., 1008
WADELICH, D. L., 801
WALKER, W. C., 435, 441, 444
WAYLAND, J., 490
WHITE, A. F., JR., 119
WHITE, H. D., JR., 318
WOODMAN, R., 972
WRIGHT, D. U., JR., 446
YAFFEE, P., 823
YOUNG, T., 490