NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Space Programs Summary No. 37-35, Volume IV

for the period August 1, 1965 to September 30, 1965

Supporting Research and Advanced Development

FACILITY FORM 602

N66-13556 (ACCESSION NUMBER) (THRU)

374 (PAGES) 1 (CODE)

CR 68793 (NASA CR OR TMX OR AD NUMBER) 30 (CATEGORY)

GPO PRICE $ _____

CFSTI PRICE(S) $ _____

Hard copy (HC) 7.00

Microfiche (MF) 2.00

ff 653 July 65

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

October 31, 1965

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Space Programs Summary No. 37-35, Volume IV

for the period August 1, 1965 to September 30, 1965

Supporting Research and Advanced Development

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

October 31, 1965

Preface

The *Space Programs Summary* is a six-volume, bimonthly publication that documents the current project activities and supporting research and advanced development efforts conducted or managed by JPL for the NASA space exploration programs. The titles of all volumes of the *Space Programs Summary* are:

Vol. I. The Lunar Program (Confidential)

Vol. II. The Planetary–Interplanetary Program (Confidential)

Vol. III. The Deep Space Network (Unclassified)

Vol. IV. Supporting Research and Advanced Development (Unclassified)

Vol. V. Supporting Research and Advanced Development (Confidential)

Vol. VI. Space Exploration Programs and Space Sciences (Unclassified)

The *Space Programs Summary,* Vol. VI consists of an unclassified digest of appropriate material from Vols. I, II, and III; an original presentation of technical supporting activities, including engineering development of environmental-test facilities, and quality assurance and reliability; and a reprint of the space science instrumentation studies of Vols. I and II.



W. H. Pickering, *Director*
Jet Propulsion Laboratory

Space Programs Summary No. 37-35, Volume IV

Copyright © 1965, Jet Propulsion Laboratory, California Institute of Technology

Prepared under Contract No. NAS 7-100, National Aeronautics & Space Administration

Contents

SYSTEMS DIVISION

I. Systems Analysis 1

A. Impulsive Earth–Mars–Earth Trajectories for 1969, 1971, 1973, and 1975

Task No. 388-30105-2-3124 (684-30-03-00), P. H. Feitis 1

B. Gravity-Turn Descent Trajectory Analysis

Task No. 388-30105-2-3125 (684-30-03-00), R. S. Schlueter 8

C. Utilization of Energy Derived From the Gravitational Field of Jupiter for Reducing Flight Time to the Outer Solar System

Task No. 388-30105-2-3124 (684-30-03-00), G. A. Flandro 12

References 23

II. Scientific Programming 24

A. Triangulation of Plane Point Sets

Task No. 329-40301-1-3140 (129-04-04-01), C. L. Lawson 24

B. A Transformation of Triangulations

Task No. 329-40301-1-3140 (129-04-04-01), F. W. Weingarten 25

References 26

GUIDANCE AND CONTROL DIVISION

III. Spacecraft Power 27

A. Switched Regulator Microcircuits

Task No. 323-30301-2-3420 (123-33-08-01), W. K. Shubert 27

B. Solar Tracking System

Task No. 323-30701-2-3420 (123-33-02-03), S. S. Luebbers 30

IV. Spacecraft Control 33

A. Horizon Scan Platform System

Task No. 384-61901-2-3440 (186-68-02-09), T. Kerner and S. Szirmay 33

B. Spinning Spacecraft Attitude Control System

Task No. 325-90501-2-3440 (125-19-04-01), B. M. Dobrotin and J. C. Nicklas . . . 35

C. Qualitative Vacuum Testing of Attitude Control Thrust Nozzles Used in a Pulse Mode

Task No. 384-64701-2-3440 (186-68-02-22), T. J. Donlin 42

D. Bag-Type Quantitative Leak Measurements Using a Helium Mass Spectrometer

Task No. 384-64701-2-3440 (186-68-02-22), R. H. Summers and E. F. Koch 44

References 48

V. Guidance and Control Research 49

A. Some Electrical and Metallurgical Properties of the Fabricated Germanium Solid-State Diodes

Task No. 329-21801-1-3450 (129-02-05-09), A. Shumka 49

Contents (Cont'd)

B. Measurement of the Second-Harmonic Generation Coefficients in Barium Titanate
Task No. 329-20101-1-3450 (129-02-05-01), M. S. Shumate 52

References 55

VI. Flight Computers and Sequencers 56

A. Reliability Techniques for Spacecraft Computer Systems
Task No. 325-70401-2-3410 (125-17-04-02), A. Avizienis 56

References 59

ENGINEERING MECHANICS DIVISION

VII. Materials 61

A. Ceramic Research
Task No. 329-31101-1-3510 (129-03-04-01), M. H. Leipold 61

B. Basal Plane Shear in Pyrolytic Carbon
Task No. 329-31601-1-3510 (129-03-04-02), W. V. Kotlensky 62

C. Modification of Hemispherical Emissometer for Low Temperature and Automatic Data Acquisition
Task No. 324-90401-2-3510 (124-09-05-01), W. M. Hall 64

References 65

VIII. Applied Mechanics 66

A. Cavity-Type Absolute Total-Radiation Radiometer
Task Nos. 324-90601-2-3530 (124-09-05-03), J. A. Plamondon; 325-40801-2-3710 (125-24-03-05), J. M. Kendall, Sr. 66

B. A Study of Convective Heat Transfer Within a Mars Lander
Task No. 324-90601-2-3531 (124-09-05-03), M. Suo 69

References 71

IX. Lunar Spacecraft Development 72

A. High Impact Technology
Task No. 384-65501-2-3550 (186-68-06-01), J. L. Adams 72

ENGINEERING FACILITIES DIVISION

X. Aerodynamic Facilities 77

A. Wind Tunnels
Task Nos. 324-70105-7-3730 (124-07-04-01), W. Marko; 324-70409-7-3730 (124-07-04-02), G. M. Goranson; 324-70103-7-3730 (124-07-04-01), B. Walker; 324-70401-7-3730 (124-07-04-02), J. Welton 77

B. Hypervelocity Laboratory
Task No. 324-70813-2-3730 (124-07-01-04), B. Riale and T. Babineaux 89

Reference 91

Contents (Cont'd)

PROPULSION DIVISION

XI. Polymer Research 93

A. Mechanical Behavior of Styrene-Butadiene Rubber Filled with Glass Beads
Task No. 328-20301-1-3820 (128-32-05-02), R. F. Fedors and R. F. Landel . . . 93

B. The Effect of Surface Modifications on the Burning Rate of a Composite Solid Propellant
Task No. 328-20301-1-3820 (128-32-05-02), B. G. Moser, R. E. Weich, Jr., and R. F. Landel 99

C. Polymer Degradation Mechanisms: C^{14}-Labeled Isocyanate Adducts of Poly(Propylene Oxide)
Task No. 328-20401-1-3820 (128-32-05-03), E. F. Kopka and J. D. Ingham . . . 104

D. Structural Integrity of Solid Retrorockets
Task No. 328-20501-1-3820 (128-32-05-04), A. San Miguel 106

E. Electrical Properties of Poly-N-Vinylcarbazole-Iodine Charge Transfer Complex
Task No. 329-30401-1-3820 (129-03-11-03), A. Rembaum, D. S. Chapin, and L. E. Creasy 115

F. The Dependence on Composition of the Optical Absorption of the Poly(N-Vinyl Carbazole)-Tetracyanoethylene (PVCA-TCNE) Charge Transfer Complex in the Solid State
Task No. 329-30401-1-3820 (129-03-11-03), J. W. Farrar and J. Moacanin . . . 118

G. Carboxyl-Ending Poly (Alkylene Oxides) Purification and Curing
Task No. 329-30501-1-3820 (129-03-11-04), S. H. Kalfayan 123

H. Temperature Dependence of Outgassing Rates in Polymeric Foams
Task No. 384-62701-2-3820 (186-68-13-03), E. F. Cuddihy and J. Moacanin . . . 125

References 127

XII. Solid Propellant Engineering 130

A. Nozzle Thrust Misalignment
Task No. 328-21101-2-3810 (128-32-06-01), L. Strand 130

B. Applications Technology Satellite (ATS) Motor Development
Task No. 724-00081-7-3810 (630-01-00-00), R. G. Anderson and D. R. Frank . . . 140

XIII. Liquid Propulsion 149

A. Advanced Liquid Propulsion Systems
Task No. 331-10101-2-3840 (731-12-03-01), R. N. Porter, H. B. Stanford, and W. F. MacGlashan; 331-10301-2-3840 (731-12-03-02), D. D. Evans 149

B. Atomization and Injection Hydraulics
Task No. 328-10701-1-3840 (128-31-06-02), G. Jaivin 158

References 160

Contents (Cont'd)

XIV. Research and Advanced Concepts 161

A. Thrust Comparisons for Swirling and Nonswirling Flows of Argon Through a Convergent-Divergent Nozzle as Determined From Wall Pressure Measurements
Task No. 329-10701-1-3831 (129-01-09-04), P. Massier 161

B. Liquid MHD Power Conversion
Task No. 320-70301-1-3830 (120-27-06-03), D. Elliot, C. Cerini, L. Hays, and E. Weinberg 165

C. A Zero-Gravity Feed System for Mercury Ion Engines
Task No. 320-60101-2-3830 (120-26-04-01), T. D. Masek and D. J. Kerrisk . . . 169

D. Spectroscopic Measurement of Plasma Electron Temperature Using the Relative Intensities of Argon II Lines
Task No. 329-11401-1-3830 (129-01-07-01), N. Nerheim 174

E. Analog Study of Thermionic Reactor Stability
Task No. 320-70201-2-3830 (120-27-06-02), H. Gronroos and G. Kikin 180

References 188

SPACE SCIENCES DIVISION

XV. Lunar and Planetary Sciences 191

A. A Pattern Range for Small-Aperture Microwave Antennas
Task No. 390-20301-2-3250 (190-42-20-16), M. L. Kellner 191

B. Infrared Spectrometry of Mars
Task No. 383-31001-2-3250 (185-37-26-01), H. T. Bull 195

C. Scanning Mode Options for a Lunar Diffraction Experiment
Task No. 367-30102-2-3250 (867-14-00-01), J. A. Dunne 199

D. Crater Densities in Mare Tranquillitatis and Alphonsus in the 1-km to 10-m Diameter Range
Task No. 367-30102-2-3250 (867-14-00-01), J. A. Dunne 202

References 205

XVI. Chemistry 207

A. Further Studies of the Properties of Chelating Agents Coupled to Cellulose
Task No. 386-50201-2-3260 (189-55-02-02), A. J. Bauman, H. H. Weetall, N. Weliky, and L. Taylor 207

B. Soil Studies—Desert Microflora. IX. Measurement of Soil Moisture Suction by Thermal Conductivity Probes
Task No. 386-50301-1-3260 (189-55-04-01), R. E. Cameron, C. W. Odd, and H. E. Britton 209

C. Soil Studies—Desert Microflora. X. Soil Properties of Samples from the Chile Atacama Desert
Task No. 386-50301-1-3260 (189-55-04-01), R. E. Cameron, G. B. Blank, D. R. Gensel, and R. W. Davies 214

Contents (Cont'd)

D. The Reaction of O(¹D) with Methane
Task No. 329-21001-1-3260 (129-02-03-02), W. DeMore and O. Raper 222

E. Atomic Integrals with the Magnetic Dipole–Dipole Operator
Task No. 329-20801-1-3260 (129-02-03-04), M. Geller 228

F. A Micro Gas Chromatograph for Descent Analysis of the Martian Atmosphere
Task No. 383-31201-2-3260 (185-37-26-03), W. F. Wilhite 228

References 234

XVII. Fluid Physics 237

A. Experimental Investigations of the Base-Flow Problem. II. Wedge
Task No. 329-10201-1-3270 (129-01-09-01), F. R. Hama 237

B. The Inviscid Stability of the Cooled Laminar Boundary Layer. Part II
Task No. 329-10201-1-3270 (129-01-09-01), L. M. Mack 241

C. The Effective Diameter of an Orifice in Low Density Flow
Task No. 329-10501-1-3270 (129-01-10-01), H. Ashkenas 242

References 244

XVIII. Physics 246

A. Construction of Invariant Scattering Amplitudes
Task No. 329-20901-1-3280 (129-02-07-02), J. S. Zmuidzinas 246

B. The JPL Spark Chamber Program
Task No. 385-60301-2-3280 (188-46-01-01), W. S. McDonald 248

C. Low-Noise Amplifier for Solid-State Detectors
Task No. 385-60301-2-3280 (188-46-01-01), R. Lockhart 251

D. Relativistic Rigid Frames in Gravitational Fields
Task No. 329-20901-1-3280 (129-02-07-02), H. D. Wahlquist and F. B. Estabrook 255

E. Thermoelectric Effect in Superconducting Lead
Task No. 329-20401-1-3280 (129-02-05-04), A. F. Hildebrandt and D. D. Elleman 256

References 258

TELECOMMUNICATIONS DIVISION

XIX. Communications Elements Research 259

A. Microwave Electronics
Task No. 451-40500-1-3330 (150-22-14-05), W. H. Higa 259

B. Solid State Circuits
Task No. 350-10400-1-3330 (150-22-11-04), R. Brantner 261

C. Optical Communications Components
Task No. 325-20101-1-3330 (125-22-02-01-55), W. H. Wells 264

Contents (Cont'd)

D. Antennas for Space Communications
Task Nos. 350-10600-1-3330 (150-22-11-06), A. C. Ludwig; 451-31100-1-3330 (150-22-13-11), R. E. Cormack 266

E. RF Techniques
Task Nos. 325-10901-1-3330 (125-21-03-04), W. V. T. Rusch and C. T. Stelzried; 451-30900-1-3330 (150-22-13-09), B. Seidel and G. S. Levy 270

References 271

XX. Spacecraft Radio 273

A. Large Aperture Antenna Study
Task No. 325-20401-2-3360 (125-22-02-02), R. M. Dickinson 273

B. Telecommunications Capsule Antenna System Study
Task No. 384-61501-2-3360 (186-68-04-06), J. F. Boreham 278

C. Multipacting and Ionization Breakdown Studies
Task Nos. 325-20501-2-3360 (125-22-02-03), S. A. Brunstein; 325-20501-2-3360 (125-22-02-03), R. Woo 282

References 287

XXI. Communications Systems Research: Mathematical Research . . . 289

A. A Generalization of Two-Level Autocorrelation
Task No. 325-10701-1-3310 (125-21-01-01), R. J. McEliece 289

B. The Notion of a (p,q,r) Automation
Task No. 325-10701-2-3310 (125-21-01-01), R. P. Stanley 292

C. Further Results on the Algorithmic Complexity of (p,q) Automata
Task No. 325-10701-2-3310 (125-21-01-01), R. P. Stanley 298

D. A Useful Partition for Rook Domains
Task No. 325-10701-2-3310 (125-21-01-01), H. Taylor 304

References 306

XXII. Communications Systems Research: Communication Statistics . . 307

A. An Integral Formula for Limiter Suppression Factor
Task No. 350-10800-2-3310 (150-22-11-08), R. C. Tausworthe 307

B. Prefix Coding for Minimizing Storage Requirements
Task No. 325-10601-2-3310 (125-21-02-03), I. Jacobs and L. Kleinrock 309

C. Lunar Radar Maps in Polar Coordinates
Task No. 350-10900-2-3310 (150-22-11-08), T. Pitcher 319

References 326

XXIII. Communications Systems Research: Data Compression 328

A. Probabilistic Metric Spaces of Infinite Entropy
Task No. 325-10701-2-3310 (125-21-01-01), E. C. Posner and H. C. Rumsey, Jr. 328

B. Detection of Mixtures of Data Sources
Task No. 350-10900-2-3310 (150-22-11-09), I. Eisenberger 331

References 338

Contents (Cont'd)

XXIV. Communications Systems Research: Coded Communications . . 339

A. Effects on the Reliability of Telemetry Data Caused by the Earth-to-Spacecraft Link

Task No. 350-10800-2-3310 (150-22-11-08), W. C. Lindsey 339

B. New Techniques for Determining Weights of Code Words

Task No. 350-10900-2-3310 (150-22-11-09), G. Solomon and R. McEliece . . . 340

References 348

XXV. Communications and Tracking 349

A. Maximum Likelihood Symbol Synchronization

Task No. 350-10800-2-3310 (150-22-11-08), J. J. Stiffler 349

B. Proper Waveforms for Subcarriers in Digital Telemetry

Task No. 350-10800-2-3310 (150-22-11-08), R. Goldstein 357

XXVI. Astrometrics 360

A. Radar Signals in Model Atmospheres

Task No. 325-10801-2-3310 (125-21-02-04), P. E. Reichley and D. O. Muhleman . . 360

B. A-Priori Statistics in Differential Correction

Task No. 325-10801-2-3310 (125-21-02-04), P. E. Reichley 363

References 365

QUALITY ASSURANCE AND RELIABILITY OFFICE

XXVII. Parts Reliability 367

A. Effect of High Impact Shock on Seven Types of Transistors

Task No. 384-80101-2-1520 (186-58-13-01), K. Martin 367

B. Diode and Resistor Screening Evaluation

Task No. 384-00501-2-1520 (186-70-01-05), M. Johnston 368

SYSTEMS DIVISION

I. Systems Analysis

A. Impulsive Earth–Mars–Earth Trajectories for 1969, 1971, 1973, and 1975

P. H. Feitis

Preliminary results are given here concerning Earth–Mars–Earth trajectories which have an impulsive velocity correction at Mars. The velocity correction is required if the approach and the departure asymptotic speeds at Mars are not equal. If they are equal, the trajectory is termed a ballistic round-trip trajectory. This type was studied extensively by Minovitch (Ref. 1), who showed that this type of trajectory to Mars was possible in both the 1971 and 1973 opportunities, having total trip times in excess of 1000 days. This article will show that the total trip time can be reduced to about 500 days and that, in terms of data transmission, the probe returning to the Earth could transmit more data during *any* time interval after Mars encounter than the best possible orbiter. The following apply for this study:

(1) The total velocity increment after spacecraft parking-orbit injection, DVT, shall not exceed 7.1 km/sec ($C_3' = 100$ km^2/sec^2).[1]

(2) The total round-trip time shall not exceed 1000 days.

(3) Both sunlit-side and dark-side Mars flybys will be considered.

[1]The total velocity increment after spacecraft parking-orbit injection, DVT, is defined as the sum of all velocity increments applied to the spacecraft beginning at parking orbit, including that needed to inject the spacecraft from the parking orbit onto the Earth–Mars transfer trajectory and that needed in the vicinity of Mars, assuming the velocity correction maneuver is carried out at closest approach. DVT may be converted to conventional geocentric energy units, C_3', by:

$$C_3' = (V_0 + DVT)^2 - 2V_0^2,$$

where V_0 is the speed in parking orbit (7.8 km/sec). C_3' determines the payload of the round-trip mission, assuming the same engine is used to inject the spacecraft and carry out the velocity correction maneuver at Mars.

(4) No restriction will be placed on the altitude of closest approach to Mars (flyby altitude), except that it be greater than zero.

(5) No restriction will be placed on the asymptotic speeds with respect to the Earth at arrival.

(6) Only 10-day increments in Earth launch date *TL*, Mars arrival date *TA*, and Earth arrival date *T'A* will be considered.

1. Physical Model for a Round-Trip Trajectory

Let each of the round-trip trajectories consist of two one-way conic trajectories between two planets. Given *TL*, *TA*, and *T'A*, the heliocentric elements of the round-trip trajectory may be calculated, including the incoming and outgoing asymptotic velocity vectors with respect to Mars. In general, the magnitudes of these velocity vectors will be different, and it is thus necessary to perform a velocity correction maneuver in the vicinity of Mars. The Mars encounter geometry is shown in Fig. 1. The trajectory in the vicinity of Mars consists of two hyperbolas patched at a distance R from the center of Mars, where the velocity correction maneuver is carried out. This notation is defined in Fig. 1.

If no restriction is placed on the flyby altitude, a thorough analysis will show that there is a unique flyby altitude and velocity increment associated with every possible value of R. (The much more complicated problem which arises if one wishes to specify the flyby altitude will be discussed in a future *SPS*, Vol. IV.) Table 1 shows the flyby altitude and velocity increment required if the maneuver is carried out at: (1) a large (infinite) distance from Mars before encounter, (2) the common periapsis of both encounter hyperbolas, and (3) a large distance from Mars after encounter. It is seen in Table 1 that the minimum velocity increment is needed if the maneuver is carried out at the common periapsis of the encounter hyperbolas. If $R = \infty$, the velocity increment will be the difference between the departure asymptotic velocity $V_{\infty 2}$ and the approach asymptotic velocity $V_{\infty 1}$. It can be shown that the associated flyby altitudes represent the upper and lower limits of this variable for all possible maneuvers. Table 1 also shows that, if $V_{\infty 2} = V_{\infty 1}$, no velocity correction maneuver is necessary and there results a single unique flyby altitude.

The lighting conditions for Mars flybys are shown in Fig. 2. If the effect of Mars is to bend the asymptote toward the Sun, the spacecraft must pass the planet on

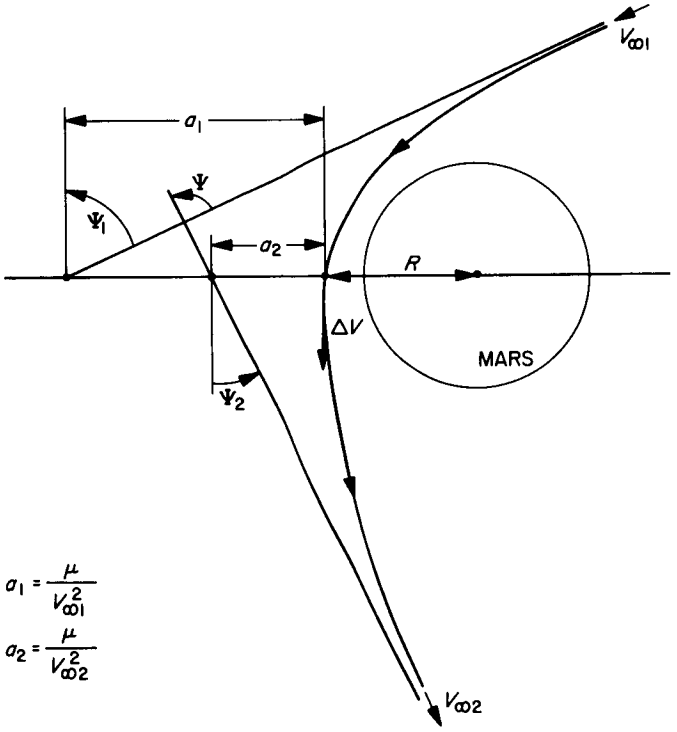


Fig. 1. Mars encounter geometry

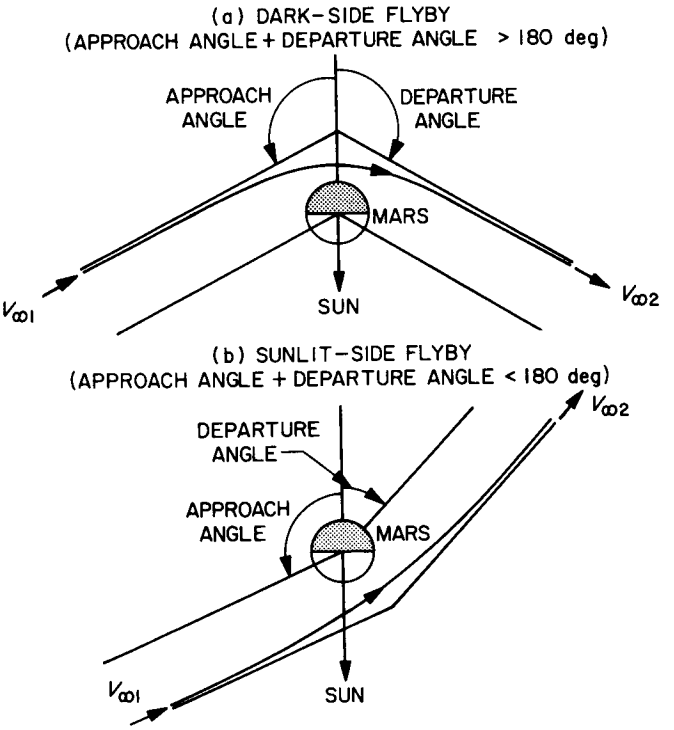


Fig. 2. Lighting conditions for Mars flybys

Table 1. Flyby altitude and required velocity increment for three types of maneuvers[a]

Point of maneuver	η_1, deg	η_2, deg	Distance from Mars, R	Flyby altitude	Velocity increment, ΔV
A large distance from Mars *before* encounter	–	$-[90 + (\Psi/2)]$	∞	$\frac{\mu}{V_{\infty 2}^2}\left[\frac{1}{\sin(\Psi/2)} - 1\right]$	$\lvert V_{\infty 2} - V_{\infty 1} \rvert$
Common periapsis of both encounter hyperbolas	0	0	R_p	R_p	$\Delta V_{\min}$
A large distance from Mars *after* encounter	$90 + (\Psi/2)$	–	∞	$\frac{\mu}{V_{\infty 1}^2}\left[\frac{1}{\sin(\Psi/2)} - 1\right]$	$\lvert V_{\infty 2} - V_{\infty 1} \rvert$

[a] Ψ = bending angle, η_1/η_2 = true anomaly of transfer point on the incoming/outgoing hyperbola, R_p = radius at the periapsis, and μ = Mars gravitational constant.

the dark side. Conversely, if the asymptote is bent away from the Sun, the spacecraft will pass Mars on the sunlit side. It should be noted that this property does not depend on where in the encounter trajectory the velocity correction maneuver is carried out.

Subject to the above constraints for a given Earth–Mars trajectory (i.e., for a given *TL* and *TA*), there exist numerous Mars–Earth trajectories which enable the probe to depart from Mars at *TA*. Each of the round-trip trajectories will pass Mars on either the sunlit side or the dark side, and each will have a different asymptotic velocity before and after Mars encounter. Due to the method of presentation, it is possible to show the trajectory characteristics of only one Earth–Mars–Earth trajectory for any given value of *TL* and *TA*. If no trajectories were found that passed Mars on the sunlit side, the data shown in Figs. 3 to 6 correspond to the round-trip trajectory having a minimum *DVT*. If, on the other hand, trajectories existed which passed Mars on the sunlit side, the particular trajectory among those passing on the sunlit side which has a minimum *DVT* is shown. In that case, trajectories which pass on the dark side having lower *DVT* may exist, but these are assumed to be "worse" than any which pass on the sunlit side.

2. Trajectory Characteristics of Earth–Mars–Earth Trajectories in the 1969, 1971, 1973, and 1975 Mars Opportunities

Figs. 3 to 6 show contour plots of *DVT*. It can be seen that the region of possible round-trip trajectories is contained within one or more oval-shaped contours. It was not possible to find any round-trip trajectories outside which satisfied the given constraints. The most important of the constraints is that the total trip time shall be between 200 and 700 days. Within the region of possible round-trip trajectories, contours corresponding to constant values of *DVT* have been drawn. Because the variation in the total flight time *T* is small, only its minimum and maximum values are indicated for each launch opportunity. Table 2 shows additional characteristics of the dark-side and sunlit-side flybys which have the lowest possible energy during each launch period.

a. 1968–1969 Launch period. Fig. 3 shows that, for sunlit-side flybys, *DVT* varies between 6.5 km/sec ($C_3' = 90$ km²/sec²) and 7 km/sec ($C_3' = 100$ km²/sec²). Much less energy is required for dark-side flybys, where *DVT* varies between 5 km/sec ($C_3' = 49$ km²/sec²) and 7 km/sec ($C_3' = 100$ km²/sec²). Total round-trip times vary between 440 and 580 days for both sunlit-side and dark-side flybys, and flyby altitudes are from 0 to 50,000 km. Over a 30-day launch window, the equivalent C_3' is seen in Fig. 3 to be 69 km²/sec² for sunlit-side flybys and 42 km²/sec² for dark-side flybys. Table 2 shows that the optimum round-trip trajectory in 1969 requires a very large velocity increment at Mars (on the order of 2 km/sec), and the energy is quite low.

b. 1970–1971 Launch period. In 1970–1971, the region of favorable *TL* and *TA* is greatly diminished. As may be seen in Fig. 4, the variations in *DVT* are similar to those for 1968–1969. Total round-trip times vary between 500 and 570 days for sunlit-side flybys and 370 and 550 days for dark-side flybys. The principal advantage in this launch period is that the required velocity increment at Mars is extremely small. Table 2 shows that the optimum sunlit-side trajectory requires a velocity increment at Mars of only 0.004 km/sec! It is possible to find a 30-day launch window during which the velocity increment is on the order of 0.03 km/sec. The corresponding value of C_3' is 66 km²/sec², and the total flight time is about 578 days. The dark-side flybys are seen to require substantially lower values of C_3', but a considerable velocity

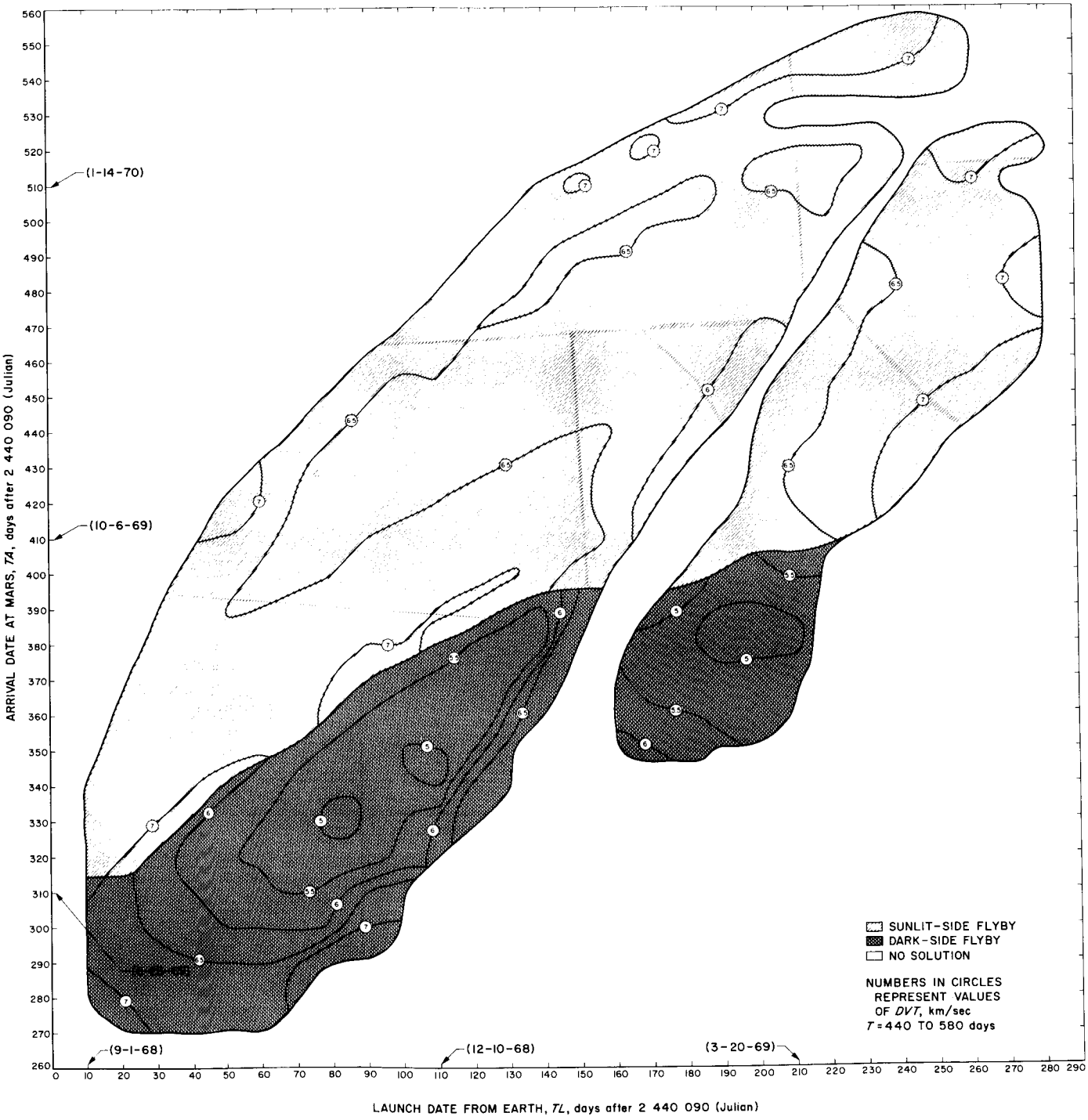


Fig. 3. Earth–Mars–Earth minimum-energy contours for 1968–1969

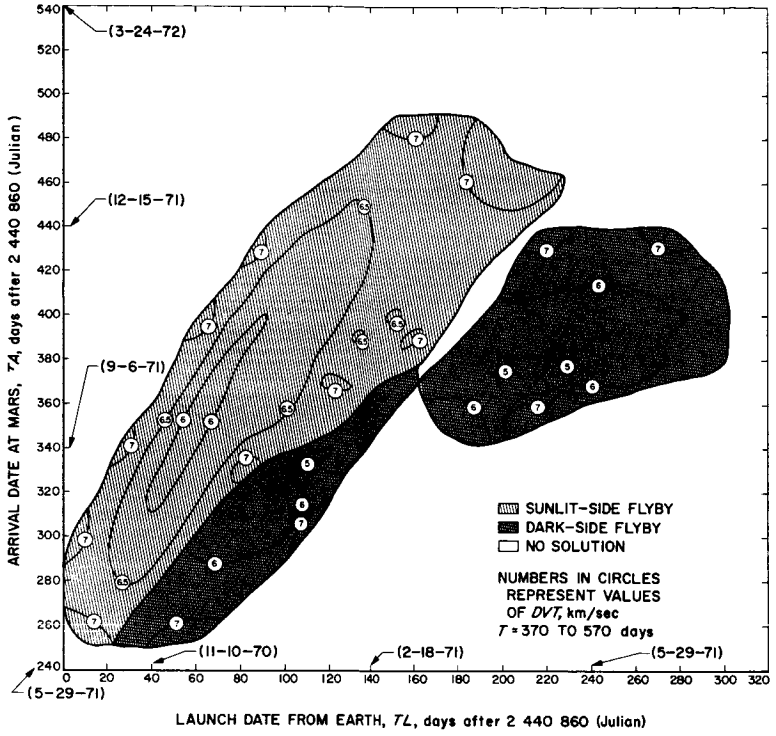


Fig. 4. Earth–Mars–Earth minimum-energy contours for 1970–1971

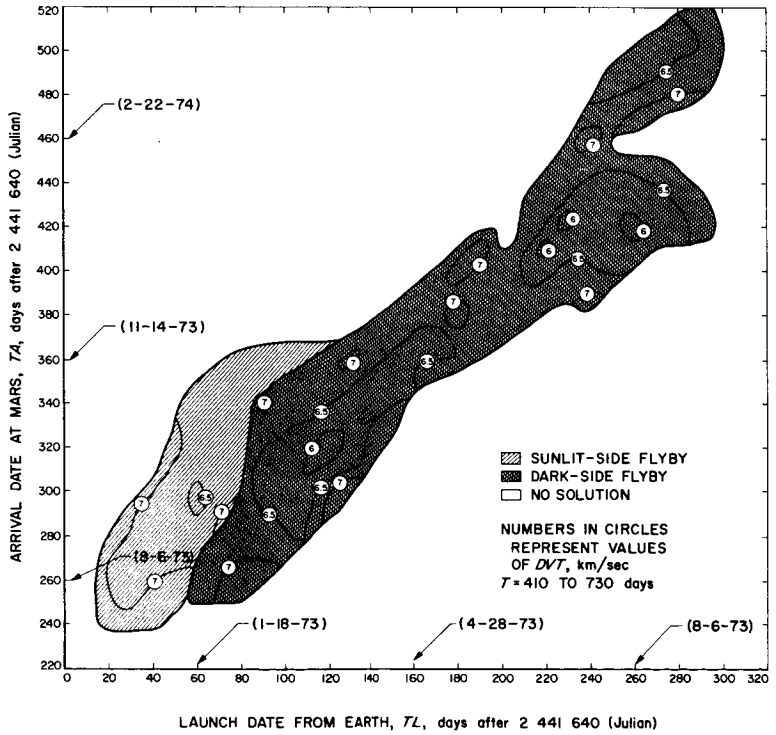


Fig. 5. Earth–Mars–Earth minimum-energy contours for 1972–1973

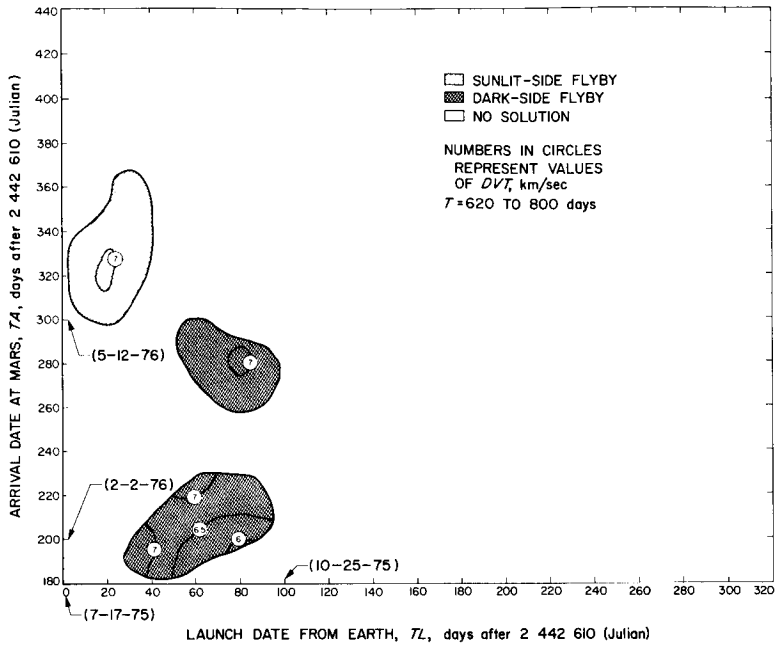


Fig. 6. Earth–Mars–Earth minimum-energy contours for 1974–1975

increment (on the order of 1.8 km/sec) must be applied at Mars. The total trip time is only 430 days.

c. 1972–1973 Launch period. Fig. 5 shows that the region of possible sunlit-side flybys is greatly reduced during the 1972–1973 Mars opportunity. This is accompanied by a substantial increase in the value of C_3'. For a 30-day launch window, C_3' would be 96 km²/sec² (DVT = 7 km/sec) and 83 km²/sec² (DVT = 6.5 km/sec) for sunlit-side and dark-side flybys, respectively. Table 2 shows that velocity increments on the order of 0.01 and 0.2 km/sec are required at Mars for sunlit-side and dark-side flybys, respectively. The flyby altitudes, unacceptably high for the sunlit-side flybys, could be reduced by increasing C_3'.

d. 1974–1975 Launch period. Fig. 6 shows that the region of possible round-trip trajectories diminishes even further during the 1974–1975 Mars opportunity. For a 30-day launch window, C_3' would be on the order of 90 km²/sec² for sunlit-side flybys and 82 km²/sec² for dark-side flybys. Table 2 shows that for sunlit-side flybys the energies at launch are relatively low (C_{3L} = 46 km²/sec²), but it is necessary to impart a very substantial velocity correction at Mars which nearly doubles the energy (C_3' = 85 km²/sec²). Total round-trip times are on the order of 800 days. Dark-side flybys require smaller energies and higher velocity corrections at Mars. Flight times for dark-side flybys are on the order of 620 days.

3. Data Return from an Earth–Mars–Earth Mission

A preliminary comparison of the data return from optimum Earth–Mars–Earth missions and optimum orbiter missions launched in the 1970–1971 launch period shows the following:

(1) The communication distance to the Earth return probe would always be smaller than that to the Mars orbiter. The Earth return probe could transmit 1.5 times as much data during the first 100 days after encounter than could the best orbiter. This ratio would increase with time.

(2) During the 2-day Earth encounter phase, the Earth return probe could transmit 600 times (aiming distance $B = 10^7$ km) to 40,000 times ($B = 10^5$ km) the amount of data transmitted by the orbiter in 200 days.

For these results, it was assumed that the Earth return probe and the orbiter would always transmit at the maximum possible data rate.

Table 2. Characteristics of the minimum-energy trajectories for each launch period

Launch period	Launch date (Julian) from Earth, *TL*	Arrival date (Julian) at Mars, *TA*	Geocentric launch energy C_{3L}, km^2/sec^2	Total velocity increment after parking-orbit injection, *DVT*, km/sec	Equivalent energy C_3', km^2/sec^2	Difference in asymptotic velocities at Mars, ΔV_∞, km/sec	Flyby altitude at Mars, km	Total flight time *T* (Earth–Mars–Earth), days
Sunlit-side flybys								
1968–1969	2 440 260	2 440 510	25	5.8	64	2.2	400	500
1970–1971	2 440 902	2 441 178	66	5.9	66	0.004	180	578
1972–1973	2 441 700	2 441 970	80	6.4	80	0.01	300,000	550
1974–1975	2 442 640	2 442 950	46	6.6	85	5.1	325	800
Dark-side flybys								
1968–1969	2 440 280	2 440 470	9	4.72	35	1.6	122	440
1970–1971	2 441 090	2 441 240	10.4	4.93	40.5	1.8	230	430
1972–1973	2 441 760	2 441 970	65	6	69	0.2	325	510
1974–1975	2 442 690	2 442 810	40	6	69	1.2	50	620

B. Gravity-Turn Descent Trajectory Analysis

R. S. Schlueter

1. Introduction

Within the framework of several simplifying assumptions and given a set of initial conditions, we wish to investigate the properties of gravity-turn descent trajectories. Gravity-turn steering implies that the direction of thrust is always opposite to the velocity vector. We will assume that the thrust is constant in magnitude from the initial time of firing until landing. As a result, the landing will be vertical and soft for a large range of initial conditions. With these assumptions to limit the problem, it is possible to do a relatively complete analysis without resorting to numerical techniques.

2. Mathematical Analysis

To begin, assume a body follows a gravity-turn trajectory such that: (1) the thrust-to-weight ratio and mass of a body are constant, (2) no atmospheric drag is present, and (3) the acceleration of gravity is constant. Since the thrust is always anti-aligned to the velocity vector, the geometry can be described in terms of the variables shown in Fig. 7. Using Fig. 8, the equation of motion in the ψ-direction can be written as

$$\frac{dV}{dt} = -ng + g\cos\psi, \tag{1}$$

where n is the local thrust-to-weight ratio. To obtain the equation of motion of ψ, differentiate

$$\tan\psi = \frac{V_x}{V_y}$$

to obtain

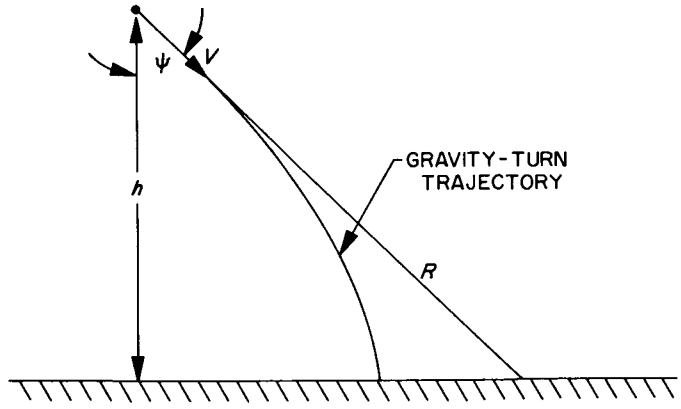
$$V\frac{d\psi}{dt} = -g\sin\psi\ . \tag{2}$$

If we divide Eq. (1) by Eq. (2), the differential equation

$$\frac{1}{V}\frac{dV}{d\psi} = \frac{n-\cos\psi}{\sin\psi} \tag{3}$$

can be integrated directly, giving the velocity as a function of ψ:

$$\frac{V}{V_0} = \left(\frac{\tan\frac{\psi}{2}}{\tan\frac{\psi_0}{2}}\right)^{n-1}\frac{\sec^2\frac{\psi}{2}}{\sec^2\frac{\psi_0}{2}}\ . \tag{4}$$



R = SLANT RANGE TO PLANET SURFACE ALONG VELOCITY VECTOR
V = VELOCITY
h = HEIGHT ABOVE PLANET SURFACE
ψ = ANGLE BETWEEN PERPENDICULAR TO PLANET AND VELOCITY VECTOR

Fig. 7. Gravity-turn geometry

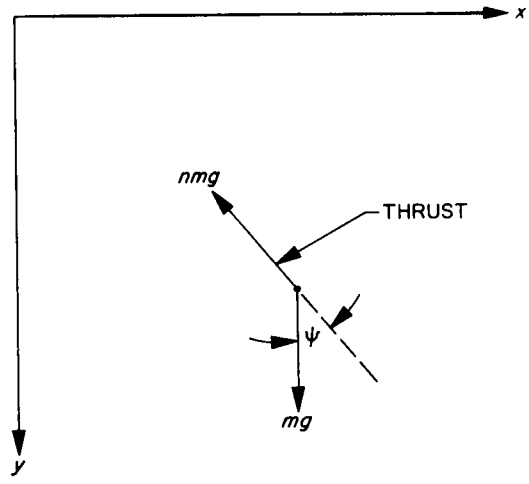


Fig. 8. Force diagram

The subscript "0" denotes the condition at time $t = 0$ of the velocity V, height h, and angle ψ. As Cheng shows in Ref. 2, when $V = V(\psi)$ is substituted into the second equation of motion, it is possible to show how the turning rate $d\psi/dt$ behaves as the body approaches the ground:

$$\psi = -g\sin\psi\left(\frac{\tan\frac{\psi}{2}}{\tan\frac{\psi_0}{2}}\right)^{1-n}\frac{\sec^2\frac{\psi_0}{2}}{\sec^2\frac{\psi}{2}}\ .$$

Since h and ψ approach zero at the same rate,

$$\lim_{h\to 0}\frac{\dot{\psi}}{} = \lim_{\psi\to 0}\psi = K\lim\psi^{2-n} = \begin{cases} 0 & \text{if } n<2,\\ K & \text{if } n=2,\\ \infty & \text{if } n>2.\end{cases}$$

Now, using Eq. (4) in Eq. (3) and integrating yields time as a function of ψ:

$$-\frac{gt}{V_0}\left(\tan\frac{\psi_0}{2}\right)^{n-1}\sec^2\frac{\psi_0}{2} = \int_{\psi_0}^{\psi}\frac{\left(\tan\frac{\psi}{2}\right)^{n-1}\sec^2\frac{\psi}{2}}{\sin\psi}d\psi\,. \tag{5}$$

If $n>1$, the integral in Eq. (5) is convergent for all ψ. Integrating and solving for t, we get

$$t = \frac{V_0\cos^2\frac{\psi_0}{2}}{(n+1)g}\left[\sec^2\frac{\psi_0}{2}+\frac{2}{n-1} - \left(\frac{\tan\frac{\psi}{2}}{\tan\frac{\psi_0}{2}}\right)^{n-1}\left(\sec^2\frac{\psi}{2}+\frac{2}{n-1}\right)\right] \tag{6}$$

Since we require the landing to be vertical, the elapsed time of flight can be found by setting $\psi=0$ in Eq. (6):

$$T = \frac{V_0}{(n^2-1)g}(n+\cos\psi_0) \tag{7}$$

where T is the powered-flight time.

We can get a parametric representation for h by integrating the following differential equation:

$$\frac{dh}{dt} = -V\cos\psi$$

$$dh = -V\cos\psi\,dt$$

$$= -V\cos\psi\frac{dt}{d\psi}d\psi$$

Integrating both sides gives

$$h+K_2 = -\int V\cos\psi\frac{dt}{d\psi}d\psi\,,$$

where K_2 is the constant of integration. Substituting for V and $dt/d\psi$ yields:

$$h+K_2 = \frac{K_1}{2}\int\left(1-\tan^4\frac{\psi}{2}\right)\left(\tan\frac{\psi}{2}\right)^{2n-3}\sec^2\frac{\psi}{2}d\psi\,, \tag{8}$$

where

$$K_1 = \frac{V_0^2\cos^4\frac{\psi_0}{2}}{g\left(\tan\frac{\psi_0}{2}\right)^{2n-2}}\,.$$

Again, the integral in Eq. (8) is convergent for $n>1$. Performing the integration and writing the result in the simplest form, we obtain

$$h+K_2 = \frac{K_1}{4(n^2-1)}\left(\tan\frac{\psi}{2}\right)^{2(n-1)} \times\left(\sec^4\frac{\psi}{2}\right)(\cos^2\psi+2n\cos\psi+1)\,. \tag{9}$$

To satisfy the initial conditions, set $h=h_0$ when $\psi=\psi_0$ and solve for K_2:

$$K_2 = \frac{V_0^2}{4(n^2-1)g}(\cos^2\psi_0+2n\cos\psi_0+1) - h_0$$

Finally, require $h=0$ when $\psi=0$, thereby implying that $K_2=0$. This gives the constraint condition between V_0, ψ_0, h_0, and n to ensure a vertical soft landing with gravity-turn steering:

$$h_0 = \frac{V_0^2}{4(n^2-1)g}(\cos^2\psi_0+2n\cos\psi_0+1). \tag{10}$$

This can be solved for n to give the required local thrust-to-weight ratio in terms of the initial conditions:

$$n = \frac{V_0^2}{4gh_0}\left\{\cos\psi_0+\left[\cos^2\psi_0 + \frac{4gh_0}{V_0^2}\left(\frac{4gh_0}{V_0^2}+\cos^2\psi_0+1\right)\right]^{1/2}\right\}. \tag{11}$$

For $0<\psi_0<90$ deg, $n>1$; therefore, the integrations performed previously are mathematically valid. With $K_2=0$, the parametric representation for h is

$$h = \frac{K_1}{4(n^2-1)}\left(\tan\frac{\psi}{2}\right)^{2n-2}\sec^4\frac{\psi}{2}(\cos^2\psi+2n\cos\psi+1). \tag{12}$$

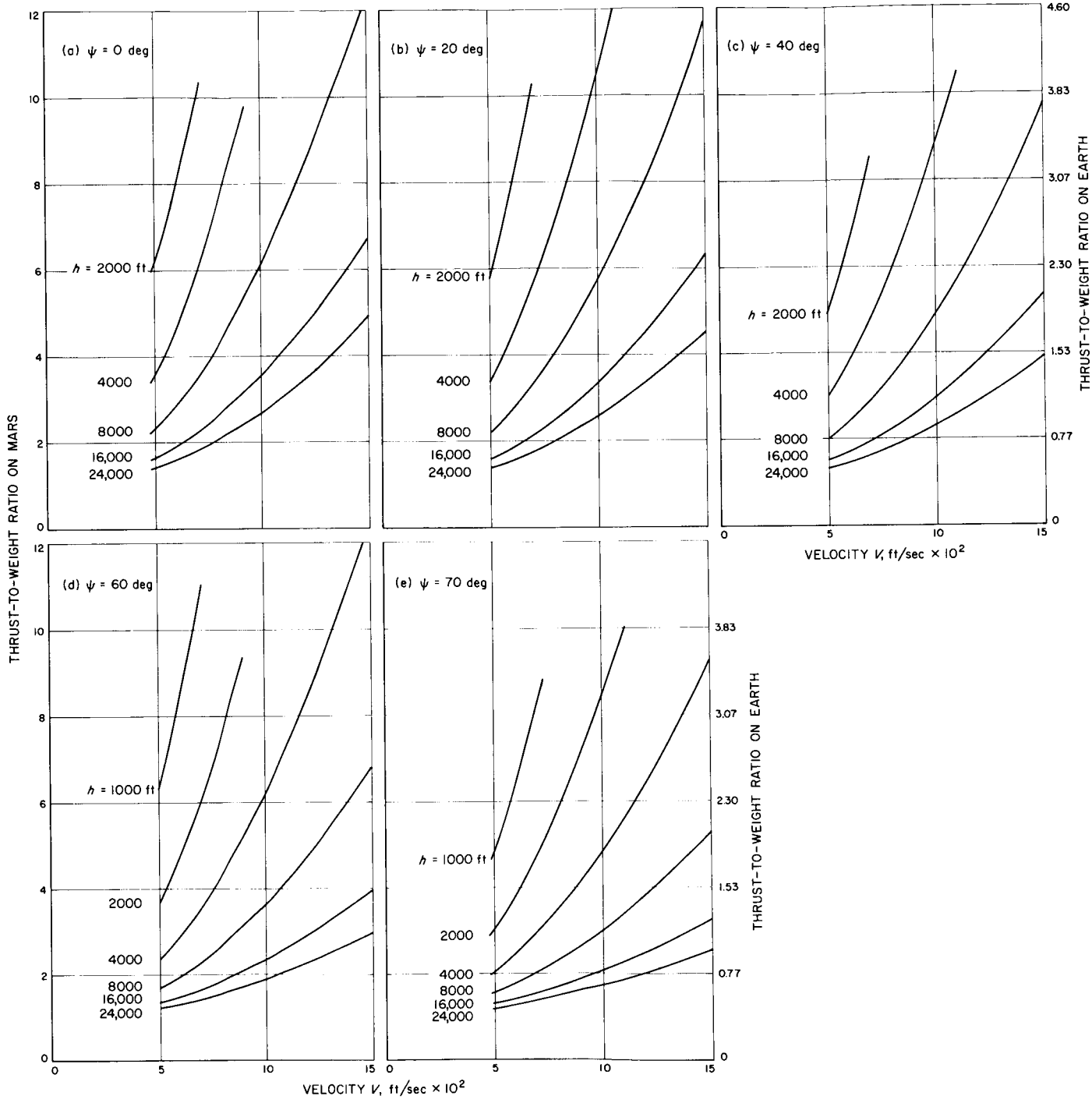


Fig. 9. Thrust-to-weight ratios on Earth and Mars as functions of velocity V

It is now possible to find the retro-thrust ΔV necessary to effect a soft landing. Since the thrust-to-weight ratio is assumed constant, the total ΔV requirement is ngT. From Eqs. (11) and (7),

$$\Delta V_R = \frac{nV_0}{(n^2 - 1)}(n + \cos\psi_0) \ . \qquad (13)$$

3. Numerical Results and Conclusions

To illustrate these results numerically, the particular example of a landing on Mars was considered. Several sets of initial conditions (V_0, h_0, ψ_0) were input in the computer program[2] written to work this example. The results are shown in Fig. 9.

In Fig. 10 are given plots of the ΔV requirement of the retromotor as a function of V. In the range of velocities considered, the ΔV requirement is nearly a linear function of V for small h. For larger values of h, the ΔV requirement increases for small V due to a longer burning time and for larger V due to a larger

[2]Described in detail by the author in "Program for Studying Gravity Turn Landing on Mars," Interoffice Memorandum 312.5-121, Jet Propulsion Laboratory, Pasadena, California, August 13, 1965.

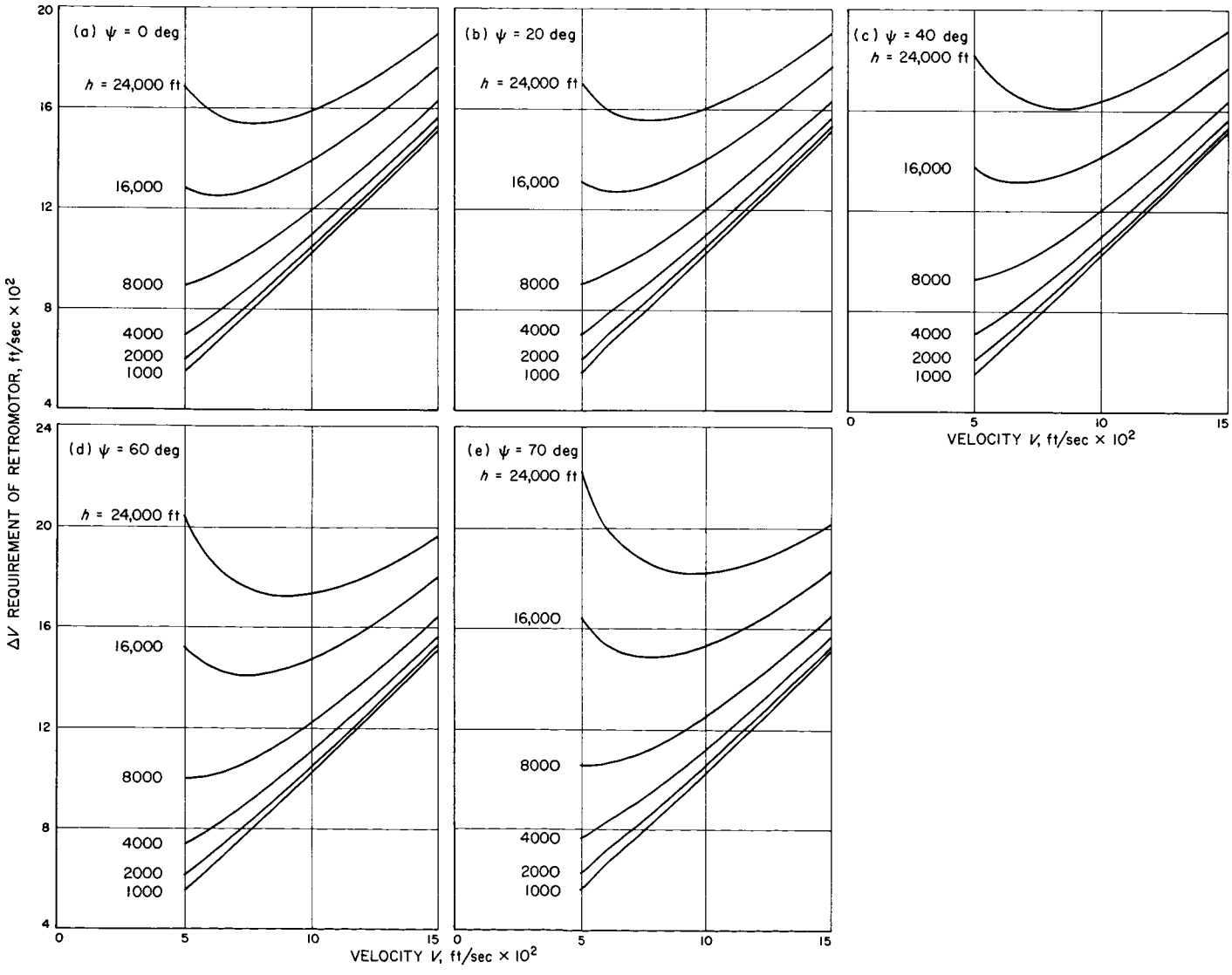


Fig. 10. ΔV requirement of the retromotor as a function of velocity V

thrust-to-weight ratio. Note also that, for a given (h_0, V_0), the ΔV requirement as a function of ψ is essentially a constant. This, the most important result of this analysis, verifies the type of approximation at small ψ that Cheng makes in Refs. 2 and 3.

C. Utilization of Energy Derived From the Gravitational Field of Jupiter for Reducing Flight Time to the Outer Solar System

G. A. Flandro

1. Introduction

Contrary to popular belief, indirect space missions using gravitational perturbations from one or more intermediate planets for trajectory shaping need not require longer flight times than those characteristic of direct missions, provided that efficient use is made of the energy which can be derived during the planetary encounters. From the point of view of a passing space vehicle, the planet appears as a field of force moving relative to the inertial coordinate system. Thus, work is done on the spacecraft, and its heliocentric energy may be increased or decreased depending upon the geometric details of the encounter. This article describes the application of energy derived in this fashion for the reduction of required launch energy and flight time for exploratory missions to the outer planets of the solar system.

The latter half of the next decade abounds in interesting multiple-planet opportunities because of the similar heliocentric longitudes of the outer planets during this time period. Trajectories using midcourse encounters with Jupiter and targeting on the outer planets Saturn, Uranus, Neptune, and Pluto were selected for detailed analysis. Many multiple-target trajectories are also possible. Of particular interest is the Earth-Jupiter-Saturn-Uranus-Neptune "grand tour" opportunity. Optimum launch dates were derived for the above trajectories with launch energies in the range $100 < C_3 < 150$ km^2/sec^2.

Multiple-planet targeting must be justified from the point of view of the encounter experiments. Each of the proposed missions involves flyby of at least two planets. Missions of this type are probably not feasible if separate instrumentation is required for each encounter. Justification for the trajectory designs presented herein is based on: (1) the great physical and chemical similarity of each of the target planets, and (2) the great similarity of approach distance and lighting conditions for each encounter.

The great communication distance involved in outer-solar-system flights causes some difficulties (Ref. 4) regarding signal propagation time, antenna size, and transmitter power requirements, for example; however, none of these should require state-of-the-art advancements of important magnitude for their solution.

The important question of guidance was not considered in the present study, but it is expected that development of planetary approach guidance techniques (Ref. 5), coupled with improved Earth-based radio guidance, should make the missions discussed herein entirely feasible by 1975. Cutting and Sturms have indicated in Ref. 6 that multiple-planet guidance problems are not as insurmountable as was once thought.

2. Heliocentric Energy Change Due to a Midcourse Planetary Encounter

Planetary encounter mechanics has been discussed elsewhere (e.g., Refs. 6–8). The purpose here is to clarify certain aspects of the problem and to define new quantities of interest in the study of multiple-planet trajectories. The mechanism of heliocentric energy change due to planetary encounter is a simple one. The planet appears to the space vehicle as a force field moving relative to the inertial heliocentric coordinate system. Thus, work is performed on the vehicle which results in an energy loss or gain relative to the Sun. Energy conservation is not violated in this process, since the energy gained or lost by the probe is lost or gained, respectively, by the planet. The percentage energy change of the planet is truly infinitesimal, while that of the vehicle may be very large.

Fig. 11 illustrates the encounter process. It can be seen that the specific energy prior to encounter is

$$E_i = \frac{\mathbf{V}_i \cdot \mathbf{V}_i}{2},$$

while that after encounter is

$$E_o = \frac{\mathbf{V}_o \cdot \mathbf{V}_o}{2},$$

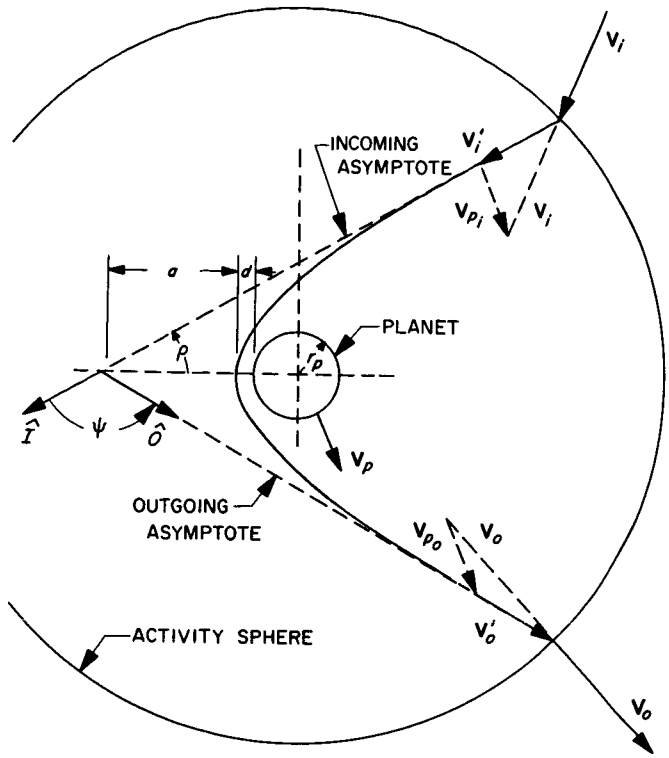


Fig. 11. Encounter hyperbola

where $\mathbf{V}_i$ and $\mathbf{V}_o$ are the pre- and post-encounter heliocentric velocity vectors of the spacecraft, respectively. Thus, the resultant energy change is

$$\Delta E = \frac{\mathbf{V}_o \cdot \mathbf{V}_o - \mathbf{V}_i \cdot \mathbf{V}_i}{2} .$$

We now introduce $\mathbf{V}'_i$ and $\mathbf{V}'_o$, the incoming and outgoing asymptotic velocity vectors relative to the planet, and note the relationship of these to the corresponding heliocentric vectors:

$$\mathbf{V}_i = \mathbf{V}_p + \mathbf{V}'_i ,$$
$$\mathbf{V}_o = \mathbf{V}_p + \mathbf{V}'_o ,$$

where $\mathbf{V}_p$ is the heliocentric velocity vector of the planet, assumed constant during the residence of the probe in the activity sphere of the planet. Assuming constant spacecraft mass and no propulsion in the activity sphere, energy is conserved relative to the planet. Thus,

$$|\mathbf{V}'_o| = |\mathbf{V}'_i| \equiv v_h ,$$

and the energy increment can be written

$$\Delta E = \mathbf{V}_p \cdot (\mathbf{V}'_o - \mathbf{V}'_i) . \qquad (1)$$

This result is only approximate for the outer planets, since $\mathbf{V}_p$ may change considerably during the encounter due to the large size of the activity spheres characteristic of these bodies (Ref. 9).

The geometrical properties of the energy equation can be clarified by setting

$$\mathbf{V}_p = v_p \hat{p} ,$$
$$\mathbf{V}'_i = v_h \hat{I} ,$$
$$\mathbf{V}'_o = v_h \hat{O} ,$$

where $\hat{p}$, $\hat{I}$, and $\hat{O}$ are unit vectors in the directions of the planet's velocity vector, incoming asymptote, and outgoing asymptote, respectively. The energy equation becomes

$$\Delta E = v_p \, v_h \left[\hat{p} \cdot (\hat{O} - \hat{I})\right] , \qquad (2)$$

where v_p is the speed of the planet and v_h is the hyperbolic excess speed of the spacecraft relative to the planet. It is evident that, for a given incoming asymptote direction, the maximum energy gain and energy loss correspond respectively to the outgoing asymptote pointing along or opposite to the planet's velocity vector. Since the vector $(\hat{O} - \hat{I})$ defines the direction of the axis of the encounter hyperbola, it can be seen that the maximum possible energy change results if the planet's velocity vector lies along this axis.

It is convenient to write the energy increment equation in the form

$$\Delta E = f E^* , \qquad (3)$$

where

$$E^* \equiv 2 v_p \, v_h \qquad (4)$$

is defined to be the characteristic energy, representing the largest energy increase for a given planet and spacecraft approach velocity. This increment is, of course, unobtainable in practice since it would require a deflection angle of 180 deg, which would in turn require passage of the vehicle very close to the center of the planet. Plots of E^* versus v_h, based on mean values of v_p, are shown in Fig. 12 for the planets of the solar system. As expected, Mercury has the highest value of E^* for a given v_h due to its close proximity to the Sun and corresponding high heliocentric velocity. The coefficient

$$f \equiv \frac{\hat{p} \cdot (\hat{O} - \hat{I})}{2} , \qquad (5)$$

defined as the energy change index, is a number between -1 and 1 which determines the magnitude and sign of

the energy change. The value of f depends greatly on the ability of the planet to deflect the trajectory. The magnitude of the deflection angle

$$|\psi| = \cos^{-1}(\hat{I} \cdot \hat{O}), \tag{6}$$

as measured positive counterclockwise from $\hat{I}$, can be written in terms of the gravitational constant of the planet, the hyperbolic excess speed, and the distance of closest approach to the surface as

$$\psi = 2\sin^{-1}\left[1 + \frac{v_h^2}{\mu}(d + r_p)\right]^{-1}, \tag{7}$$

with r_p as the radius of a sphere enclosing the planet such that

$$r_p = R_p - \delta a + \delta g, \tag{8}$$

where R_p is the equatorial radius of the planet, δa is the maximum depth of sensible atmosphere, and δg is the guidance error contingency. The maximum possible deflection angle corresponds to a flight path grazing the forbidden sphere. Thus,

$$\psi_{max} = 2\sin^{-1}\left(\frac{\mu}{\mu + v_h^2 r_p}\right). \tag{9}$$

Fig. 13 shows plots of ψ_{max} versus v_h for all planets of the solar system except Pluto, whose gravitational constant is not precisely known. The calculations were made for $\delta a = \delta g = 0$. As expected, Jupiter is capable of the largest flight path deflection for a given asymptotic speed due to its great mass.

The maximum possible energy change index for a given incoming asymptotic direction can now be computed.

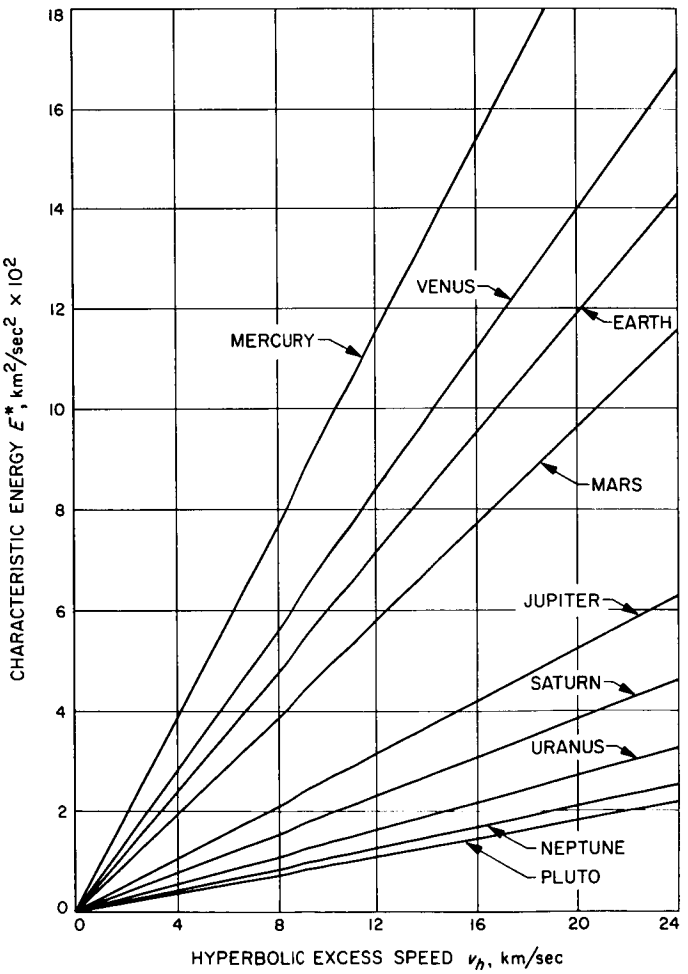


Fig. 12. Characteristic energy vs hyperbolic excess speed of spacecraft

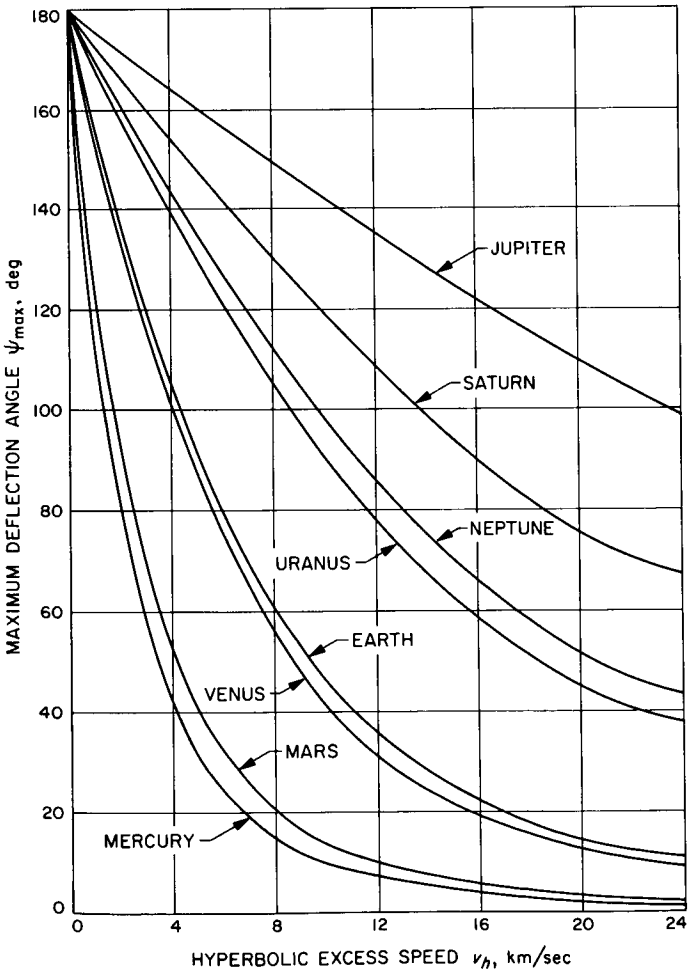


Fig. 13. Maximum deflection angle vs hyperbolic excess speed of spacecraft

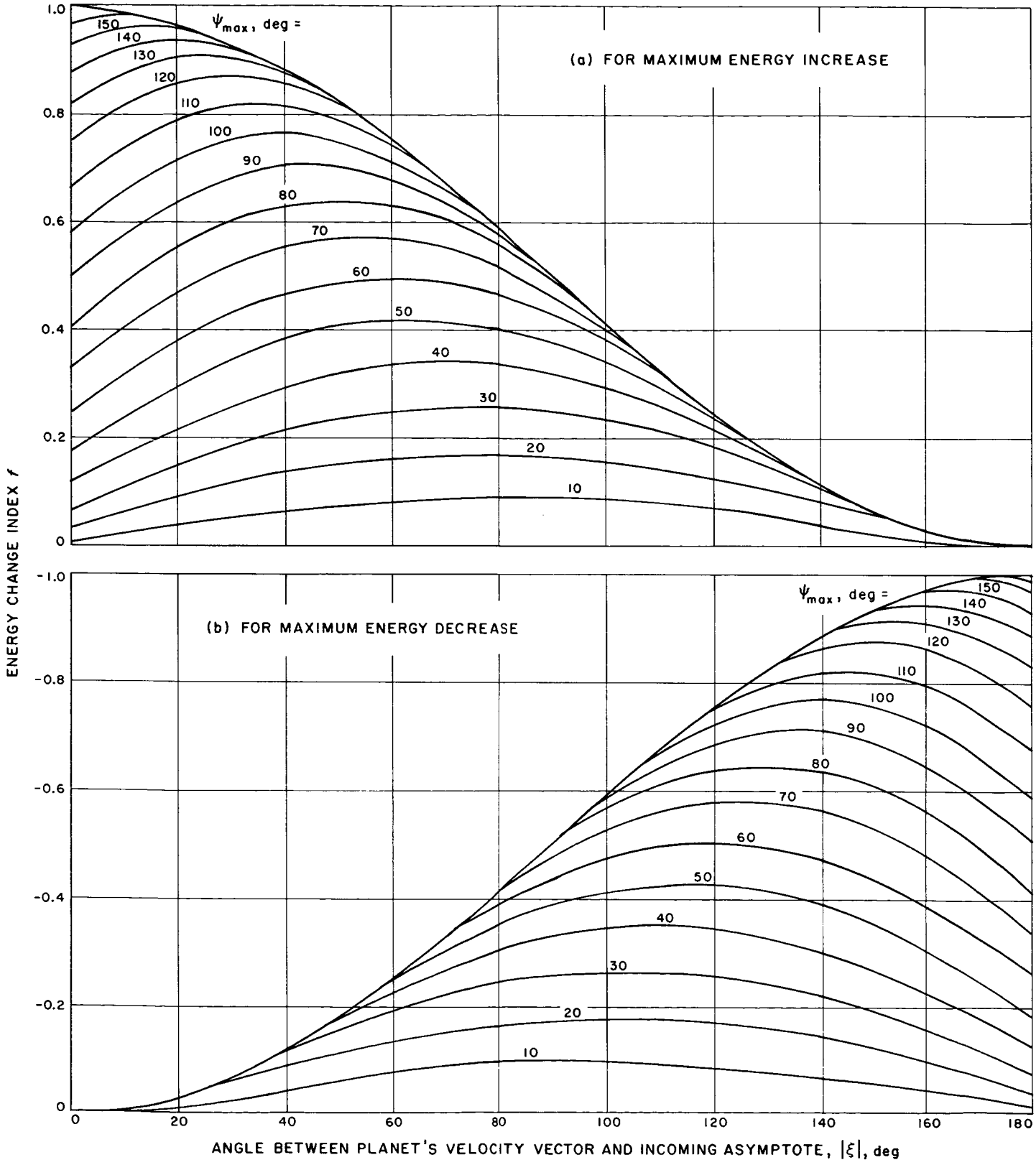


Fig. 14. Energy change index vs angle between planet's velocity vector and incoming asymptote

The angle between the encounter planet's velocity vector and the incoming asymptote is denoted as ξ, where

$$|\xi| \equiv \cos^{-1}(-\hat{I} \cdot \hat{P}) . \tag{10}$$

ξ is measured positive counterclockwise from $\hat{P}$ to $\hat{I}$, as viewed from the north pole of the celestial sphere. The maximum energy gain which can be achieved corresponds to

$$f_{\text{max energy gain}} = \begin{cases} \dfrac{\cos\xi + 1}{2}, & |\xi| \geq \pi - \psi_{\max}, \\ \dfrac{\cos\xi - \cos(\psi_{\max} + \xi)}{2}, & |\xi| < \pi - \psi_{\max}. \end{cases} \tag{11}$$

Values of f for maximum energy loss relative to the Sun for a given incoming asymptote direction are

$$f_{\text{max energy loss}} = \begin{cases} \dfrac{\cos\xi - 1}{2}, & |\xi| \leq \psi_{\max}, \\ \dfrac{\cos\xi - \cos(\psi_{\max} - \xi)}{2}, & |\xi| > \psi_{\max}. \end{cases} \tag{12}$$

These relationships are plotted in Fig. 14 for the cases of energy increase and energy decrease.

It is of interest to assess the abilities of the various planets to influence the trajectories of passing space vehicles. This can be accomplished by comparing the maximum available energy increments at given values of approach speed v_h. It is often assumed that Jupiter, because of its great mass, is capable of producing the largest energy change. However, this is not the case under some conditions; e.g., for low hyperbolic excess speeds, say $v_h < 3$ km/sec, the much smaller planet Mars is capable of a larger energy increment than Jupiter. This effect is the result of the higher heliocentric speed of planets which are closer to the Sun. Thus, E^* dominates at low approach speeds. At higher speeds, the ability of a more massive planet to deflect the trajectory can be utilized. Therefore, the values of the energy change index f are larger, and a larger energy increment results. Fig. 15 shows the maximum possible energy change which can be produced by encounters with the planets of the solar system. Only in exceptional cases would it be possible to utilize the maximum available energy boost because of the geometric constraints imposed by the required approach and departure asymptotic velocity vectors. The trajectory designer must endeavor to find those flight paths linking the launch, midcourse, and terminal planets which can utilize as large a percentage of the available energy as possible in order to reduce the over-all launch

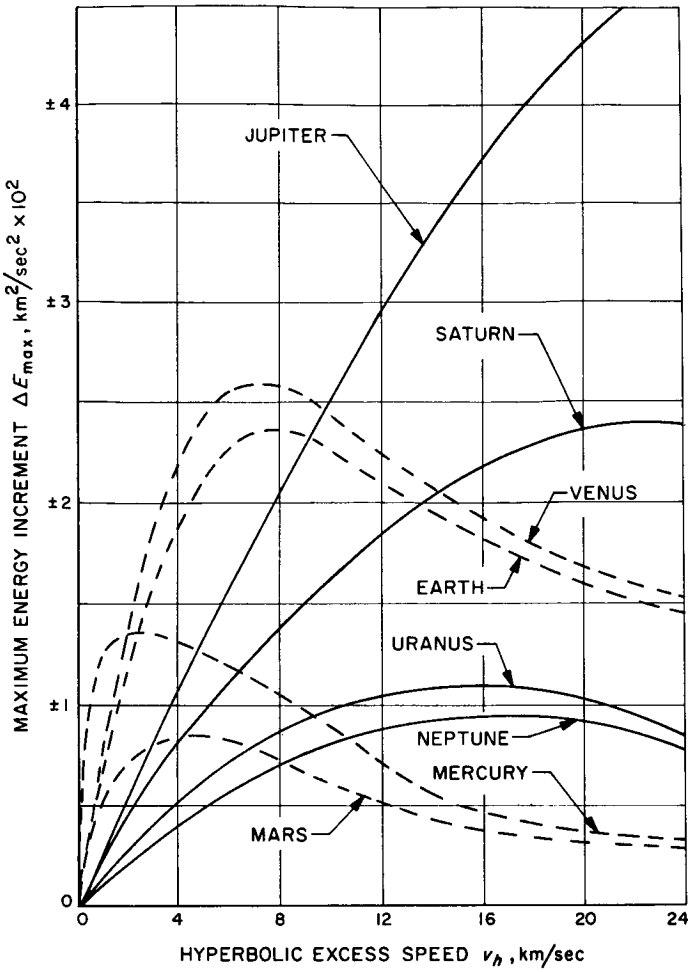


Fig. 15. Maximum energy increment vs hyperbolic excess speed of spacecraft

energy and flight time requirements of the mission. A measure of his success is the factor of merit

$$\text{FM} = \frac{f}{f_{\text{max energy change}}}, \tag{13}$$

where f is the energy change index actually achieved in the encounter, and $f_{\text{max energy change}}$ is the theoretical maximum energy gain or loss as given by Eqs. (11) and (12). Eq. (11) is used if f is positive. Eq. (12) is used if an energy loss is involved.

3. Trajectory Results

Conic trajectory programs were utilized in studying multiple-planet opportunities for the latter half of the next decade with Jupiter as an intermediate planet. Vehicles with launch energies in the range $100 < C_3 < 150$ km²/sec² were assumed. It was discovered that, for

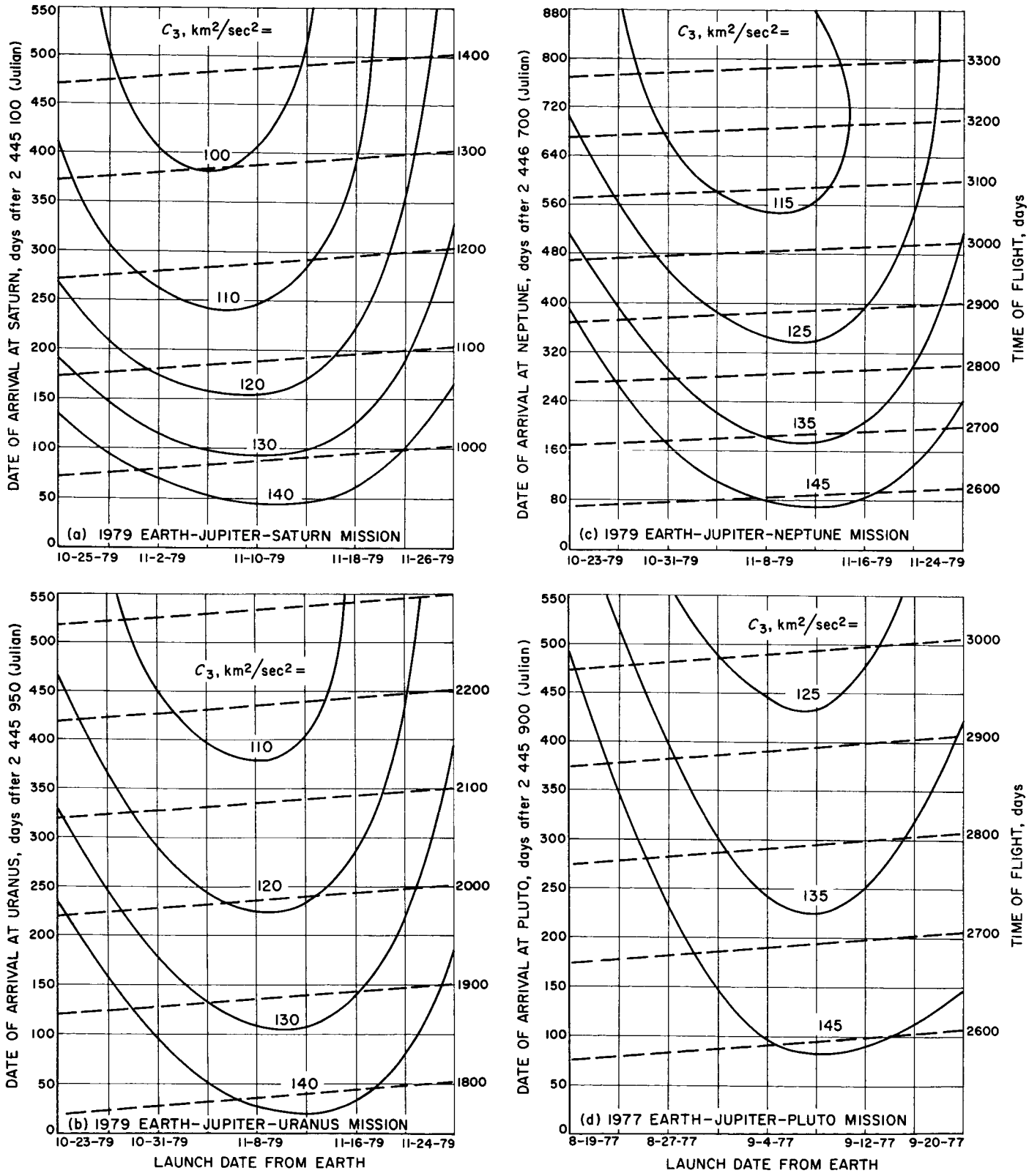


Fig. 16. Constant launch energy contours for 1979 Earth–Jupiter–Saturn, 1979 Earth–Jupiter–Uranus, 1979 Earth–Jupiter–Neptune, and 1977 Earth–Jupiter–Pluto missions

a given mission, several consecutive Earth–Jupiter launch opportunities resulted in acceptable trajectories to the outer planets; however, only the optimum launch years are discussed here.

a. Earth–Jupiter–Saturn trajectory characteristics. Typical arrival date versus launch date plots are shown in Fig. 16(a) for the best 30-day launch period in 1979. The solid curves in Fig. 16(a) are constant launch energy contours, while the dashed lines represent the flight time T_f. Good opportunities are available in 1978 and 1980 also, but these require about 10% more flight time than the 1979 trajectories. Fig. 17(a) shows a typical 1979 Earth–Jupiter–Saturn flight path projected in the plane of the ecliptic. It will be noted that the trajectory after Saturn flyby is depicted as escaping from the solar system. The actual post-Saturn flight path would depend on the targeting conditions at Saturn selected by the designer, but solar system escape energy is available. (This is true of all trajectories considered in this article.)

The required closest approach distances at Jupiter for the 1979 Earth–Jupiter–Saturn trajectories can be derived from Fig. 18(a). None of the trajectories considered required passage closer than 2 Jupiter radii from the surface. It is of interest to assess the trajectories using the theory discussed above. For a typical 1979 trajectory with $C_3 = 140$ km^2/sec^2, the maximum available energy change index is

$$f_{\text{max energy increase}} \approx 0.69\,,$$

while that actually achieved is $f = 0.61$. Thus, the factor of merit for these trajectories is typically FM = 0.88, which indicates good usage of the energy available from the Jupiter encounter. Earth–Jupiter–Saturn launch opportunities recur with a period of about 20 yr.

b. Earth–Jupiter–Uranus trajectory characteristics. Launch opportunities to Uranus by way of Jupiter are available in 1978, 1979, and 1980. The 1979 flights are far superior with regard to required flight time for a given launch energy. Fig. 16(b) shows constant launch energy contours for the 1979 opportunity, and Fig. 17(b) shows the planetary configuration typical of these missions. The required distance of closest approach to Jupiter is quite small, as indicated in Fig. 18(b). High energy trajectories approach to within one-half of a Jupiter radius from the surface of the planet. The factor of merit for a 1979 launch with $C_3 = 110$ km^2/sec^2 is FM = 0.97, while that corresponding to a launch energy of $C_3 = 140$ km^2/sec^2 is typically FM = 0.99. Thus, efficient use is made of the available encounter energy increment. Uranus opportunities by way of Jupiter are available about every 14 yr.

c. Earth–Jupiter–Neptune trajectory characteristics. 1979 appears to be the optimum launch date for trajectories to Neptune by way of Jupiter also. This is the result of the similar heliocentric longitudes of Saturn, Uranus, and Neptune during the latter half of the next decade. Arrival date versus launch date curves for the 1979 opportunity are shown in Fig. 16(c) for several launch energies. Fig. 17(c) shows the typical planetary constellation for the mission. Plots of closest approach distance at Jupiter versus launch date are shown in Fig. 18(c). A typical factor of merit for this mission is FM = 0.88, indicating good use of the Jupiter encounter energy boost. Earth–Jupiter–Neptune opportunities exhibit a periodicity of about 13 yr.

d. Earth–Jupiter–Pluto trajectory characteristics. The optimum launch year for Earth–Jupiter–Pluto missions was found to be 1977. Fig. 16(d) shows the arrival date versus launch date plots for this opportunity, and Fig. 17(d) illustrates the corresponding planetary constellation. Fig. 18(d) shows the distance of closest approach required for deriving the necessary energy during the Jupiter encounter. A high-energy transfer orbit with a flight time of about 7 yr would require the vehicle to pass within 2 radii of Jupiter's surface. The short flight times to Pluto are the result of efficient usage of the available energy increment (FM = 0.91), coupled with the fact that Pluto is near perihelion at closest approach of the probe. Although launch opportunities are available every 12.5 yr, the optimum conditions characteristic of the 1977 opportunity will not be repeated until the year 2224. The 1990 launch opportunities should also be favorable, however.

e. Earth–Jupiter–Saturn–Uranus–Neptune "grand tour" trajectory characteristics. Many interesting multiple-target trajectories to the major planets are available during the latter half of the next decade due to the similar heliocentric longitudes of these bodies during that time period. Typical of these are the Earth–Jupiter–Saturn–Uranus–Neptune opportunities. Fig. 19 shows the arrival date at each planet versus the launch date from Earth for the best 30-day launch period. 1979 was again found to be the optimum launch year. Fig. 20 shows the planetary constellation which is typical of the mission. The required flight time from Earth to the terminal planet Neptune with launch energy of $C_3 = 140$ km^2/sec^2 is

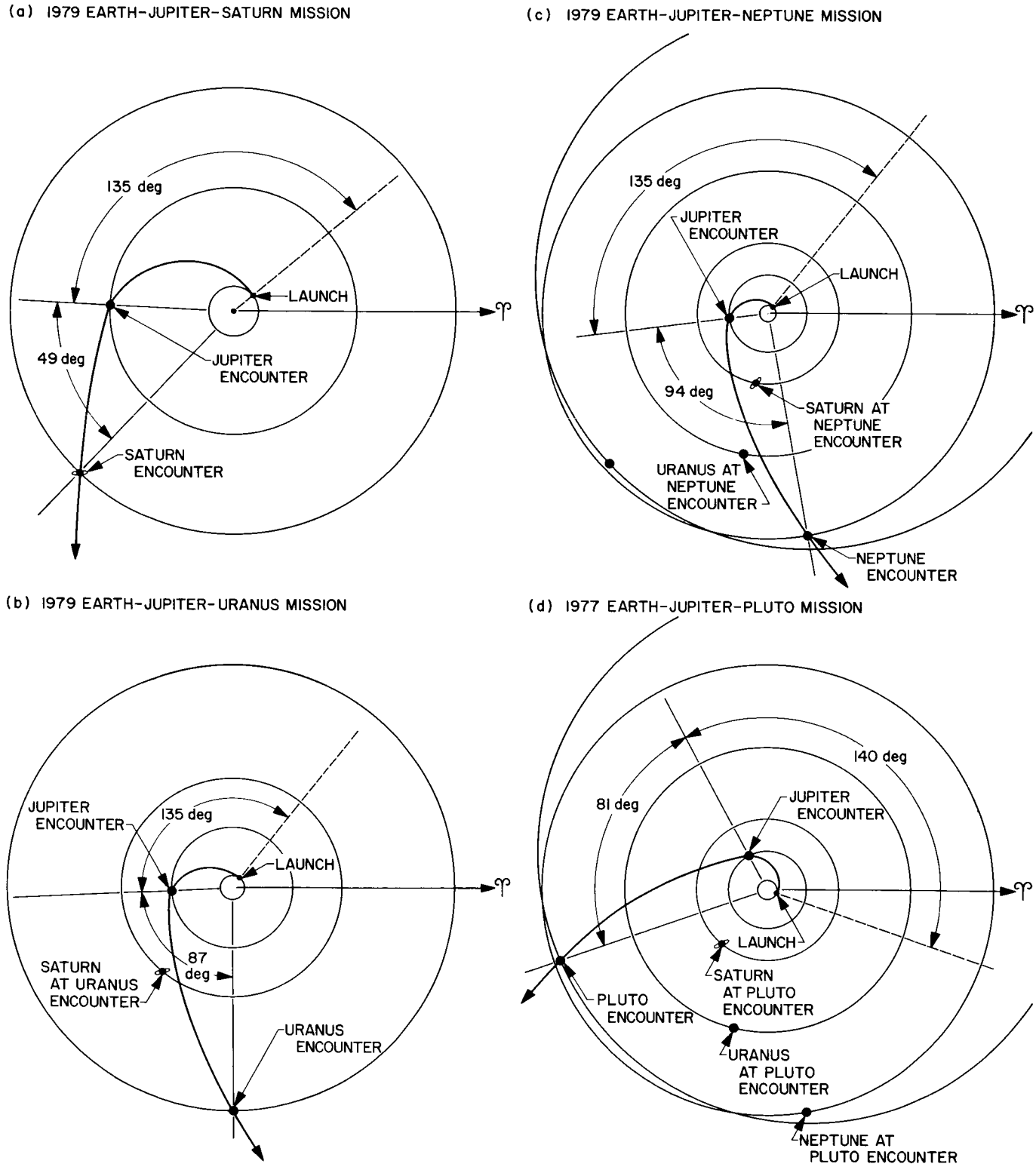


Fig. 17. Typical planetary configurations for 1979 Earth–Jupiter–Saturn, 1979 Earth–Jupiter–Uranus, 1979 Earth–Jupiter–Neptune, and 1977 Earth–Jupiter–Pluto missions

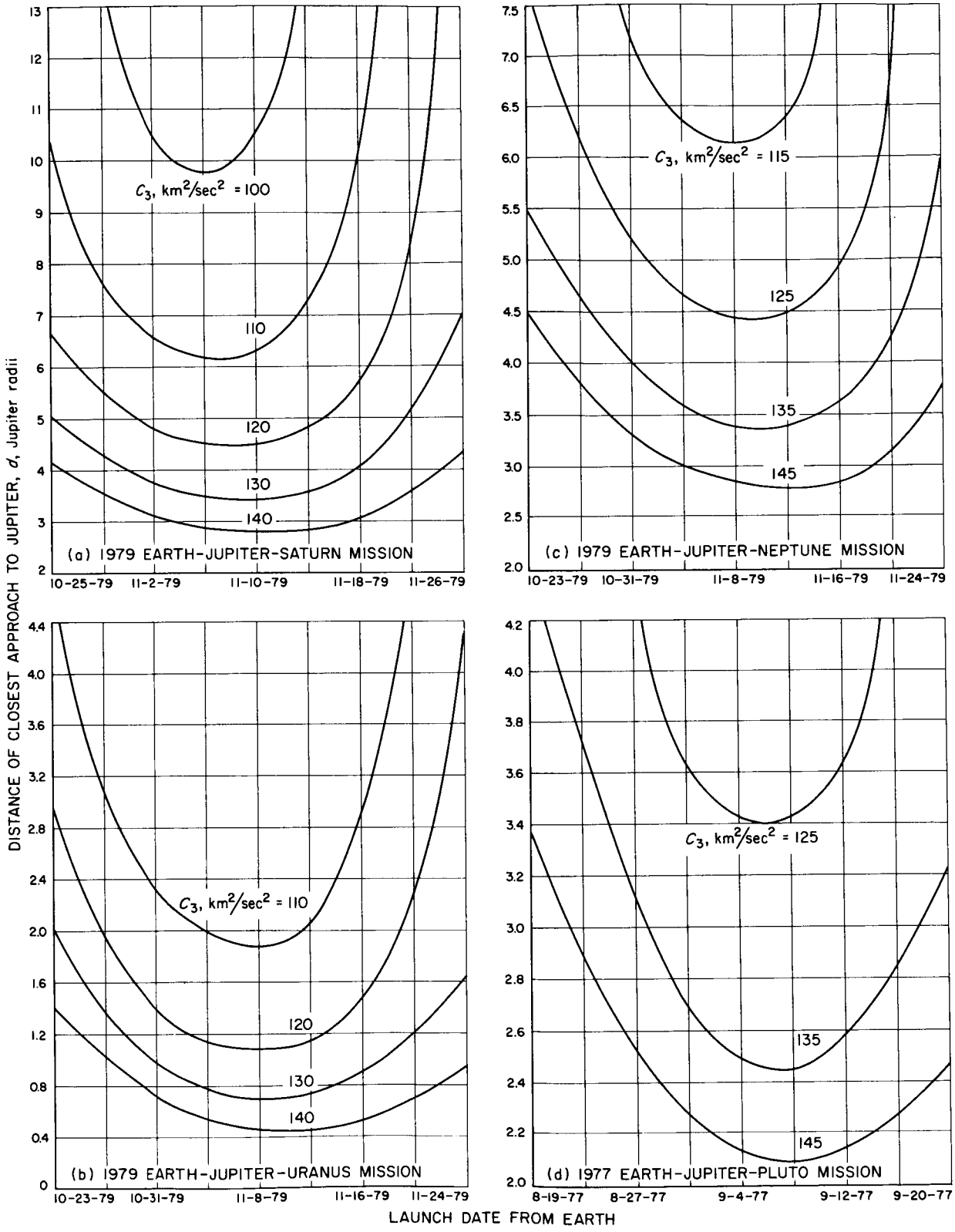


Fig. 18. Distance of closest approach to Jupiter vs launch date from Earth for 1979 Earth–Jupiter–Saturn, 1979 Earth–Jupiter–Uranus, 1979 Earth–Jupiter–Neptune, and 1977 Earth–Jupiter–Pluto missions

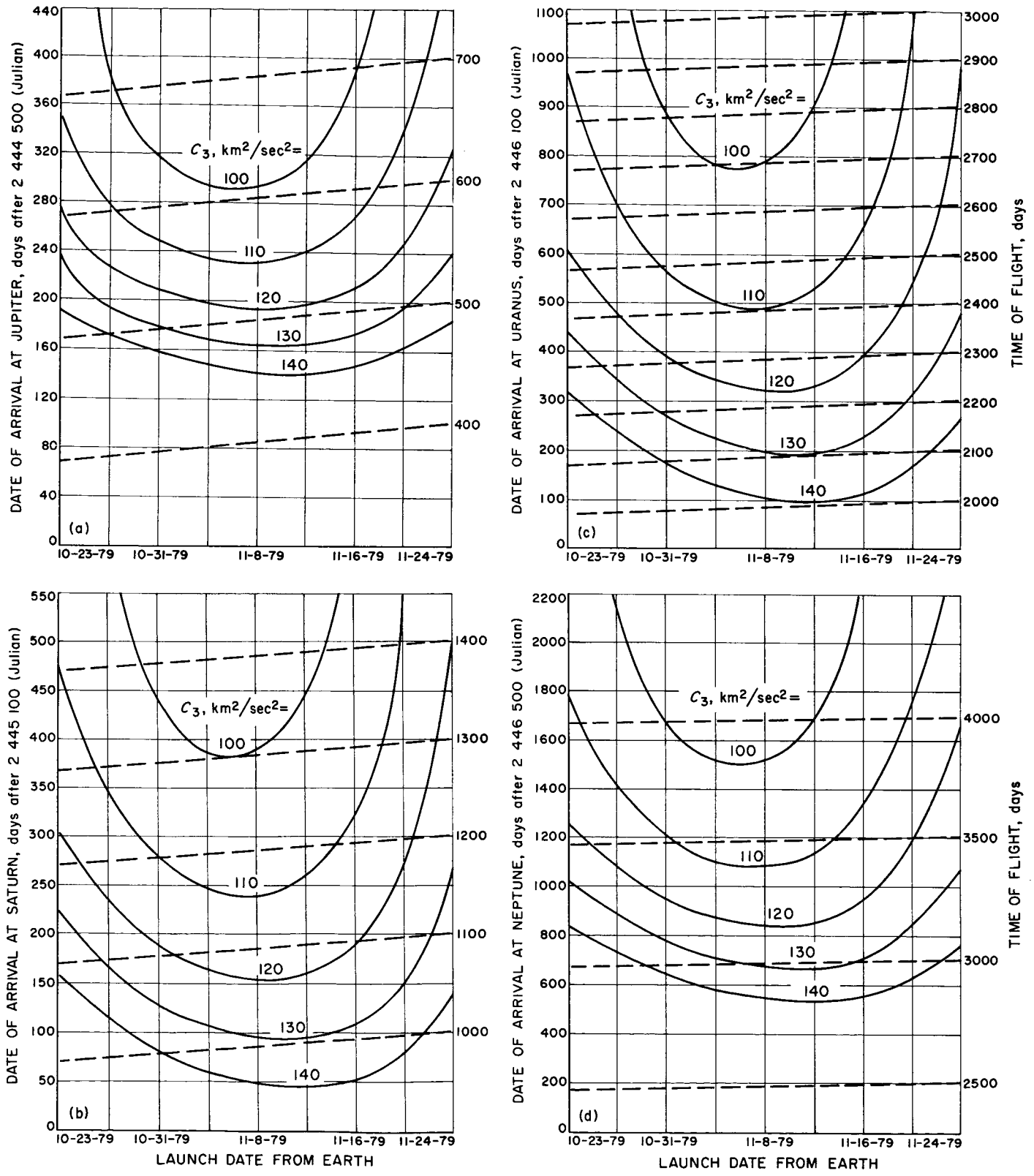


Fig. 19. Date of arrival at planet vs launch date from Earth for a 1979 Earth–Jupiter–Saturn–Uranus–Neptune mission

Table 3. Trajectory parameters for 1979 Earth–Jupiter–Saturn, 1979 Earth–Jupiter–Uranus, and 1979 Earth–Jupiter–Neptune missions

Type of trajectory	Earth–Jupiter–Saturn mission		Earth–Jupiter–Uranus mission		Earth–Jupiter–Neptune mission	
	C_3, km²/sec²	T_f, yr	C_3, km²/sec²	T_f, yr	C_3, km²/sec²	T_f, yr
Direct to terminal planet; minimum energy	130	4.1	130	7.6	135	20.5
Indirect (multiple-planet)	100	3.8	100	6.3	110	8.8
	140	2.9	140	4.9	140	7.2
Direct to terminal planet, using corresponding T_f and C_3 of multiple-planet mission	135	—[a]	155	—[a]	190	—[a]
	150	3.3	180	7.7	230	17.8

[a] Mission cannot be accomplished under these conditions.

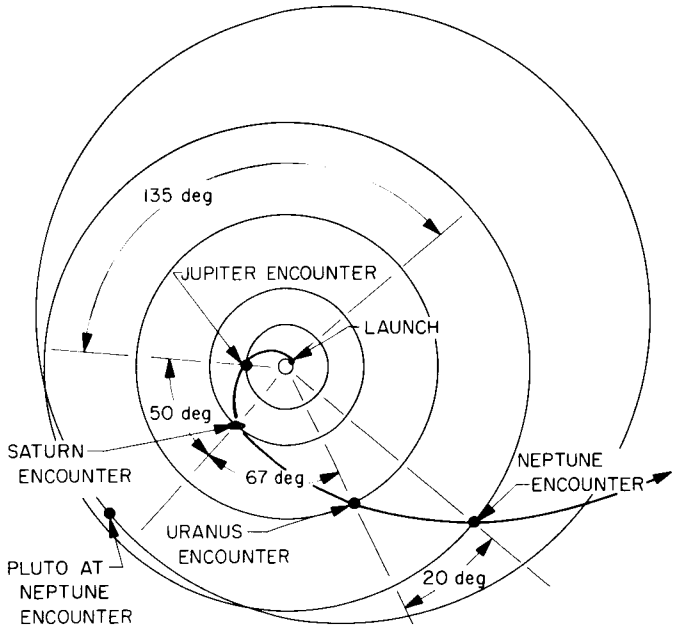


Fig. 20. Planetary constellation for a 1979 Earth–Jupiter–Saturn–Uranus–Neptune mission

about 7.8 yr, which compares favorably with the 7.2 yr for an Earth–Jupiter–Neptune mission with the same launch energy. The longer flight time for the four-planet mission is the result of a loss of heliocentric energy in the Uranus encounter due to a requirement for the vehicle to pass in front of the planet to achieve the proper Uranus–Neptune flight path. The periodicity of Earth–Jupiter–Saturn–Uranus–Neptune missions is about 175 yr due to the constraints imposed by the required positions of Uranus and Neptune.

4. Conclusions

The 1975–1980 time period is characterized by many interesting multiple-planet trajectories which efficiently utilize energy derived from a close approach to the planet Jupiter. 1979 appears to be an especially favorable year for missions to the outer planets using launch vehicles capable of energies in the range $100 < C_3 < 150$ km²/sec².

The trajectories discussed here are characterized by very short flight times in comparison to those for direct flights from Earth to the target planets considered. Characteristics of direct and multiple-planet trajectories are compared in Table 3. Although higher launch energies are suggested for some of the multiple-planet flights, the additional cost of this energy might be offset by the great savings afforded by the short flight times. On the other hand, some multiple-planet missions can reach some targets with *less* launch energy than would be required for a direct "minimum energy" flight.

References

1. Minovitch, M. A., *The Determination and Characteristics of Ballistic Interplanetary Trajectories Under the Influence of Multiple Planetary Attractions,* Technical Report No. 32-464, Jet Propulsion Laboratory, Pasadena, California, October 31, 1963.

2. Cheng, R. K., *Surveyor Terminal Guidance,* Hughes Aircraft Company, Space Systems Division, El Segundo, California, April 7, 1965.

3. Cheng, R. K., *Application of Surveyor Soft-landing Techniques to a Mars Probe,* Hughes Aircraft Company, Space Systems Division, El Segundo, California, September 8, 1964.

4. Kirsten, C. C., "Data Acquisition at Planetary Ranges," *Proceedings of the NASA-University Conference on the Science and Technology of Space Exploration,* Vol. 1, NASA SP-11, 1962.

5. Gates, C. R., and Gordon, H. J., "Planetary Approach Guidance," *Journal of Spacecraft and Rockets,* Vol. 2, No. 2, pp. 160–166, March–April 1965.

6. Cutting, E., and Sturms, F. M., Jr., *Trajectory Analysis of a 1970 Mission to Mercury Via a Close Encounter with Venus,* Technical Memorandum No. 312-505, Jet Propulsion Laboratory, Pasadena, California, December 7, 1964.

7. Minovitch, M. A., *Determination and Characteristics of Ballistic Interplanetary Trajectories Under the Influence of Multiple Planetary Attractions,* Technical Report No. 32-464, Jet Propulsion Laboratory, Pasadena, California, October 31, 1963.

8. Minovitch, M. A., *Utilizing Large Planetary Perturbations for the Design of Deep-Space, Solar-Probe, and Out-of-Ecliptic Trajectories,* Technical Memorandum No. 312-514, Jet Propulsion Laboratory, Pasadena, California, February 15, 1965.

9. Ehricke, K. A., *Space Flight, Vol. I,* D. Van Nostrand Company, Inc., New York, 1960.

II. Scientific Programming

A. Triangulation of Plane Point Sets

C. L. Lawson

A study has been initiated for the development of computer algorithms for performing certain numerical operations on empirical data which are available at irregularly located points in a plane. Applications considered include interpolation, quadrature, and the construction of contour maps. A construct which appears to be useful in this connection is the triangulation of a planar point set.

Let P be a set of n distinct points $p_1, \cdots, p_n$ in the xy-plane, not all of which lie on the same straight line. Let H be the convex hull of P; i.e., let H be the smallest convex polygon containing P. A triangulation, τ, of P will be defined to be a partition of H into nondegenerate triangles such that each triangle contains exactly three points of P, one point at each vertex. For a given point set P, there may be many possible triangulations, although each triangulation of P will have the same number of triangles and the same number of line segments (Ref. 1). This latter fact, of significance in the design of a computer algorithm, may be verified as follows:

Let n be the number of points in P; n_b, the number of points of P on the boundary of H; and $n_i = n - n_b$. Let t and s be the number of triangles and line segments, respectively, in a triangulation τ. The n_b line segments lying in the boundary of H each belong to one triangle, while the remaining $(s - n_b)$ line segments each belong to two triangles. It follows that

$$3t = n_b + 2(s - n_b) .$$

Using Euler's formula $(t - s + n = 1)$, the formulas

$$t = n_b + 2n_i - 2, \quad s = 2n_b + 3n_i - 3$$

result; these express t and s in terms of quantities which depend only on P and not on the particular triangulation τ.

An upper bound for the number of different triangulations of a point set P is given by

$$\frac{2(2n_b - 3)!\,(4n_i + 2n_b - 5)!}{(n_b - 1)!\,(n_b - 3)!\,(3n_i + 2n_b - 3)!\,n_i!} .$$

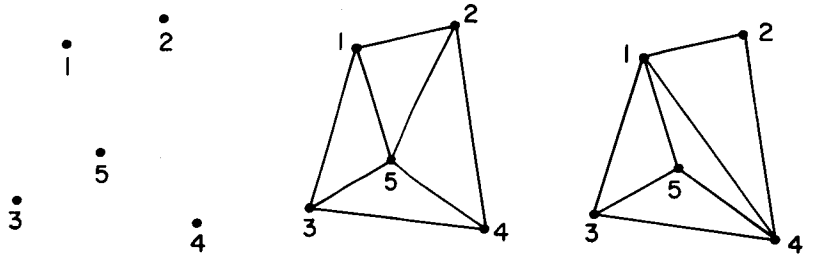
This formula was obtained from Ref. 2, where it was derived as the exact solution of a slightly different enumeration problem.

Some probabilistic results relating to this problem are given in Ref. 3. For instance, if R is a bounded convex plane region and if the point set P consists of n points selected at random from a uniform distribution in R, then the expected value of n_b is $0(n^{1/3})$ if R has a smooth boundary and $0(\log n)$ if R is a polygon.

B. A Transformation of Triangulations

F. W. Weingarten

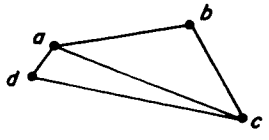
Let a triangulation of a plane point set be defined as in the preceding article "Triangulation of Plane Point Sets." Then the point set illustrated in Part (a) below permits the two triangulations (b) and (c).



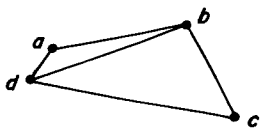
(a) (b) (c)

Consider the quadrilateral formed by points 1, 2, 4, and 5. The diagonal of this quadrilateral in Part (b), line $<5,2>$, may be exchanged with a new diagonal $<1,4>$ to give the triangulation shown in (c).

Proceeding from this example, we can define a simple transformation of triangulations. We define a quadrilateral to be *strictly convex* if it is convex and if no three vertices lie in the same straight line. If a,b,c,d is a strictly convex quadrilateral with diagonal $<a,c>$ as shown in Part (a) below, we define a T-transform as the exchange of diagonal $<b,d>$ for $<a,c>$, as shown in (b).



(a)



(b)

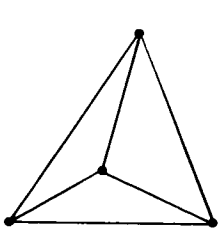
Now any triangulation of a set of points contains as many quadrilaterals as there are internal line segments, for any internal line segment forms the boundary between two triangles. The remaining sides of these triangles form a quadrilateral. Thus, we can talk about T-transforms of a general triangulation.

***Theorem:** Let τ_1 be a triangulation of a set of points and τ_2 be any other triangulation of the same set of points. Then, starting with τ_1, the triangulation τ_2 may always be obtained by a finite number of T-type transformations.*

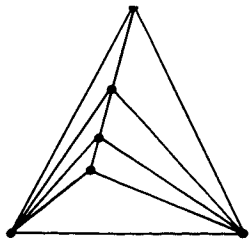
This theorem, proved during a study of interpolation of data in two dimensions, has several implications for an algorithmic study of triangulations: (1) It puts forth a surprisingly simple transformation as a universal "stepping stone," so to speak; and (2) it provides a scheme by which all triangulations may be determined in an algorithmic manner. A direct corollary is as follows:

***Corollary:** Let τ be a triangulation of a set of points such that no T-transform is possible. Then τ is the unique triangulation of these points.*

An example of this corollary is given in Part (a) below, and any extension as in Part (b) is also an example.

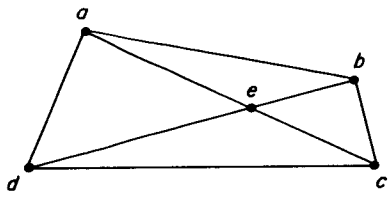


(a)



(b)

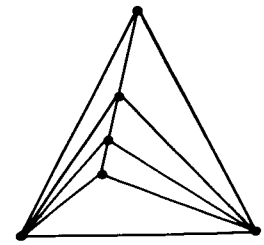
Aside from the general form shown in Part (b) above, only one other distribution of points has been found for which the triangulation is unique. Let a,b,c,d be a strictly convex quadrilateral, and let e be the intersection point of the diagonals $<a,c>$ and $<b,d>$. Then, the set of points (a,b,c,d,e) has the unique triangulation shown below:



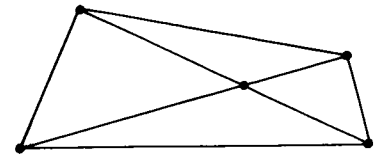
We may conclude with the following conjecture:

Conjecture: Let P be a set of points such that no three points in the boundary of the convex hull of P are co-linear. Then, if n_b is the number of points in the boundary of the convex hull, we have the following conditions for a unique triangulation:

(1) ***If $n_b = 3$, the points have the following arrangement (all interior points and a vertex lie on a line):***



(2) ***If $n_b = 4$, the points have the following arrangement (one interior point on the intersection of the diagonals):***



(3) ***If $n_b = 5$, there is no arrangement of points for which the triangulation is unique.***

References

1. Tutte, W. T., "A Census of Planar Triangulations," *Canadian Journal of Mathematics*, Vol. 14, pp. 21–38, 1964.

2. Brown, W. G., "Enumeration of Triangulations of the Disk," *Proceedings of the London Mathematical Society*, Vol. 3, pp. 746–768, 1964.

3. Rényi, A., and Sulanke, R., "Über die konvexe Hulle von n zufällig gewählten Punkten," *Z. Wahrscheinlichkeitstheorie and Verw. Gebiete*, Vol. 2, pp. 75–84, 1963, and Vol. 3, pp. 138–147, 1964 (*Mathematics Reviews*, p. 1185, June 1964, and p. 1200, June 1965).

GUIDANCE AND CONTROL DIVISION

III. Spacecraft Power

A. Switched Regulator Microcircuits

W. K. Shubert

1. Introduction

The switched regulator's high efficiency is ideally suited to space power applications. High efficiency over a wide input voltage variation allows the power source to be smaller. Low internal heat generation does not require a heavy heat sink with high heat conductivity.

There are two disadvantages to the switched regulator: the heavy *L-C* filter and the circuit complexity required to generate a variable duty cycle pulse train. For constant output ripple, the filter can be made acceptably small by increasing the switching frequency, but changing switching frequency does not change the large parts count in the low-level circuitry. This complexity would be less objectionable if it could be fabricated as a microcircuit module.

Microcircuitry has many advantages. The reliability of a microcircuit approaches that of a single transistor. Fabrication of the regulator would involve the purchase and screening of the microcircuit, its installation in a chassis, and the addition of a few higher power components. The size and weight would be considerably reduced in comparison to conventional circuit assemblies.

2. Microcircuit Design Requirements

The objective of the present program is to put all of the low-level circuitry for a switched regulator in one small container. The device should use as few silicon chips as possible, a maximum of three. For utility in more than one design, it has to be useful in both booster regulators and down regulators. The microcircuit will be enclosed in a hermetically sealed container with welded leads.

The graphical representations of down regulators (Fig. 1) and booster regulators (Fig. 2) show that both have the

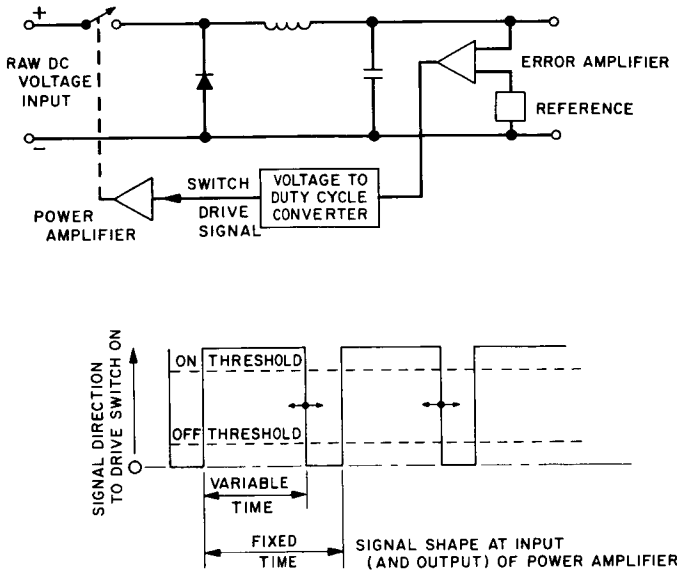


Fig. 1. Down-regulator functional diagram

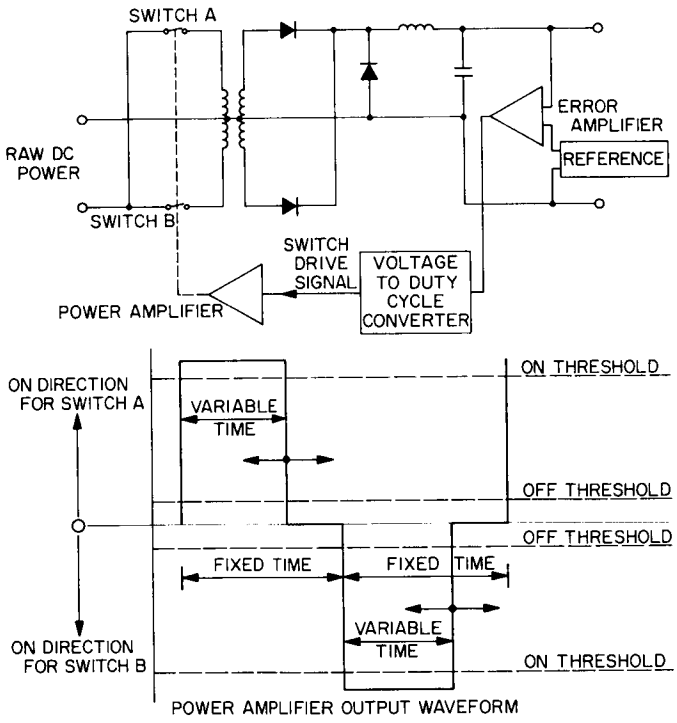


Fig. 2. Booster functional diagram

same low-level functions. These are the voltage reference, the error amplifier, and the analog voltage to duty cycle converter. The basic difference is the output waveform of the duty cycle generator. For use in either regulator type, the device must accommodate both signals. If the

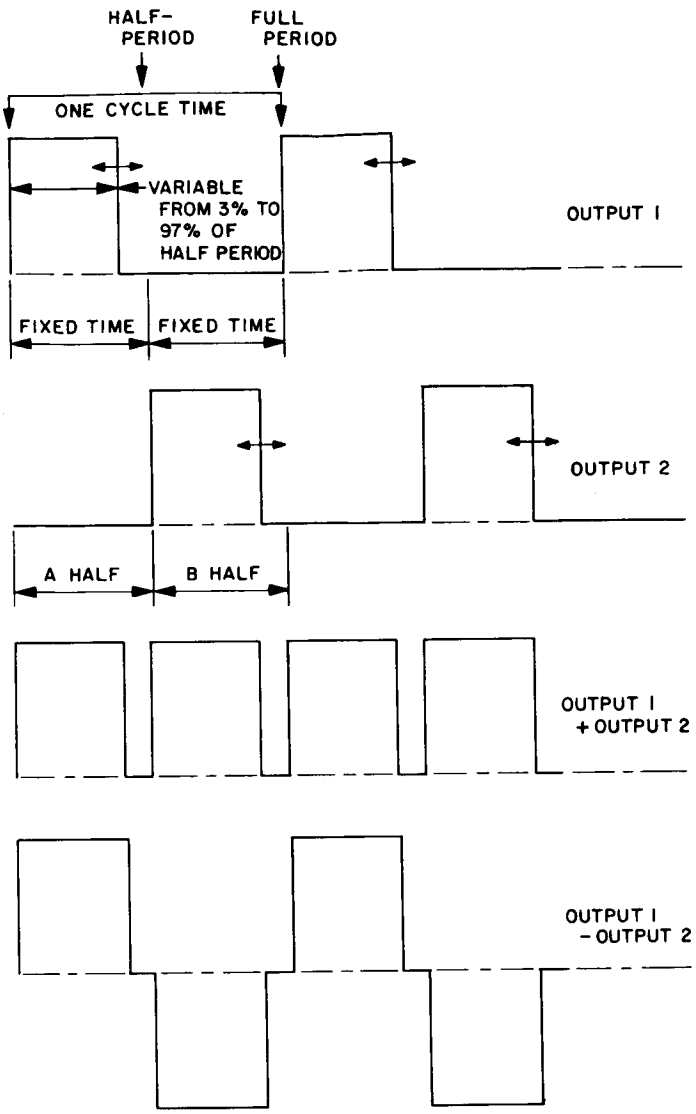


Fig. 3. Generation of booster and down-regulator waveforms

output of the functional block is composed of two outputs as shown in Fig. 3, either type of signal could be derived by proper external manipulation. Therefore, this output scheme was selected.

Since space applications require a basic regulation of at least $\pm 1\%$, the device must have enough gain to yield this regulation. In order to accomplish this regulation, the pulse output waveform should change from almost 100% to nearly 0% when the signal input to the device varies $\pm 1\%$. In order to allow for temperature variation and line voltage drops, $\pm 0.5\%$ signal change was specified for an output duty cycle change from 3% to 97%. The complete electrical specifications are listed in Table 1.

3. Device Configuration

Normally, to obtain a duty cycled output waveform electronically, an analog ramp generator or a controlled pulse length monostable multivibrator is used. The use of either approach with a repetition frequency of 2.5 Kc requires time constants of RC networks on the order of 400 μsec. Capacitors of 500 pf and 800KΩ resistors (400 μsec time constant) are both bordering on the state-of-the-art in microcircuitry. Therefore, another approach must be used.

The design selected is quite novel and could not be used for a circuit with discrete components because of the high parts count and circuit complexity. Fig. 4 is a preliminary functional diagram. Using a 640 Kc oscillator and digital logic, a 128-step digital ramp is generated. This ramp approaches linearity much closer than most RC-type ramp generators. The ramp is compared with the amplified input error signal. When amplified and limited, the difference between the two signals becomes a constant-frequency, variable pulse width, digital signal. This signal is applied to the two "and" gates. Each "and" gate has as a second signal an output of flip-flop 8. The outputs of the "and" gates are those shown in Fig. 3.

4. Summary

Microcircuits as complex as this analog voltage to duty cycle generator have not been made to date; such circuits will require several multilayer interconnections on the silicon chip. Under another JPL contract, Westinghouse is building a microcircuit power synchronizer with multilayer interconnections. As it is almost as complex as the duty cycle generator, the new device will require only an extension of available techniques. The development and fabrication of the device is estimated to require 1 yr.

This device will facilitate all areas of switched regulator use. The designer can use a block diagram approach, without tedious circuit designing and breadboard testing. Fabrication of the regulator will also be simplified when compared to discrete component methods.

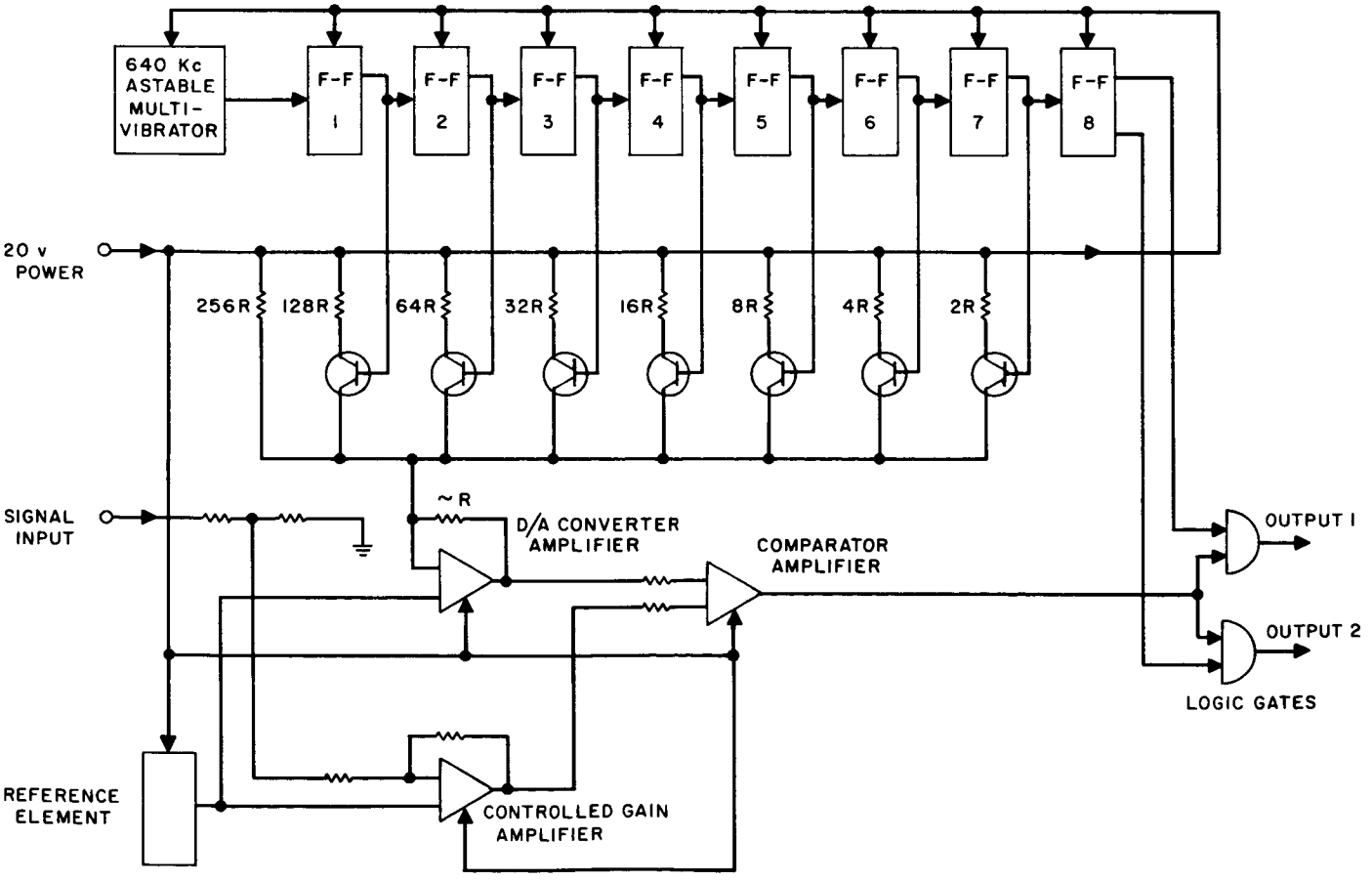


Fig. 4. Circuit block diagram of duty cycle generator

Table 1. Switched regulator specifications

Repetition rate for either output	2.5 Kc ±10%
Output level	the "one" level is 5 ±0.25 v
Output load	zero to 50 ma
Output voltage regulation	The ±0.25v tolerance on output level shall be maintained through any and all of the following: input voltage variation, temperature range, and full load variation.
Input power	20 v DC ±1 v; this implies internal regulation.
Input signal	20 v DC ±0.1 v at a nominal 10K source impedance.
Output duty cycle swing	From 97% or more, on part of duty cycle for lowest input signal (− 0.1 v) to 3% or less, on part of duty cycle for maximum input signal (+ 0.1 v).
Output duty cycle linearity	If the transfer function of output duty cycle (ordinate) versus input signal (abscissa) is drawn, ideally it should be a straight line; we shall define a linearity tolerance to be a vertical displacement of the true value from a straight line drawn through the 97% and 3% true values; the maximum vertical departure at any input signal from the straight line shall be + 2.5 duty cycle percentage points. (Use the ordinate scale to determine the departure.)
Output time symmetry	During steady-state conditions (any constant value input signal), the "on" time of the signal from one output line shall be within 0.5% of the "on" time of the other output line; in addition, the half-period time (Fig. 3) which determines the start of output number 2 shall be within ±0.25% of the absolute value of half-period time.
Temperature	− 10 to 75°C
Output voltage rise and fall time	0.5 μsec maximum from zero to specified output, and vice versa.

B. Solar Tracking System

S. S. Luebbers

1. Introduction

In order to evaluate solar heated thermionic power sources, an accurate solar tracking system is required. The tracking system must provide a Sun-oriented structure upon which solar thermionic experiments may be conducted. In the following paragraphs, the basic components and structures of a solar tracking system, currently being fabricated, are discussed and compared to a system presently in use at the JPL Table Mountain test facility.

2. Design Goals

Design and fabrication of a new tracking system was undertaken to alleviate several problems encountered in the operation of the existing tracking system. The problems encountered thus far have been (1) lack of structural rigidity, (2) torque limitation with wind velocities in excess of 5 mph, and (3) drift of the tracker from normal solar orientation. Since these problems have significantly reduced the effectiveness of sunlight testing of solar thermionic systems, it was decided that an improved tracking system should be developed to overcome these limitations. The design specifications for this improved tracker include: minimum drift, sufficient torque for operation in a constant wind velocity of 15 mph and in gust velocities up to 25 mph, maximum deflection of the experimental platform of 0.040 in., and provisions for mounting a 12-ft-D solar concentrator.

3. Tracking System Components

The basic solar tracking system is composed of the following components: (1) Mobile base structure. (2) Azimuth drive actuator. (3) Outer yoke structure. (4) Inner yoke structure. (5) Elevation drive actuator. (6) Experiment platform. (7) Associated electronics and instrumentation. Each of the above components is discussed in the following summary.

a. Mobile base structure. As may be seen in Fig. 5, the existing solar tracker utilizes ordinary searchlight components to provide the mobile base. Not shown are the leveling screws which provide solid ground support. One major structural rigidity problem is caused by supporting the outer yoke on the small-diameter drum. In order to alleviate this problem, a 5-in. caliber-38 gun mount was selected as the base structure for the improved system. In Fig. 6, the improved base structure is shown in an inverted position. Also shown are the cross members to which steel casters are attached.

b. Azimuth drive actuator. The prototype tracking system utilizes a Minneapolis–Honeywell DC motor as a prime mover for the azimuth drive train. Even though a large gear reduction is provided, the tracking system becomes torque-limited under relatively small wind loadings. To circumvent this problem, a motor with higher torque is provided on the improved system. The higher



Fig. 5. Prototype solar tracker

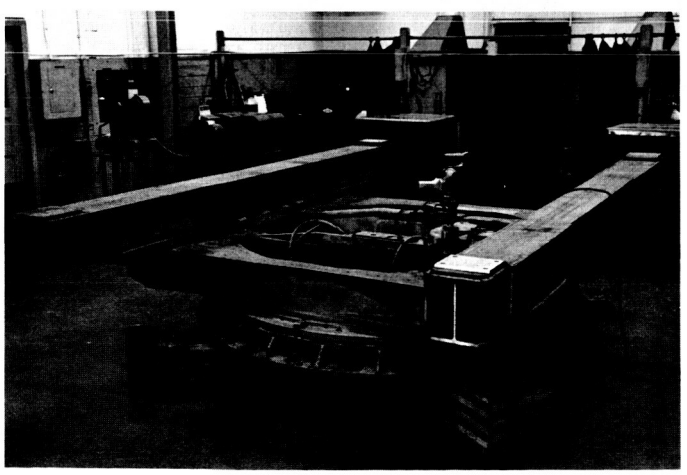


Fig. 6. Improved base structure

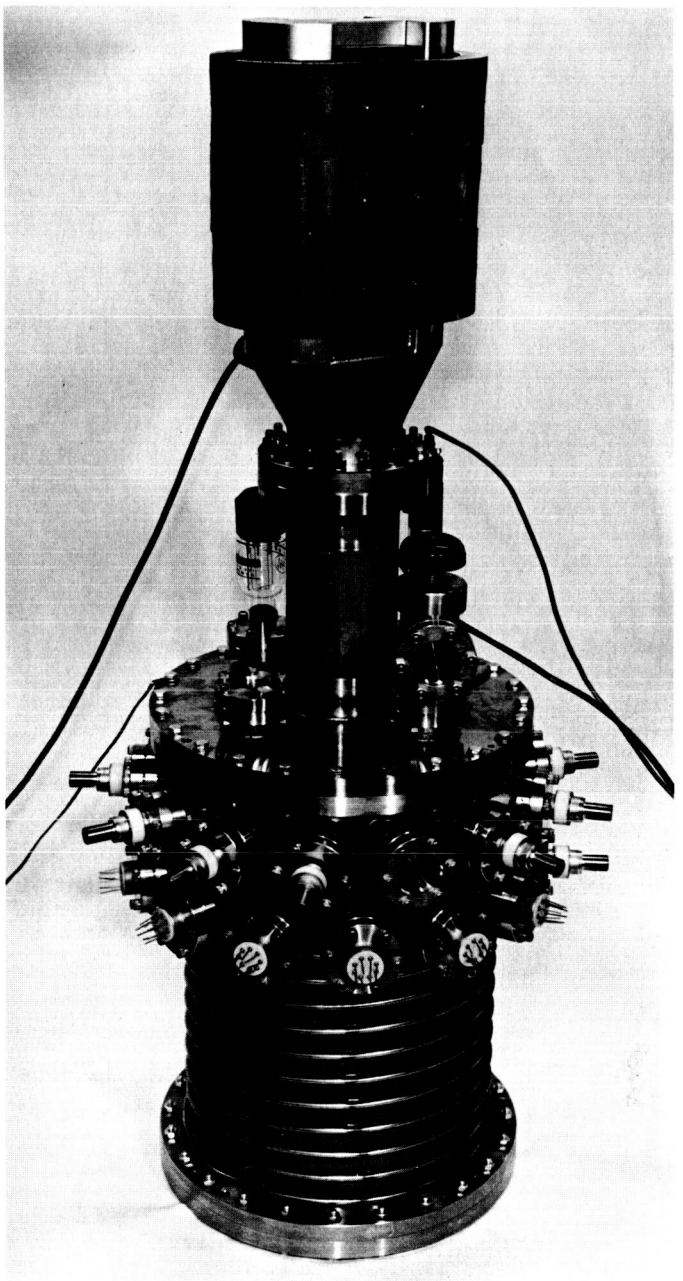


Fig. 7. Vacuum chamber

torque motor is a double-differential synchronous motor which provides full torque from zero rpm to slewing speeds.

c. Yoke structures. The searchlight outer-yoke structure was modified by cutting the small-diameter searchlight yoke near the base. The stubs were then boxed to provide support for an 10-ft-D inner yoke or ring. The solar concentrator is attached to the inner yoke (see Fig. 5). The inner yoke also supports the experiment platform.

Thermionic generator testing must be done in a vacuum environment. This requires the attachment of a vacuum chamber, shown in Fig. 7, to the experiment platform. Concentrated solar energy is admitted to the chamber by a pyrex window attached to the front of the chamber. The total chamber weight is approximately 500 lb.

Excessive platform deflection has necessitated the design of a heavier inner-yoke structure. The newly developed structure is shown in Fig. 8. As may be seen in the

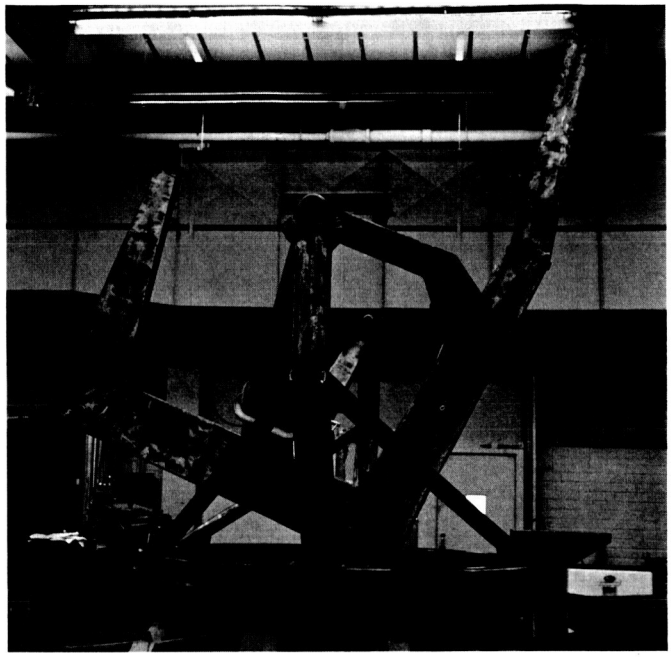


Fig. 8. Improved yoke structure

figure, the ring concept has been abandoned and a simple support is provided for the 12-ft-D solar concentrators.

d. Elevation drive actuator. On the prototype system, the elevation is provided by a bull gear attached to one of the inner-yoke pivot points. The drive motor and gear train combination is similar to the azimuth drive; however, electromagnetic braking is provided.

In order to allow greater elevation torque, the improved system utilizes a cable drive. A cable trough (not shown) is welded to the inner yoke, and a drive cable is stretched around an actuated drum.

e. Experiment platform. On both the prototype and improved tracking systems, x-y-z positioning of the experiment platform is provided. This movement is required to locate the concentrated solar energy at the desired point.

f. Electronics and instrumentation. The prototype tracking system utilizes a radiation tracking transducer (RTT) as a source of error signal. The error signal is chopped, amplified, and used to provide the control voltage for the Minneapolis–Honeywell motor. The RTT has a serious problem of drift. The Sun is imaged directly on a semiconductor which inherently changes characteristics with heating. To avoid this, a shadow bar transducer is employed on the improved system.

The improved control loop consists of the shadow bar transducer, amplifiers, choppers, logic circuitry, and a variable control oscillator (VCO). The VCO provides the necessary control signal to the differential motor.

The control console for the prototype tracking system is shown in the foreground of Fig. 9. The remaining six cabinets provide the necessary instrumentation for testing thermionic generators.

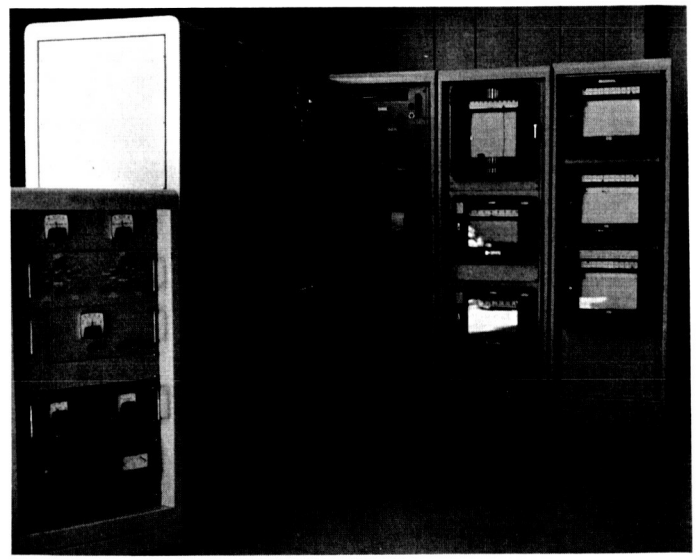


Fig. 9. Control console and experiment instrumentation

4. Improved Tracker Status

The improved tracker is in the final fabrication and assembly stage and will be ready for checkout during the month of October. It is believed that the problems encountered with the prototype system have been rectified by this improved tracker, and a functional system should result from the current design.

IV. Spacecraft Control

A. Horizon Scan Platform System

K. Kerner and S. Szirmay

The goal of the scan platform system development effort is to provide a mounting base for scientific instruments which will be capable of an orientation independent of that of the spacecraft. A light-weight low-power and high-reliability scan system is sought that can be programmed or commanded to search, acquire, and track or scan a target planet. As shown in Fig. 1, the platform subsystems should be comprised of a planet sensor and/or a programmer, a controller, a motor, and an actuator.

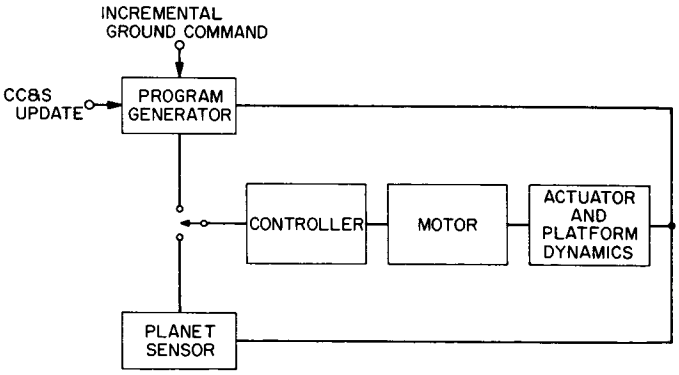


Fig. 1. Horizon scan platform position control system

The sensor used to control the platform position is an infrared horizon scanner now under development by Barnes Engineering under contract from JPL. The salient features of the scanner follow.

The Barnes scanner consists of four optical scanning heads (a pair for each axis) and the signal-processing electronics. Each head contains a 10×81 deg array of detectors. The instantaneous field-of-view of each detector, neglecting image aberrations, is 0.25×10 deg. The arrays consist of 90 detectors spaced 0.9 deg apart on center, giving the total array size of 10×81 deg. A scan in one axis consists of simultaneously sampling the detectors in two opposing heads, starting with the detectors farthest apart and proceeding towards the scanning nadir. If the target planet should happen to be positioned on the optical axis of the scanner, both heads will give a concurrent signal indicating planet presence as the horizon is crossed, and no error signal will be generated. If a

pointing error exists, one head will provide a presence signal first; the number of detectors that must be sampled by the opposing head until it also indicates planet presence will be proportional to the pointing error. The scan period is 100 msec, defining the frequency with which the error signal is updated. Fig. 2 shows the scan geometry of the horizon scanner.

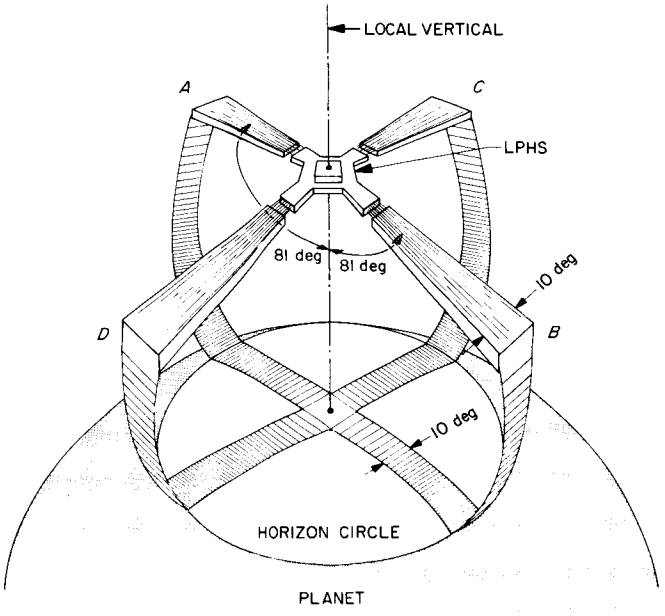
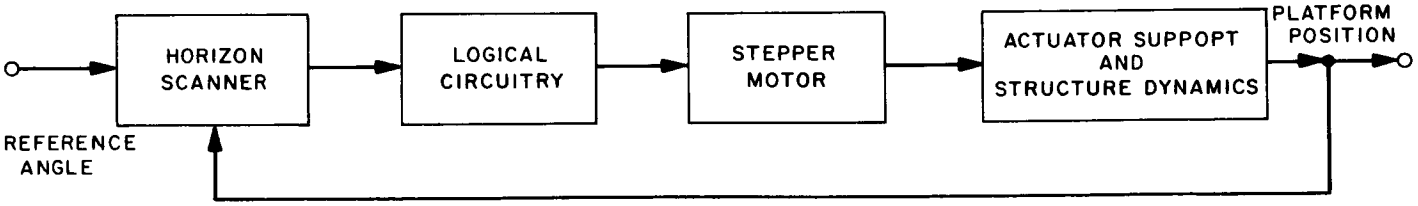


Fig. 2. Scan geometry of the horizon scanner

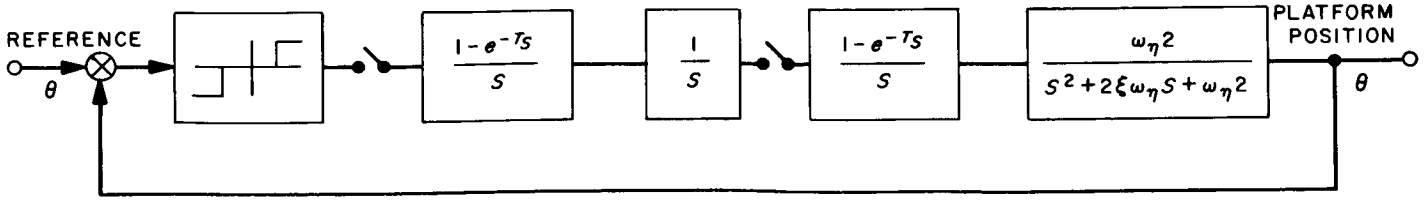
Two alternate approaches in mechanizing a platform that shall satisfy the science instrument requirements are being explored. One approach, and from the standpoint of mechanization the simpler one to implement, is shown in Fig. 3(a). The horizon scanner output is processed by logical circuitry in a manner where any error, regardless of its magnitude, drives the motor one step in a direction dictated by the polarity of the error. The transfer functions of the several blocks are shown in Fig. 3(b). The horizon scanner is represented by a nonlinear amplifier fed into a sampling switch and followed by a zero-order hold circuit. The stepper motor is represented by an integrator, sampler and zero-order hold circuit. The boom-mounted platform is characterized by a second-order system transfer function.

This type of system defines the maximum stepping rate of the motor as 10 steps/sec. From the orbital consideration, as presently formulated, the maximum tracking rates required from the scan platform are in the order of 0.2 deg/sec. Allowing for a 100% margin to accommodate the system dynamics, an incremental rate of 0.04 deg/step can be realized. The system is being simulated on the analog computer to investigate the stability characteristics.

The proposed mechanization of the system is shown in Fig. 4. Heads "A" and "B" of Fig. 4 refer to the two complementary optical scanning heads that comprise a one-



(a) FUNCTIONAL DIAGRAM



(b) ANALYTICAL DIAGRAM

Fig. 3. Stepper motor driven platform

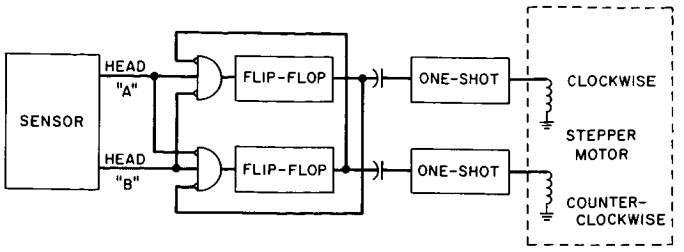


Fig. 4. Stepper motor drive system mechanization

axis sensor device. A signal output from Head "A" followed by a delayed output from Head "B" corresponds to some error in the platform orientation from the local vertical. From Fig. 4, it is seen that an output from Head "A" actuates its associated flip-flop which in turn inhibits the other flip-flop from being energized for the remainder of the scan and also actuates a "one shot" whose "on" time is of a duration enabling the stepper motor to be driven one step. A simultaneous output from both heads, corresponding to no error, is inhibited. A new scan resets both flip-flops.

An alternate approach being studied is the use of a standard two-phase servomotor as the actuator. The correspondent block diagram is shown in Fig. 5(a), while Fig. 5(b) shows the correspondent transfer function representation. An analog computer simulation of the above is in process.

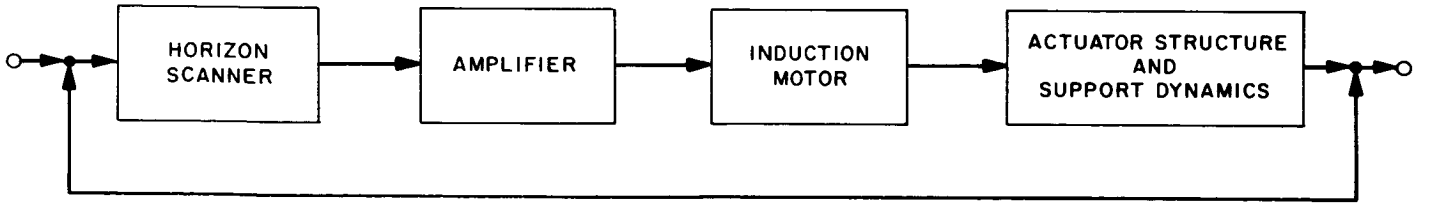
A sample and zero-order hold circuit has been breadboarded and tested. The remaining circuitry necessary to complete the loop is being designed.

B. Spinning Spacecraft Attitude Control System

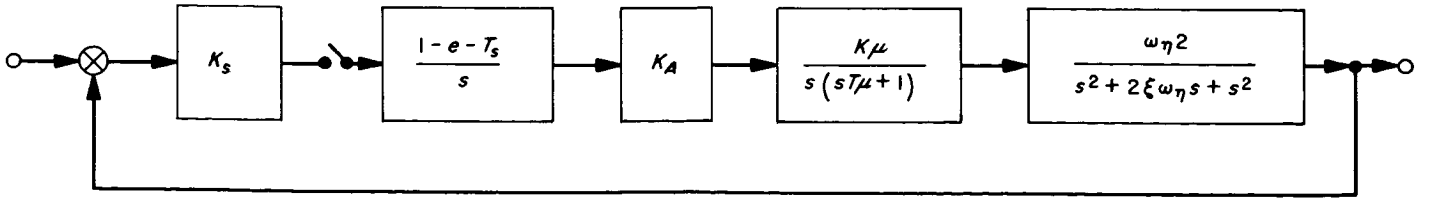
B. M. Dobrotin and J. C. Nicklas

1. Introduction

This report presents a method of analysis for a spinning spacecraft attitude control system. This analysis is a continuation of the study reported in *SPS 37-34*, Vol. IV, pp. 66–76. The study showed that spin stabilization provides a simple method for fixing the attitude of the spacecraft in inertial space, and it also described some of the constraints imposed upon the spacecraft by spin stabilization. It was pointed out that, while spin stabilization offers the advantage of simplicity, it does fix the attitude with respect to inertial space. In order to change the pointing of the spacecraft, an active control system must be used. A simple Sun-pointing attitude control was described in *SPS 37-34*, and some of the general parameters of such a system were obtained. However, the specific



(a) FUNCTIONAL DIAGRAM



(b) ANALYTICAL DIAGRAM

Fig. 5. Induction motor driven platform

operation of the control system as it approached the Sun orientation was not defined. The analysis presented below outlines a simple, easily mechanized method for determining the performance of a single-pulse jet that is to be controlled by a Sun sensor. In addition, the results of several computer runs are presented, and variations of the basic control system are discussed.

2. Analysis

The basic system to be analyzed is shown in Fig. 6(a). The spacecraft is represented by an oblate ellipsoid of revolution with moments of inertia of I_R about the roll axis and I about the pitch and yaw axes. The unit vectors $\hat{i}$, $\hat{j}$ and $\hat{k}$ are attached to the spacecraft pitch, yaw and spin axis, respectively. The Sun sensor is aligned with the $\hat{j}$ unit vector and is a slit Sun sensor with a field of view extending from the $-\hat{k}$ axis to the $+\hat{k}$ axis, less a deadzone, γ. The control jet is located on the $\hat{i}$ axis and is triggered whenever the Sun is in the field of view of the Sun sensor. Ideally, this control system would function so that the spin axis would be pointed at the Sun within a cone of half-angle γ and would remain within this cone with no further operation of the control system.

The attitude of the spacecraft may be referred to the Sun as shown in Fig. 6(b). Assuming that the spacecraft is in the between-pulse time interval, it will be a torque-free body and will precess with a fixed cone angle θ about the angular momentum vector **H** which is fixed in inertial space. Additionally the Sun position may be represented by a unit vector $\hat{S}$ which is also fixed in inertial space. Thus the $\hat{X}$, $\hat{Y}$, $\hat{Z}$ coordinate system is defined as an inertially fixed system with the $\hat{Z}$ axis along the angular momentum vector and the Sun vector $\hat{S}$ contained in the $\hat{Y}$-$\hat{Z}$ plane. The body axes are related to the inertial axes through the roll, pitch, roll Euler angle sequence represented by the angles ψ, θ, ϕ. The sunline is related to the inertial axes through the angle α. For the purposes of this analysis, the Sun sensor is represented by a plane contained in the $\pm\hat{k}$, $+\hat{j}$ plane. Every time the Sun sensor field of view contains the Sun, the control jet is triggered and delivers a torque impulse of magnitude a. This is expressed for purposes of analysis as "the sunline must lie in the Sun-sensor plane."

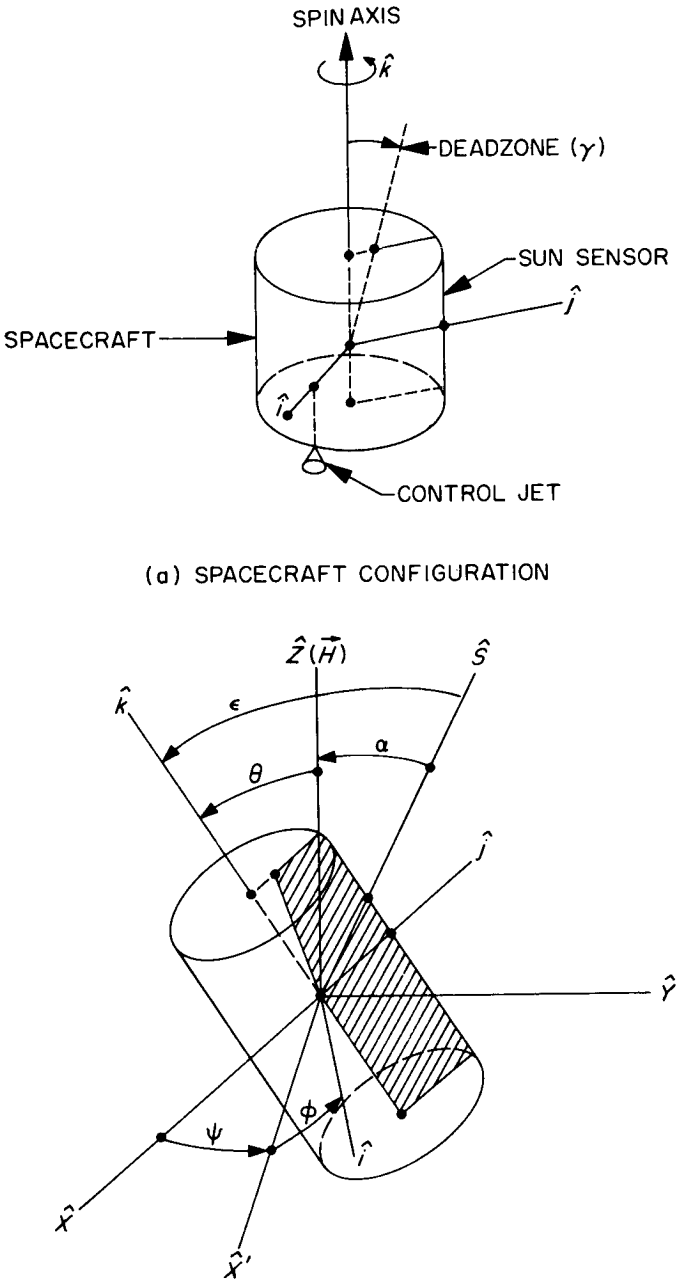


Fig. 6. Configuration and coordinate systems of spinning spacecraft

The condition that the sunline will lie in the Sun-sensor plane will occur when

$$\hat{S} \times \hat{k} = \sin\epsilon\,\hat{i} \geq \sin\gamma\,\hat{i} \qquad (1)$$

This is true since the cross product of any two vectors must be perpendicular to the plane formed by the two vectors. The angle ϵ is defined as the error angle between $\hat{S}$ and $\hat{k}$. For the general condition,

$$\hat{S} = [A]\,(\hat{X}\,\hat{Y}\,\hat{Z})^T$$

and

$$(\hat{X}\,\hat{Y}\,\hat{Z})^T = [B]\,(\hat{i}\,\hat{j}\,\hat{k})^T$$

where A and B are transformation matrices. Then

$$\begin{aligned}\hat{S} &= [A][B]\,(\hat{i}\hat{j}\hat{k})^T \\ &= a\hat{i} + b\hat{j} + c\hat{k} \\ &= [(\text{Cos}\,\phi\,\text{Sin}\,\psi + \text{Sin}\,\phi\,\text{Cos}\,\theta\,\text{Cos}\,\psi)\,\text{Sin}\,\alpha \\ &\quad + \text{Sin}\,\theta\,\text{Sin}\,\phi\,\text{Cos}\,\alpha]\,\hat{i} \\ &\quad + [\text{Sin}\,\alpha\,(\text{Cos}\,\phi\,\text{Cos}\,\psi\,\text{Cos}\,\theta - \text{Sin}\,\phi\,\text{Sin}\,\psi) \\ &\quad + \text{Sin}\,\theta\,\text{Cos}\,\phi\,\text{Cos}\,\alpha]\,\hat{j} \\ &\quad - [\text{Sin}\,\theta\,\text{Cos}\,\psi\,\text{Sin}\,\alpha - \text{Cos}\,\theta\,\text{Cos}\,\alpha]\,\hat{k} \qquad (2)\end{aligned}$$

and

$$\hat{S}\times\hat{k} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ a & b & c \\ 0 & 0 & 1 \end{vmatrix} = b\hat{i} - a\hat{j}$$

$$\begin{aligned}&= [\text{Sin}\,\alpha\,(\text{Cos}\,\phi\,\text{Cos}\,\psi\,\text{Cos}\,\theta - \text{Sin}\,\phi\,\text{Sin}\,\psi) \\ &\quad + \text{Sin}\,\theta\,\text{Cos}\,\phi\,\text{Cos}\,\alpha]\,\hat{i} \\ &\quad - [(\text{Cos}\,\phi\,\text{Sin}\,\psi + \text{Sin}\,\phi\,\text{Cos}\,\theta\,\text{Cos}\,\psi)\,\text{Sin}\,\alpha \\ &\quad + \text{Sin}\,\theta\,\text{Sin}\,\phi\,\text{Cos}\,\alpha]\,\hat{j} \qquad (3)\end{aligned}$$

Eqs. (1) and (3) may be linearized with respect to θ and α since satisfactory operation of the control system requires that θ remain small and $\alpha\to 0$. Thus,

$$\begin{aligned}\hat{S}\times\hat{k} &= [\alpha\,\text{Sin}\,(\phi+\psi) + \theta\,\text{Cos}\,\phi]\,\hat{i} \\ &\quad - [\alpha\,\text{Cos}\,(\phi+\psi) + \phi\,\text{Sin}\,\phi]\,\hat{j} \qquad (4)\end{aligned}$$

For rigid-body torque-free motion it may be shown that (for small θ):

$$\dot{\phi} = -\lambda\omega_R \qquad (5a)$$

$$\dot{\psi} = (1+\lambda)\omega_R \qquad (5b)$$

where

$$\lambda = \left(\frac{I_R}{I} - 1\right)$$

$$\omega_R = \text{initial spin rate}$$

Combining Eqs. (5a) and (5b), integrating and substituting into Eq. (4):

$$\begin{aligned}\hat{S}\times\hat{k} &= [\alpha\,\text{Cos}\,(\omega_R t) + \theta\,\text{Cos}\,\lambda\omega_R t]\,\hat{i} \\ &\quad - [\alpha\,\text{Sin}\,(\omega_R t) - \theta\,\text{Sin}\,\lambda\omega_R t]\,\hat{j} \qquad (6)\end{aligned}$$

Now changing to complex notation, with $\hat{j}$ as the imaginary axis:

$$\begin{aligned}\hat{S}\times\hat{k} &= \alpha[\text{Cos}\,(\omega_R t) - j\,\text{Sin}\,(\omega_R t)] \\ &\quad + \theta[\text{Cos}\,(\lambda_R t) + j\,\text{Sin}\,(\lambda\omega_R t)] \\ &= \alpha e^{j\omega_R t} + \theta e^{j\lambda\omega_R t} \qquad (7)\end{aligned}$$

where $j \equiv \sqrt{-1}$. Thus, $\hat{S}\times\hat{k}$ reduces to the sum of two vectors rotating in opposite directions. Additionally, the requirements for triggering the Sun sensor, Eq. (1), are met when the vector sum is positive real and greater than γ.

Since Eq. (7) was derived under the assumption that the spacecraft was a torque-free rigid body, the effect of the torque impulse on the body must be determined so that the initial conditions for Eq. (7) may be met. First, some physical significance must be attached to Eq. (7). Forming the cross product of any two unit vectors:

$$\hat{\xi}\times\hat{\eta} = \sin\beta\,\hat{\zeta} \cong \beta\,\hat{\zeta} \qquad (8)$$

for small β; β is the included angle between $\hat{\xi}$ and $\hat{\eta}$ (measured positive from $\hat{\xi}$ to $\hat{\eta}$) and $\hat{\zeta}$ is a unit vector orthogonal to the plane of $\hat{\xi}$ and $\hat{\eta}$. Thus $\beta\hat{\zeta}$ is a rotation vector which will produce a rotation from $\hat{\xi}$ to $\hat{\eta}$ and may be defined as the moment of the angle β. Examining Fig. 6(a), it may be seen that the product $\hat{S}\times\hat{Z}$ produces the moment of α, $\hat{Z}\times\hat{k}$ the moment of θ, and $\hat{S}\times\hat{k}$ the moment of ϵ. These may now be treated as vectors: $\alpha\hat{x}$, $\theta\hat{x}'$ and $\epsilon\hat{z}$; $\hat{z}$ is defined as the unit vector orthogonal to $\hat{S}$ and $\hat{k}$, and $\hat{z} = \hat{i}$ for Sun sensor triggering. Again referring to Fig. 6(b), these vectors may be added:

$$\epsilon\hat{z} = \alpha\hat{x} + \theta\hat{x}' \qquad (9)$$

The vectors on the right-hand side of Eq. (9) may be transformed into the body-fixed coordinate system. When α and θ are referenced to the $\hat{i}$ axis using complex notation (i.e., $e^{j\rho}$ where ρ is the rotation angle), Eq. (9) becomes Eq. (7). Eq. (7) may then be plotted in the complex plane as shown in Fig. 7. Fig. 7 is then the analogy of Fig. 6(b) with α and θ linearized. In the complex plane, the spin axis $\hat{k}$ may then be considered to be located at the origin, **H** located at the end of the $\boldsymbol{\theta}$ vector, and $\hat{S}$ at the end of $\boldsymbol{\alpha} + \boldsymbol{\theta}$. The rotation of $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$ represent the apparent motion of the Sun with respect to the spacecraft.

When the conditions of Sun sensing are met, Eq. (1), an impulsive torque is applied which produces a component $\Delta\mathbf{H}$ of angular momentum perpendicular to the total angular momentum vector $\mathbf{H}_1$, as shown in Fig. 8(a):

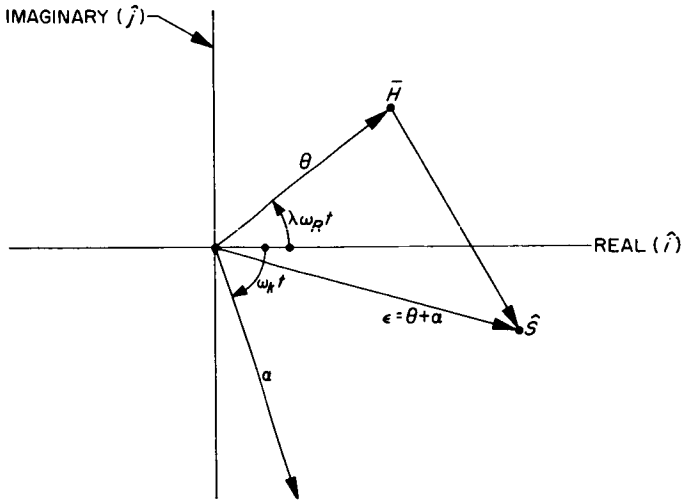


Fig. 7. Complex plane representation

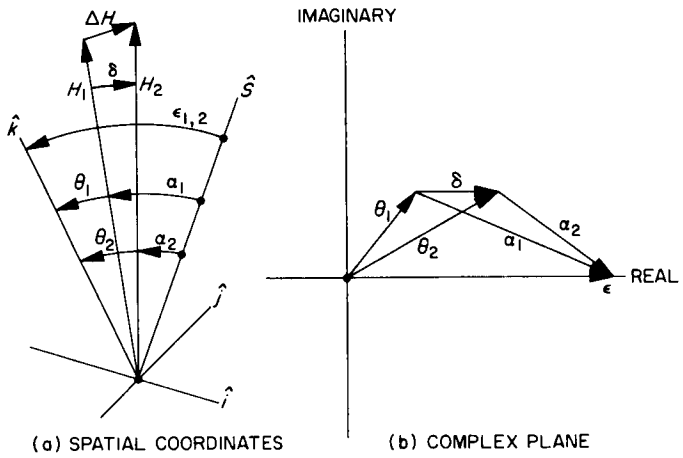


Fig. 8. Torque application to spinning spacecraft

$$\mathbf{H}_2 = \Delta\mathbf{H} + \mathbf{H}_1 \tag{10}$$

where

$$|\Delta\mathbf{H}| = a \tag{11}$$

Further, since the control jet is fixed within the body, the torque vector will be along the $\hat{j}$ axis and $\Delta\mathbf{H}$ will also lie along the $\hat{j}$ axis:

$$\mathbf{H}_2 \times \mathbf{H}_1 = \left(\tan^{-1}\frac{|\Delta\mathbf{H}|}{|\mathbf{H}_1|}\right)\hat{i} \equiv \delta\,\hat{i} \tag{12}$$

Thus, each time the control jet is triggered, **H** is rotated by an increment δ. Since the position of the body with respect to the Sun cannot change instantaneously, the position of $\hat{S}$ with respect to $\hat{k}$ must remain constant. From this it is seen that $\hat{S} \times \hat{k}$ remains constant but $\hat{Z} \times \hat{k}$ is increased by the amount δ. On the complex plane representation in Fig. 8(b), this is accomplished by incrementing θ by $+\delta$ and incrementing α by $-\delta$ along the real axis. This will generate the initial conditions for the next torque-free rotation of the spacecraft Assuming that initially **H** and $\hat{k}$ are aligned, the first pulse will generate θ such that θ and α are directly additive. During the next rotation period of the spacecraft, θ and α will have rotated so that θ does not directly add to α and the total error ϵ is reduced. For the case of the sphere ($I_R = I$, $\lambda = 0$), θ and α rotate at the same rate. As a result, the θ generated by each subsequent pulse will be directly attitive to α each time the Sun sensor is triggered, and θ will increase until the vehicle tumbles ($\theta \rightarrow 90°$). For the case of $I_R = 1.5\,I$, $\lambda = 0.5$, θ rotates at one-half the rate of α, and every other pulse cancels out the θ generated by the preceding pulse.

3. Simulation

Eq. (7) and the constraints imposed by Eqs. (1) and (12) were mechanized for the analog computer. The block diagram is given in Fig. 9 for reference. Initial runs with the system described above indicate satisfactory operation until the deadzone is reached. At this time, depending on the initial conditions, the cone angle θ diverges, and the total pointing error increases. This is illustrated in Fig. 10, which shows the trace of ϵ, the total error, in the complex plane for a spacecraft with $\lambda = 0.5$.

Since Fig. 10 only shows the total error ϵ, Fig. 11 is provided to give a better idea of the variation of θ, α and ϵ by displaying the real and imaginary components of the various quantities. As shown when Im $\epsilon = 0$ and Re $\epsilon > 0$, the Re θ is increased by δ and Re α is decreased by δ. This continues with the magnitude of θ increasing and decreasing as the phase angle between α and θ changes. Fig. 10 shows that after ϵ enters the deadzone, it will start increasing and continues to increase until the vehicle tumbles. Fig. 11 indicates that this is due to the increase in θ.

The explanation is as follows: Though the illustration is given for $\lambda = 0.5$, it remains the same for any $\lambda > 0$.

After the $(n+1)$th revolution, the spacecraft will be in the following condition, as shown in Fig. 12(a):

$$\epsilon > \gamma$$

$$\text{Re}\,\alpha_{n+1} < \gamma$$

$$\text{Re}\,\theta_{n+1} = \epsilon_{n+1} - \text{Re}\,\alpha_{n+1} > 0$$

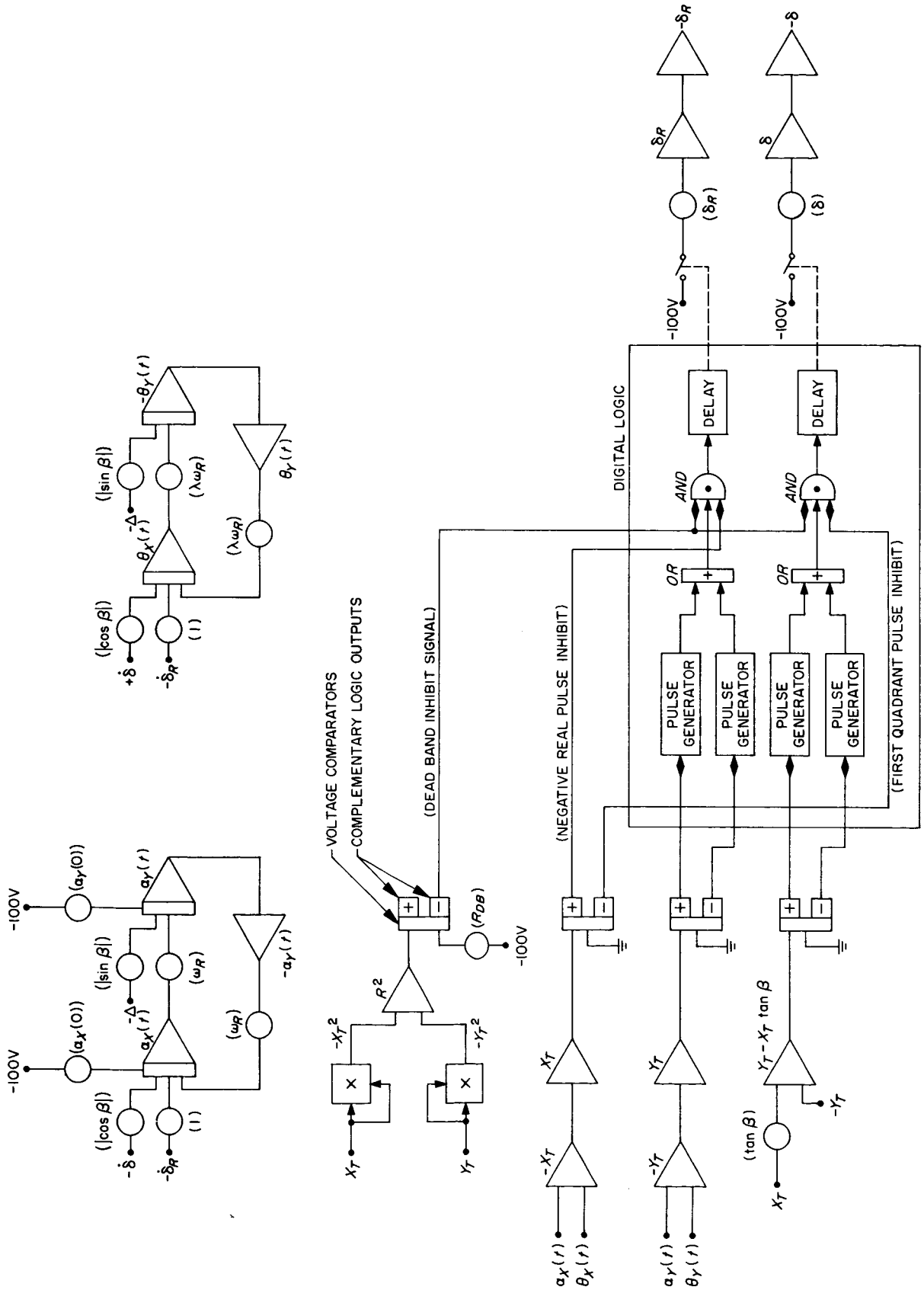


Fig. 9. Two-pulse spin-axis stabilization

At the $(n+2)$th revolution, Fig. 12(b),

$$\text{Re}\,\theta_{n+2} < 0$$
$$\epsilon_{n+2} < \gamma$$

and the Sun sensor is not triggered. After the $(n+3)$th revolution, the spacecraft is in the same position it was at the $(n+1)$th revolution, and the Sun sensor is triggered, adding the increment δ to θ. This will continue, with the Sun sensor triggering on every other revolution until the vehicle tumbles. At some point, Fig. 12(d), Re α will be negative at the time of Sun sensing. This indicates that the sunline is now located between the momentum vector and the spin axis and now **H** will be rotated away from the Sun. This illustrates the fact that the control system rotates **H** away from the spin axis, not toward the Sun, due to the body-fixed torque.

Even though the above example specifically describes a spacecraft which has a λ of 0.5, the analogy is applicable to any $\lambda > 0$. Thus, θ will always increase in magnitude

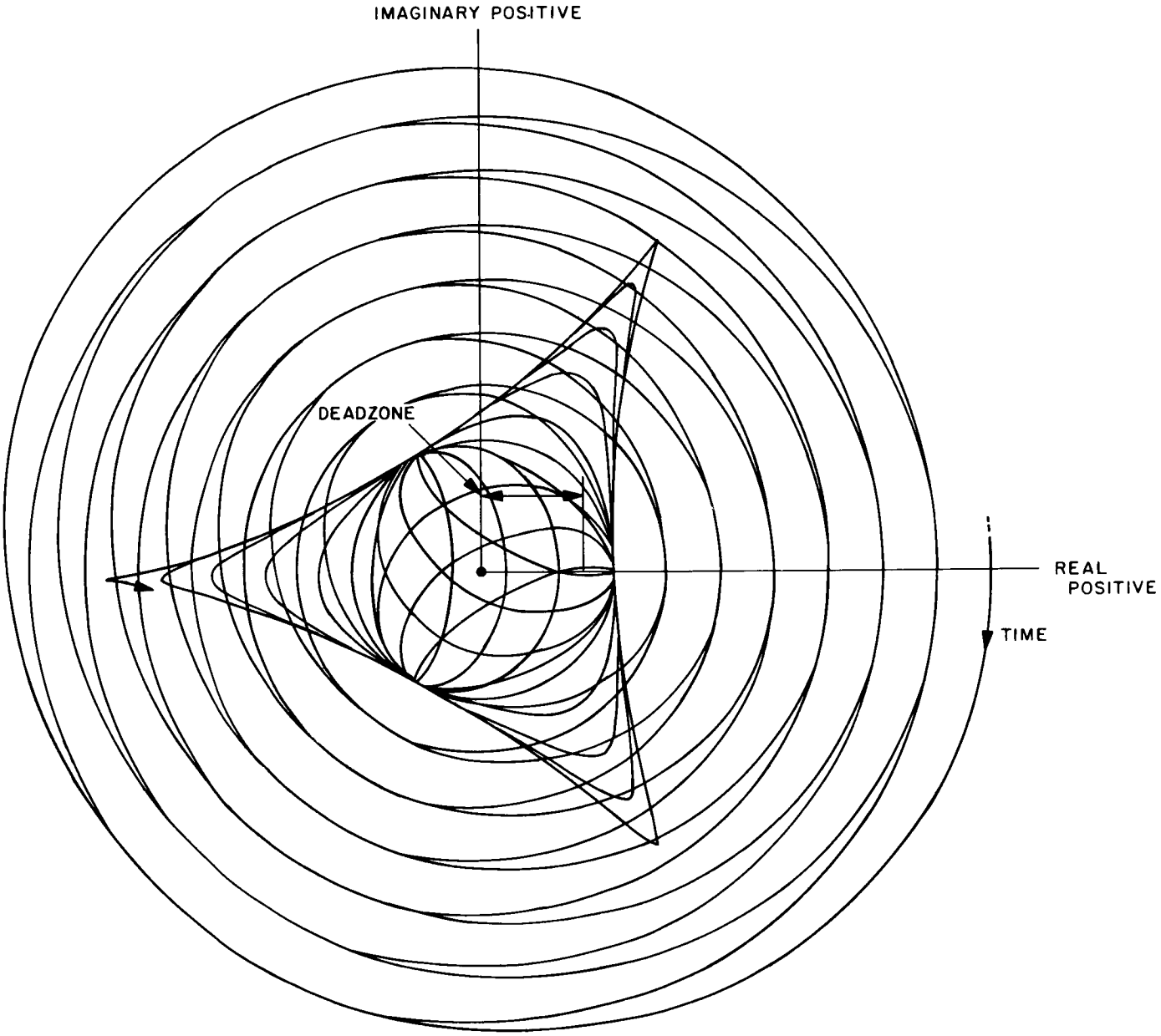


Fig. 10. Error angle versus time for $\lambda = 0.5$

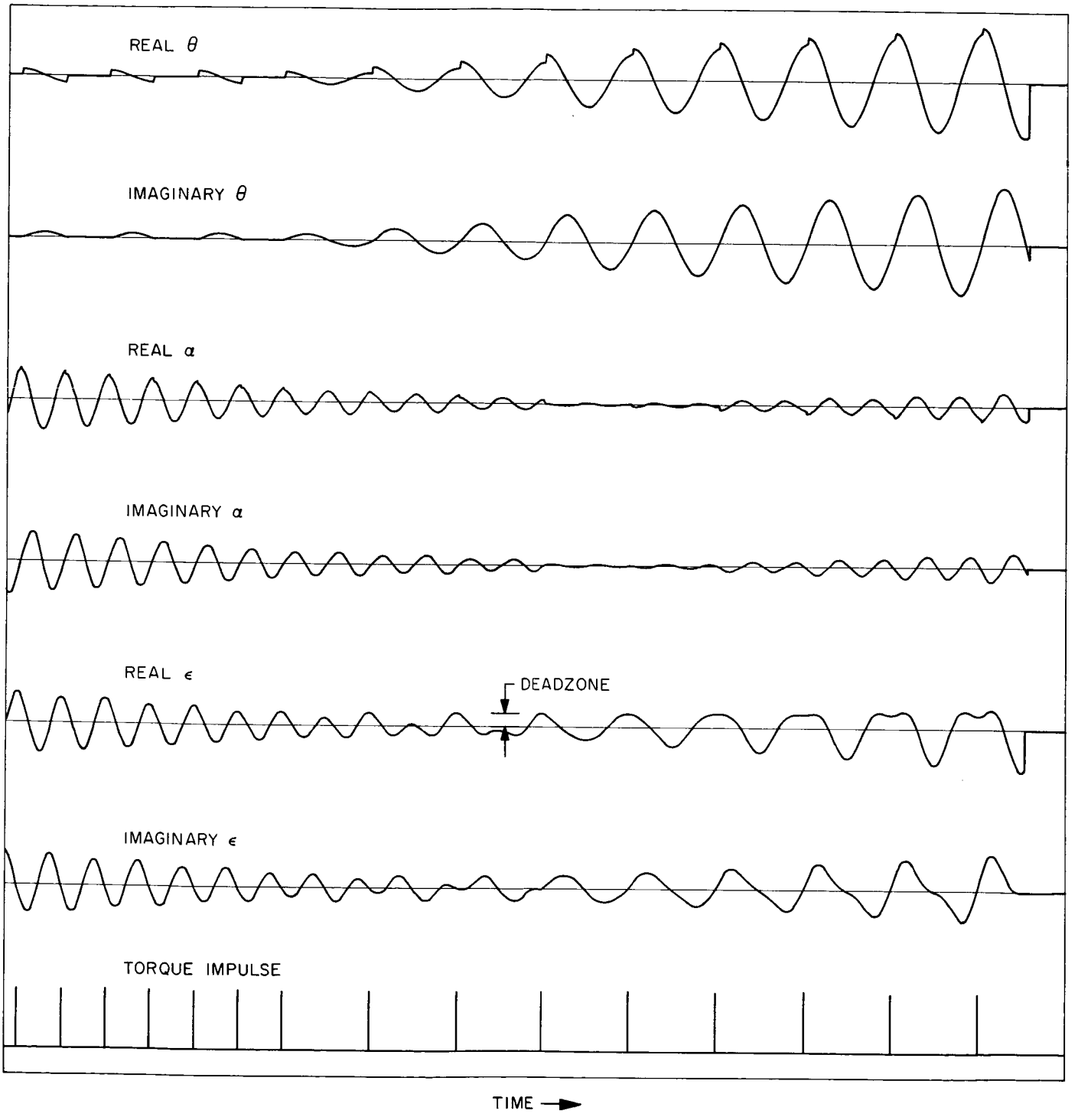


Fig. 11. Error components versus time

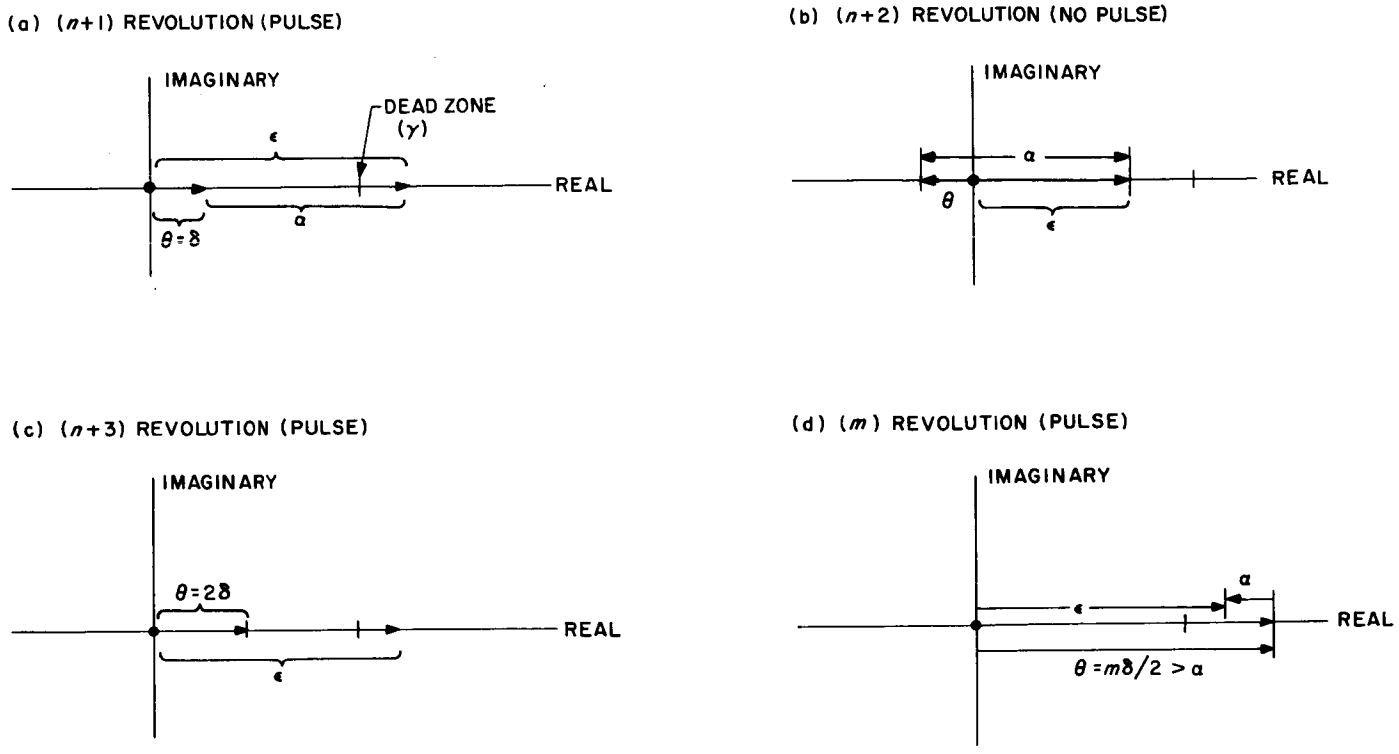


Fig. 12. Operation of control system in deadzone, $\lambda = 0.5$

after the deadzone is reached. A given set of conditions may be postulated wherein the control system produces stable deadzone operation, but these require rigid constraints on the body configuration and initial conditions and cannot be taken as a practical situation.

4. Summary

The simple attitude control system postulated thusfar must be modified to prevent the vehicle from tumbling. Several different modifications may be made, with both advantages and disadvantages. Reference 2 uses a rate gyro to sense the cross axis rate, and thus measure θ independently of the total error. Then the control jet may trigger on a linear combination of α and θ. Reference 3 utilizes a time delay and offsets the torque axis to insure that δ will decrease θ as the deadzone is entered, instead of increasing θ. A simpler method is to use an auxiliary Sun sensor to determine when the sunline has first entered the deadzone. At this time the control system is disabled and a passive damper eliminates the cone angle.

Several assumptions have been made, and the usefulness of the analysis may be extended by generalizing these assumptions. First, the torque generator was assumed to provide a perfect impulse. For a real system this is not true, as the torque is applied over a finite time. This may be accomplished in the simulation by changing the digital comparator circuit to allow a trigger band equal to the Sun sensor look angle. The δ input to the θ and α linear oscillators is then rescaled, and the simulation proceeds as before. The second assumption is that the spacecraft is symmetric. As shown in *SPS 37-34*, Vol. IV, pp. 66–76, this assumption is justified for asymmetry up to $\sim 3\%$. Over this value, the single θ vector may be replaced by two vectors, as described in *SPS 37-34*.

C. Qualitative Vacuum Testing of Attitude Control Thrust Nozzles Used in a Pulse Mode

T. J. Donlin

Flight telemetry data returned on spacecraft attitude control has clearly shown that an accurate measurement of the dynamic thrust of the attitude control thrust nozzles in a vacuum is required. The impulses imparted to the pitch, roll, and yaw axes of the *Mariner* spacecraft by the attitude control gas jets were higher than anticipated.

Subsequent analysis isolated the conditions leading to this effect to the thruster valve and nozzle. An accurate direct thrust and impulse measurement would have exposed the condition before flight. The indirect mechanical and pressure measurements coupled with analysis did not isolate this condition when the valve and nozzle thrusters were used in a pulse mode. The thrusters for attitude control were used on an intermittent basis in flight with a nominal 20-msec pulse requested by the control system. The dynamic thrust provided by the nozzle is not a square wave as the pressure drop across the nozzle does not have an instantaneous build-up and drop-off. Any pressure, other than the pressure for which the thrust nozzle was designed, reduces the efficiency of the nozzle so that total impulse cannot be directly related to gas mass expulsion. In practice, the pressure drop-off after the valve closes provides a longer period of thrust and an angular velocity imparted to the spacecraft other than the actual velocity desired. The dynamic thrust levels of interest are from 0.1 lb to 0.0001 lb.

To measure dynamic thrusts of these levels requires the following conditions to be imposed on the measurement means: (1) Rapid response. (2) Vacuum environment. (3) Frictionless thruster mount and/or calibrated mounting. (4) Accurately readable parameter directly related to thrust. The means of measurement must also be decoupled from ambient vibration and protected from acoustical energy.

Fig. 13 shows an entire installation at one period of construction. The scheme is to mount the valve and nozzle, with its nitrogen (N_2) gas pressure source, on a gas-bearing table with all electrical connections across the gas bearing

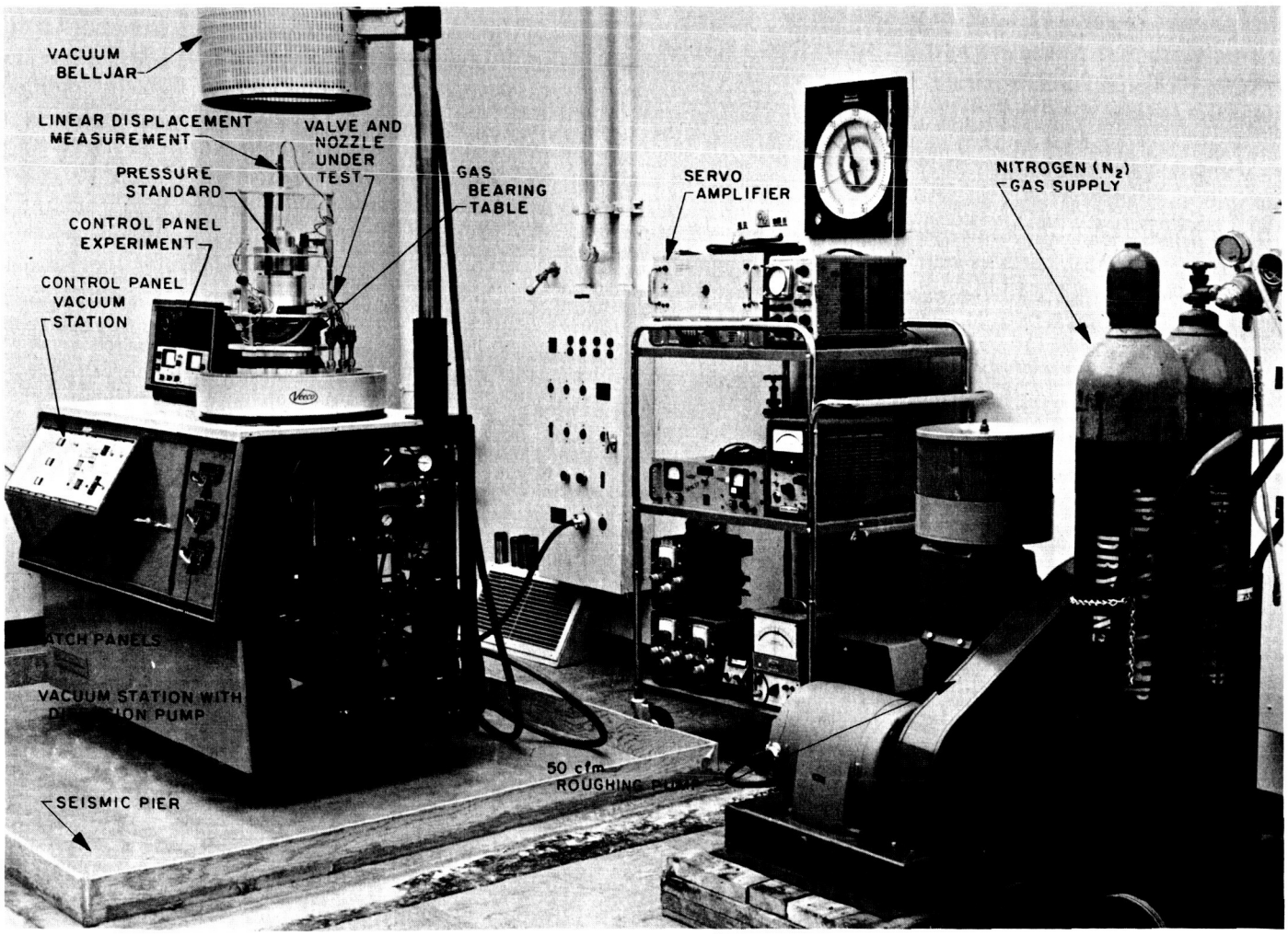


Fig. 13. Attitude control dynamic thrust measuring station under construction

through mercury puddles. The pressure source is N_2 gas filled initially, closed off and external connections removed. Pressure is obtained with weight upon a piston, piston travel assuring a constant pressure to the valve and nozzle until the gas in the pressure source is depleted. A clearance of 40 microinches (μ in.) between the piston and its cylinder reduces gas leakage to a negligible level. The N_2 gas volume, a function of piston travel, is measured during the testing of the valve and nozzle.

The gas-bearing table, shown in Fig. 14, has its rotor floating on N_2 gas cushions in such a manner that the rotor will not blow out of the stator when the table is placed in a vacuum. Flight-style Sun sensors are mounted on the stator with louvers mounted on the rotor. When light shines through the louvers to the Sun sensors, the shift and direction from the zero position are detected, providing an error signal. A 40 μ in. displacement from zero position is detected by the Sun sensor.

The error signal provided is directed into an adjustable-gain DC servoamplifier with a frequency response from DC to 10 Kc. Maximum output capability of the servoamplifier is 4 amp at 100 w.

The output of the servoamplifier is returned to the torquer on the air-bearing table. The torquer is built with an air core coil mounted upon the stator with a permanent magnet inserted into the coil. The magnet mount is on the rotor of the air-bearing table. The frequency response of the torquer is up to 40 Kc. The current into the torquer coil can be directly and accurately related to any rotational force on the rotor of the air-bearing table provided the torquer force is capable of matching the imposed force and the servoamplifier gain is properly set.

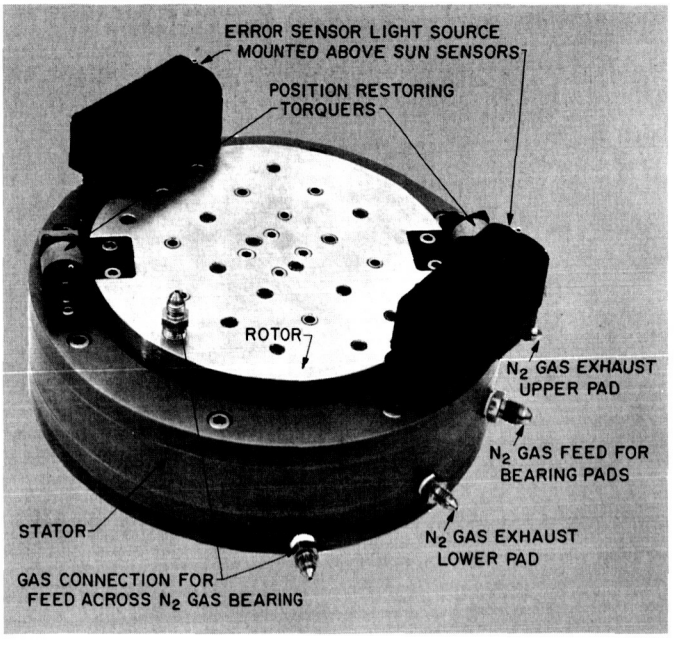


Fig. 14. Bearing table adapted for use in a force balance system

The vacuum system selected has an 18 × 30 in. belljar vacuum chamber. The vacuum in the chamber is initially pulled by a 50 ft³/min pump, then a diffusion pump replaces the roughing pump. The vacuum obtained is 10^{-4} torr pressure. While this does not duplicate the vacuum of space, the inaccuracy reflected in the results is negligible.

The test setup has been placed upon a concrete block 6 × 8 × 5 ft, resting on the earth 5 ft below the floor of the building. The only coupling between the building and the block is through the earth 5 ft below the floor level. There is no attachment to the block along its 5-ft sides.

It is expected that accurate measurements of instantaneous thrust of the attitude control thrusters will be obtainable in the very near future. When the vacuum thrust test station is in complete operation, a nozzle and valve test program will be started. A cooperative effort with the JPL aerodynamicists is planned. The test program will be to investigate attitude control thrusters used in a pulse mode. Such variables as pulse duration, nozzle sizing, feed pressure, and fuel will be investigated and correlated. This experimental data can then be applied to a particular attitude control flight requirement. A properly designed valve and nozzle can be readily designed to accurately meet the needs of attitude control. The finished valve and nozzle can also be accurately checked for proper performance before flight.

D. Bag-Type Quantitative Leak Measurements Using a Helium Mass Spectrometer

R. H. Summers and E. F. Koch

1. Purpose

The purpose of developing this measuring technique is to be able to determine in the shortest possible test time the total leakage from a complex pressure system required to operate unattended and without recharging for a long period.

2. Description

The test consists of two parts and requires two specialized equipments, (1) a helium mass spectrometer leak detector, and (2) a reasonably tight bag to enclose the equipment to be tested (Figs. 15, 16). The use of a specially fitted bag with a minimum volume permits the detection of smaller leaks in a shorter time. The first step of the test consists of calibrating the indicator on the spectrometer against a known helium leak. This is accomplished by allowing the helium to escape into a bag, which contains the test unit, and measuring the increase

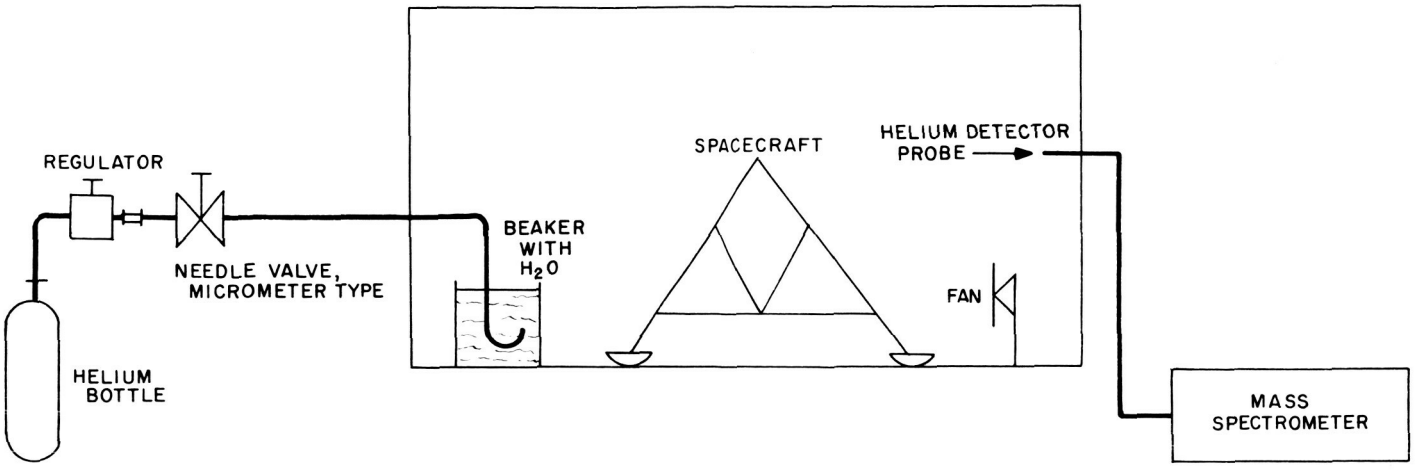


Fig. 15. Test configuration for bag-type leak measurements

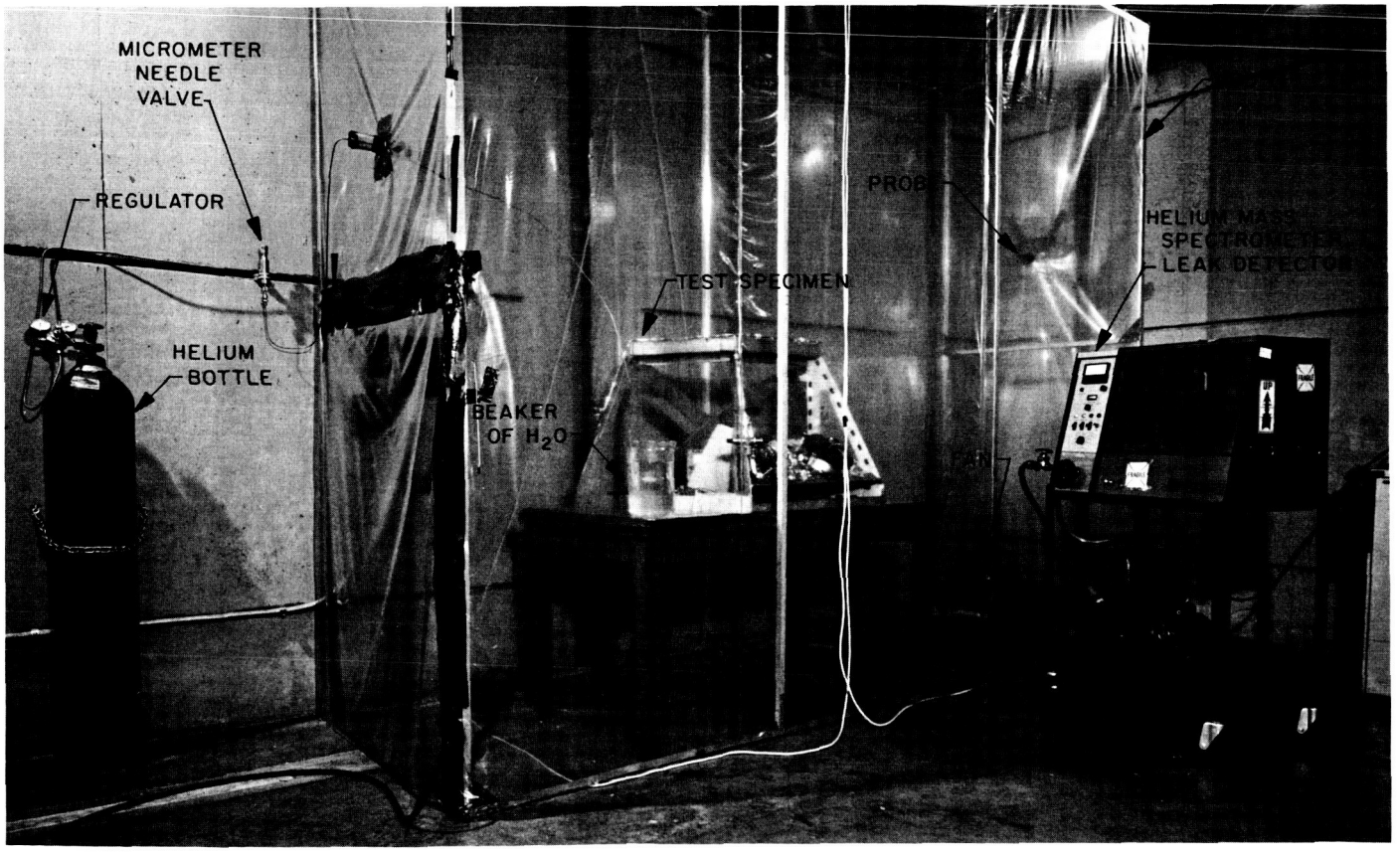


Fig. 16. Test setup for simulated leak measurements

in helium concentration with time. A plot of the concentration reading versus time should have a straight-line portion whose slope is "divisions per unit time." A careful measurement of the test leak by bubbling it through a beaker of water into a filled inverted graduated cylinder for a measured time interval provides a quantitative leak value for the slope of the calibration data. The second step consists of pressurizing the system to a known value with helium or a known proportion of helium and the working gas; then, a plot of the mass spectrometer readings against time will give another slope. By setting a simple ratio of the slope from the calibration data and the value of the known leak equal to the ratio of the test data slope to the unknown leak rate, the unknown value can be determined. Fig. 15 schematically shows the test configuration, and Fig. 16 shows the test set-up.

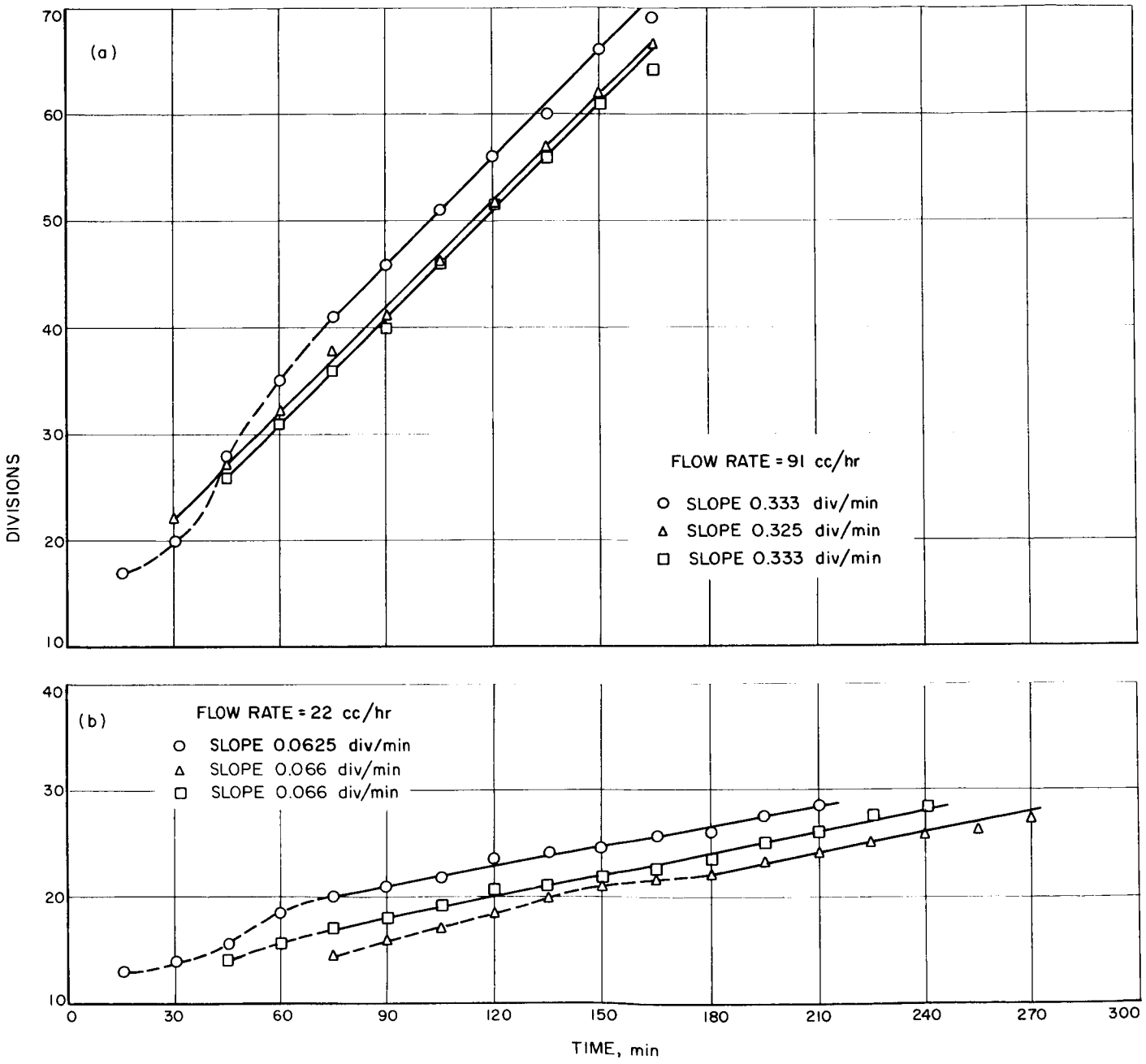


Fig. 17. Calibration data for test equipment

To date, the tests performed have simulated an actual test by simply changing a needle value setting from the known calibrated setting to a new position for the unknown leak. By measuring this leak rate after completing the mass spectrometer readings, the accuracy of the test can be evaluated. In addition several calibration runs have been made at each of three known leak rates in order to evaluate consistency of the testing with a secondary aim of determining the smallest leak detectable in the given bag, which had a volume of 480 ft^3.

3. Results

For the highest leak rates measured, 91 cc/hr of pure helium, the calibration of the mass spectrometer–bag combination gave values of 0.333 div/min, 0.333 div/min, and 0.325 div/min for three runs. Each test was conducted for 2.75 hr, with data at the end of 2 hr producing a satisfactory slope. The lowest calibrating rate at which repeatable data have been obtained is 22 cc/hr, with values of 0.066 div/min, 0.066 div/min, and 0.0625 div/min obtained. These tests were conducted for about 3.5 hr. Plots of the calibration data are shown in Fig. 17.

A calibrated slope of 0.066 div/min was obtained from a measured leak of 22 cc/hr in two separate tests. Each test was followed by a change in leak rate to simulate an unknown leak. In each case the plotted data had a slope of 0.20 div/min and the leak was later measured at 57 cc/hr. Using the slopes and initial calibration, the leak rate was calculated to be 66.7 cc/hr, as compared to the 57 cc/hr measured, or about 15% high. The results from these two tests are shown in Table 1, and a typical plot of the data in Fig. 18. Early test data, although plagued by mass spectrometer difficulties, indicates the possibility of measuring leaks as small as 8 cc/hr in the 480 ft^3 bag used.

These results required a total test time of less than 8 hr plus the time required to pressurized and later purge the system tested. It was concluded that the accuracy of the tests conducted was within $\pm 15\%$, and the consistency of calibration was excellent.

4. Discussion

Quantitative measurements of leakage can be made by other methods. One is to pressurize the system while connected to a known leak-tight system with a shutoff valve

Table 1. Data from simulated leak test

Test	Run	Plotted slope, div/min	Measured leak rate, cc/hr	Calculated leak rate, cc/hr
1	1	0.066	22	—
	2	0.20	57	66.7
2	1	0.066	22	—
	2	0.20	57	66.7

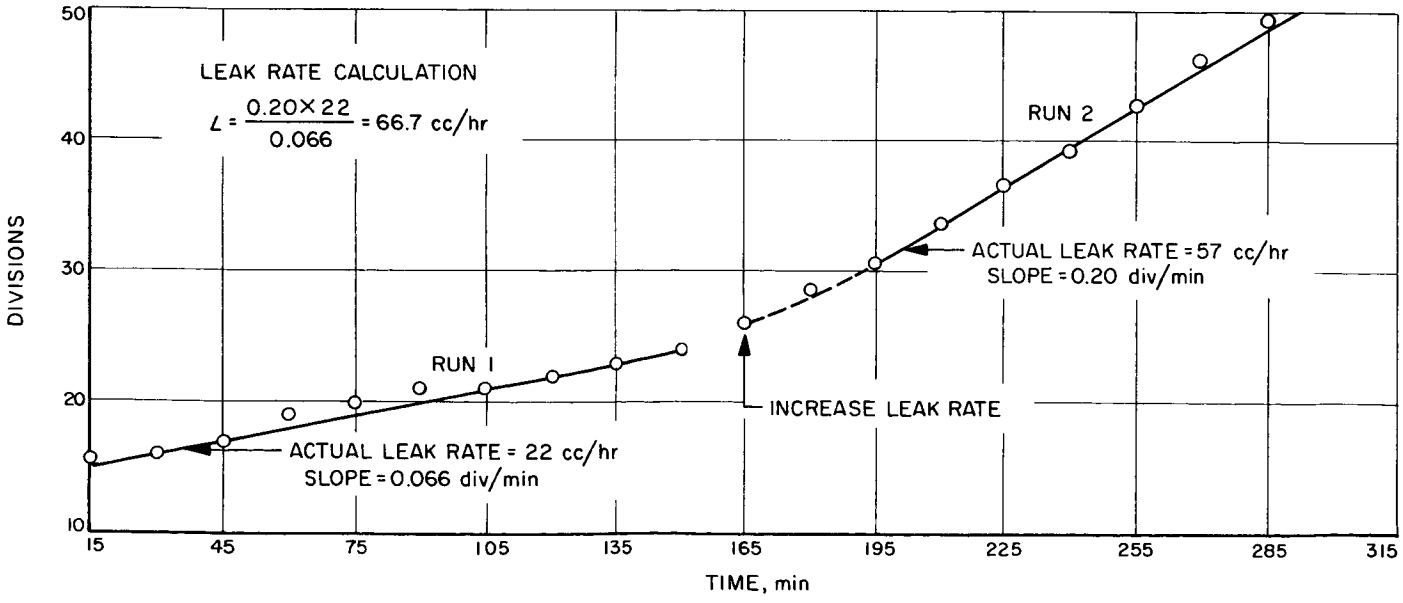


Fig. 18. Data from simulated leak test

and a differential pressure gage in parallel between them. Once the temperature has stabilized and the pressure equalized, the valve is closed. Now, if the internal volume of the system being checked is known and the temperature of both the test system and the leak-tight system are the same, readings of the differential pressure gage and the elapsed time since closing the valve will permit calculation of a total leak rate. In general, for the low leak rates required for *Ranger* and *Mariner* missions, several days were required to get a significant measurement. Another very similar method is to substitute an accurate pressure gage. The accurate pressure gage is substituted for the known leak-tight system and differential pressure gage combination. The differential pressure gage is not used. This normally takes even longer because the least count of the gage measuring total pressure is greater than that of the differential pressure gage. Both of these methods are very sensitive to temperature fluctuations which can cause pressure changes greater than the leak. Still other methods are possible but they have limitations of portability, or require very special equipment or long test times.

The basic assumption of this bag test technique is that all gross leaks are sealed and therefore the leakage is by molecular flow; that is, the hole has a radius approximately equal to the mean free path of the molecules. Then the leak will be directly proportional to the pressure differential across the hole. Unless the pressure in the system causes physical distortion of the parts to open existing and cause new leaks, tests can be run at pressures below the working pressure and can be converted to final conditions. Other modifications to the test data may be necessary to evaluate the final total leakage. The simplest is to account for the proportions of helium and the working gas used to pressurize the system. It is also necessary to correct for the difference in molecular flow of the gases used if helium is not the working gas. For nitrogen, this factor is 2.65 and is determined as

$$\left(\frac{\text{molecular wt. } N_2}{\text{molecular wt. He}}\right)^{1/2} = \left(\frac{28}{4}\right)^{1/2} = 2.65$$

The special merits of this test technique do not arise from the most general application of the helium mass spectrometer leak detector, where either the external or internal area of the system is evacuated to a hard vacuum. Here the system to be tested is too large to place in the evacuated chamber easily and cannot itself be evacuated because of the operational characteristics of the pressure regulator. By pressurizing the system with helium or a mixture of helium and another gas to the working pressure, the actual service leaks will be checked. Sampling the surrounding atmosphere limited by a close-fitting bag with a probe which can be evacuated provides means for quantitative evaluation of the sum total of the leaks.

Emphasis on a close-fitting bag is justified by the observation that it is helium concentration detected by the mass spectrometer and a very small leak into a large volume changes the concentration very slowly. A large volume bag could limit the minimum total leak detectable. Previously the bag was described as reasonably tight. Since gas is flowing into it through the leaks, the bag will become pressurized and will leak also. If the leakage of the bag is large and that of the system small, it is possible that the helium concentration will not be increased above the sensitivity of the mass spectrometer due to the diffusion and leakage of helium through the bag. This same situation causes the requirement of a close-fitting bag so loss of helium does not limit the test. It is assumed here that little disturbance of the bag will take place between the calibration and actual test, or that joints can be sealed effectively in a repeatable manner so the effects of bag leaks are accounted for in the calibration run.

Variations in test and calibration are possible. They include filling internal voids to reduce the volume of the bag, calibration by ejecting a fixed known quantity of helium into the bag and measuring an equilibrium concentration, and using a constant flowing inert (N_2) atmosphere in the bag, and sampling the helium in the outflow. These have not been evaluated because of increased complexity over the methods used. A go, no-go type test can be made by simply setting the calibrating leak at the maximum value permitted and checking that the test data fits a line plotted below the calibration curve, i.e., with a lower slope of div/min.

References

1. Garner, H. D., Reid, Jr., H. J. E., *Simulator Studies of Simple Attitude Control for Spin Stabilization Vehicles*, NASA TN D-1395, National Aeronautics and Space Administration, Washington, D. C., September 1962.

2. Le Compte, G. W., and Bland, J. G., "Simply Mechanized Attitude Control for Spinning Vehicles," *Journal of Spacecraft and Rockets*, Vol. 1 (No. 6): pp. 593–8, 1964.

V. Guidance and Control Research

A. Some Electrical and Metallurgical Properties of the Fabricated Germanium Solid-State Diodes

A. Shumka

1. Introduction

Space-charge-limited (SCL) hole current is being investigated in near-intrinsic high-purity germanium. For this investigation it is required that alloyed germanium n^+-π-n^+ planar solid-state diodes be fabricated. The procedure for fabricating such structures is described in *SPS 37-32*, Vol. IV, pp. 64–67. Electrical and metallurgical tests were performed to determine whether or not these structures contain flat, parallel and abrupt n^+-π junctions which exhibit predictable electrical properties.

2. Evaluation of the Solid-State Diodes

a. I-V Characteristics of the n^+-π junctions. Reverse *I-V* characteristics are measured separately for the two junctions of each solid-state diode to check for any serious defects in the junction structure and for surface leakage. Any observed low-voltage breakdown is attributed to junction defects, while low reverse-bias resistance is attributed to leakage effects. The former effect, if consistently present, would indicate the need for a major modification in the alloying process, while the latter would indicate the need for surface retreatment. The breakdown

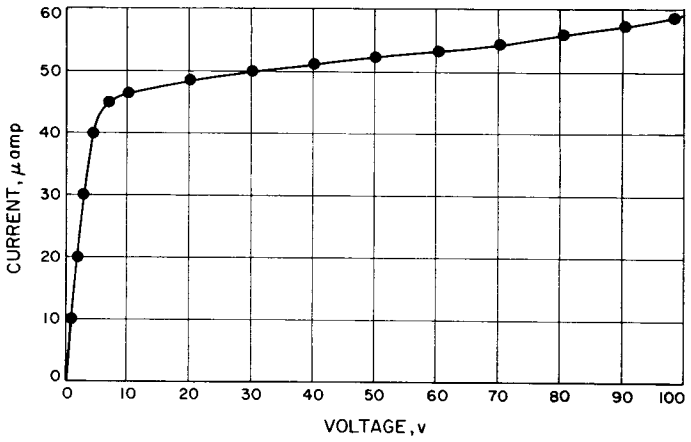


Fig. 1. A reverse-biased n^+-π junction *I-V* characteristic of a typical n^+-π-n^+ solid-state diode

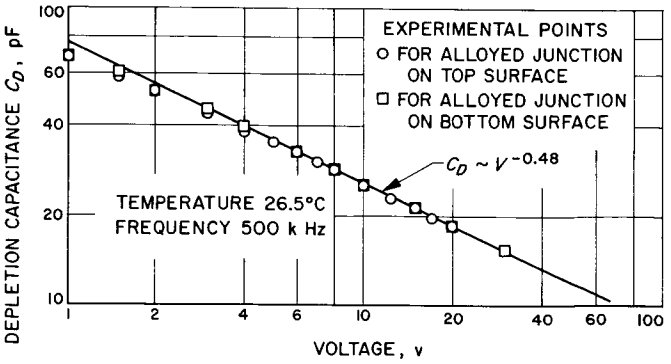


Fig. 2. Junction depletion capacitance as a function of voltage

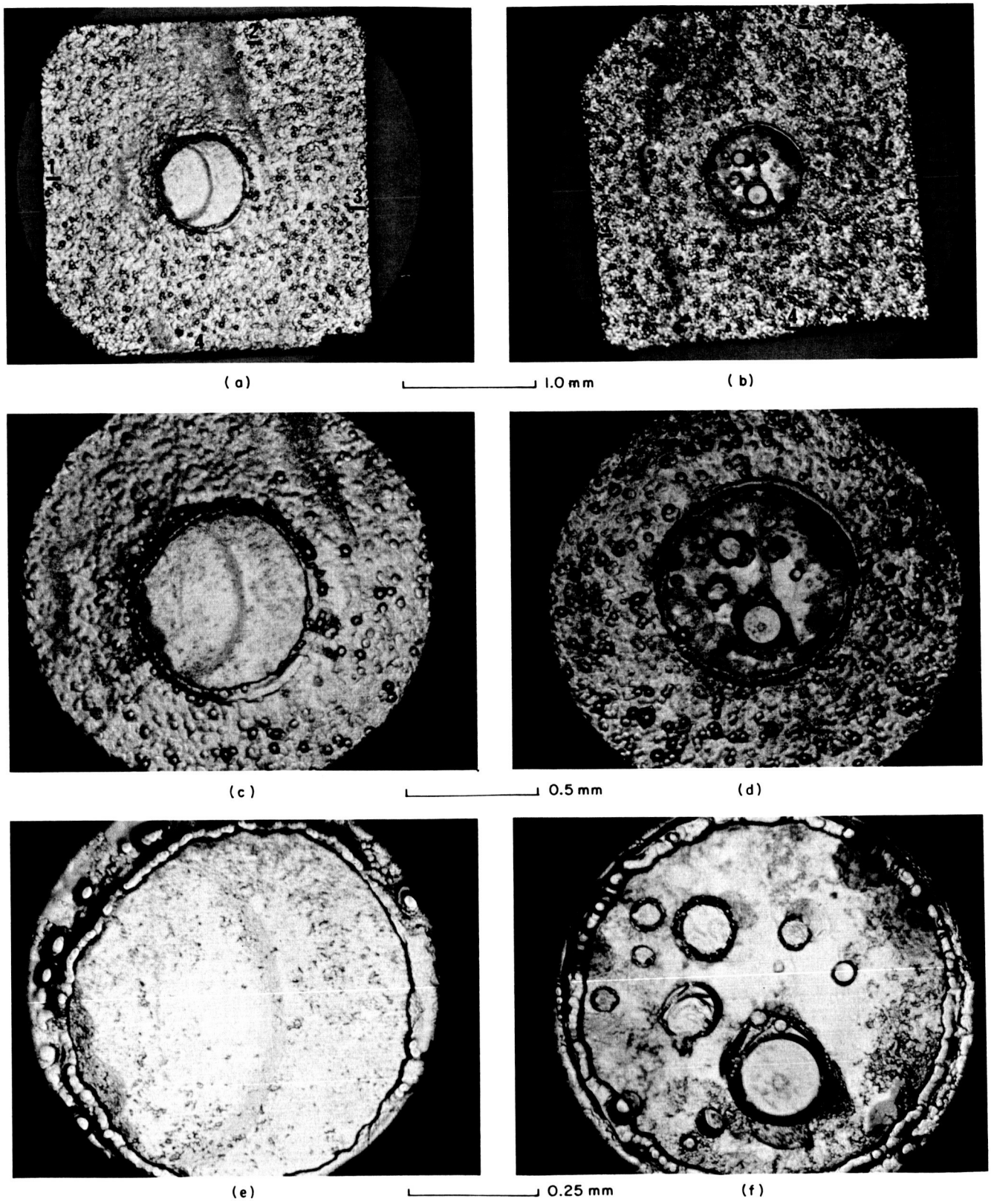


Fig. 3. Recrystallized surfaces of a typical solid-state diode, where photomicrographs (a), (c), (e) are of the top surface and (b), (d), (f) are of the bottom surface

voltage was less than 120 v in only two of eighteen fabricated solid-state diodes. It was necessary to chemically retreat the surface in only one solid-state diode. The current at 50 v for each reverse-biased junction of a given solid-state diode did not differ by more than 5%, indicating close similarity between the two junctions in a given diode. A typical reverse *I-V* characteristic is shown in Fig. 1.

b. Depletion capacitance measurement. Abruptness in the impurity profile of the alloyed junctions is determined from depletion capacitance measurements. Two diodes with the same junction area were alloyed on 40-mil 1.7–1.2 Ω-cm *p*-type germanium to determine whether or not the junctions alloyed on the bottom of the wafer and on the top of the wafer (*SPS* 37-32) exhibit similarly abrupt impurity profiles. Depletion capacitance measurements on these diodes, as shown in Fig. 2, indicate that the bottom and the top of the fabricated solid-state diodes are similar. The capacitance-voltage relation $C_D \sim V^{-0.48}$ is in good agreement with $C_D \sim V^{-0.5}$ which is expected from the abrupt junction theory.

c. Metallurgical analysis. Metallurgical cross-sections of the solid-state diodes make it possible to determine unambiguously whether or not parallel and planar alloyed junctions are obtained and to determine accurately the separation distance between the junctions (base width). Also, a visual inspection of the recrystallized surfaces may qualitatively indicate variations in the alloying process from sample to sample as well as differences in the alloying process between the top and the bottom of each wafer.

The recrystallized surfaces of a typical solid-state diode are shown in Fig. 3. These surfaces are exposed by completely dissolving the lead-antimony mixture in a solution containing equal parts of glacial acetic acid and hydrogen peroxide. The degree of alignment and the area of the two junctions are measured from the photomicrographs in Figs. 3(a) and (b), and in Figs. 3(c) and (d), respectively. The photomicrographs in Figs. 3(e) and (f) clearly indicate a marked difference between the two recrystallized surfaces. The top recrystallized surface shown in the photomicrograph, Fig. 3(e), contains a grain structure while the bottom surface, Fig. 3(f), contains pits. From the metallurgical cuts it has been determined qualitatively that if there are any nonuniformities at the junction interface, they appear to occur at grain boundaries. It is also observed that the pits do not cause any discontinuities at the junction interfaces. At present, an attempt is being made to eliminate any thermal gradients that may occur along the wafer during the alloying process with the intent to minimize the number of boundaries on top of the surface. Metallurgical cuts are made in the following way: solid-state diodes, which are encased in a clear epoxy holder, are orthogonally cross-sectioned by lapping and then polished to remove all scratches. The polished surface is then etched in a solution of 20 ppV H_2O_2 and 1 ppV

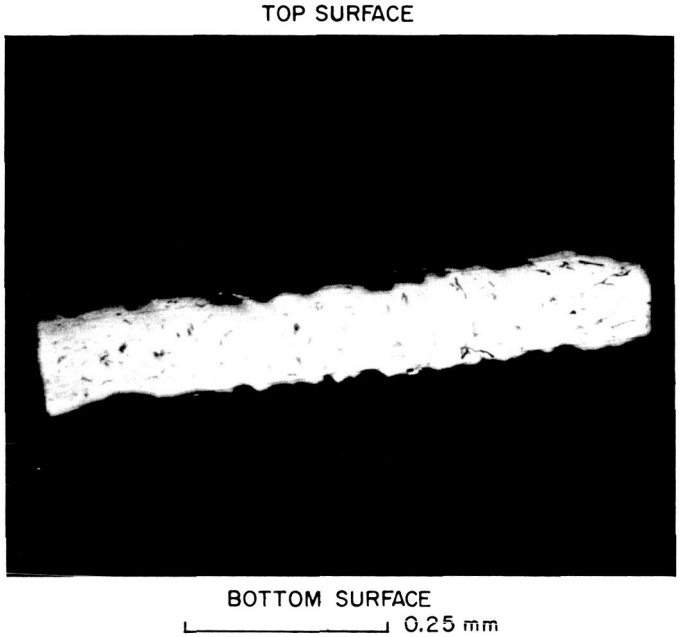


Fig. 4. A cross-sectional cut of a solid-state diode with parallel junction interfaces

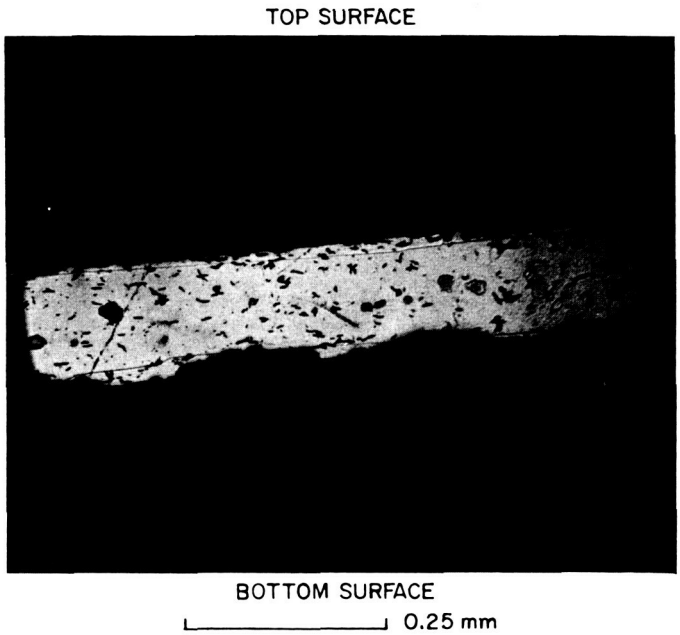


Fig. 5. A cross-sectional cut of a solid-state diode with non-parallel junction interfaces

HF for 45 sec. A photomicrograph of a cross-sectional cut of the solid-state diode of Fig. 3 is shown in Fig. 4. It is evident that the displayed junction interfaces are parallel and planar even though the edges of the cross section had been chipped while being lapped. The photomicrograph also clearly shows that the two junctions are not perfectly aligned. In almost all cases the alignment is much better.

From the series of tests described above it was usually found that the structures did contain flat, parallel and abrupt n^+-π junctions which exhibited predictable electrical properties. However, in some samples non-parallel junctions, as shown in Fig. 5, with low breakdown voltage were observed. In order to prevent the occurrence of such effects, the alloying process, as already mentioned, is being modified.

B. Measurement of the Second-Harmonic Generation Coefficients in Barium Titanate

M. S. Shumate

An experimental apparatus for determining the optical second-harmonic generation (SHG) coefficients of barium titanate has been constructed, and the results of the measurements performed are reported here. The theory of optical second-harmonic generation in barium titanate has been reported previously (*SPS 37-31*, Vol. IV, pp. 95–97). The SHG coefficients are determined by illuminating a crystal sample with radiation of known intensity from a ruby optical maser and measuring the intensity of the second-harmonic radiation which is induced at the front surface of the sample.

1. Application of the Theory to the Experiment

The expressions for the magnitude of the electric field strength of the second harmonic light wave reflected from the front face of an x-cut crystal of barium titanate, as derived in *SPS 37-31*, Vol. IV, are:

$$E_{\perp}^{(2\omega)r} = \frac{-4\pi \left(E_{\perp}^{(\omega)s}\right)^2 \sin^2\theta_t \sin\theta_s}{\sin(\theta_r + \theta_t)\sin(\theta_s + \theta_t)\sin\theta_r} \times \left[(d_{15} + d_{31})\sin\phi\cos^2\phi + d_{33}\sin^3\phi\right] \quad (1a)^*$$

$$E_{\parallel}^{(2\omega)r} = \frac{4\pi \left(E_{\perp}^{(\omega)s}\right)^2 \sin\theta_s \sin^2\theta_t \cos\theta_r}{\sin(\theta_t + \theta_s)\sin(\theta_t + \theta_r)\cos(\theta_t - \theta_r)\sin\theta_r} \times \left[(d_{33} - d_{15})\sin^2\phi\cos\phi + d_{31}\cos^3\phi\right] \quad (1b)^*$$

where the angles θ_s, θ_t, and θ_r are functions of the angle of incidence, and are fixed for this experiment. The angle ϕ, the angle between the sample's c-axis and the plane of incidence, is the independent variable, and the coefficients d_{15}, d_{31} and d_{33} are the second-harmonic generation coefficients which are to be determined.

The type of optical maser to be used for the light source radically affects the way in which the above equations are interrupted. The output of an unmodified optical maser usually consists of many transverse and longitudinal modes simultaneously. It has been previously pointed out (Refs. 1, 2) that the effect of multimoding of the optical maser generally leads to a reduction in the amount of second harmonic produced due to interference effects. If it is assumed that the output of the optical maser consists of many independent frequency components, the magnitude of E_1^2 in Eqs. (1a) and (1b) is reduced by a factor of $\sqrt{2}$ (Ref. 3). The effects of the multimoding can be reduced further by averaging the second harmonic output over many pulses from the optical maser.

Averaging is accomplished by measuring the total energy in each beam (incident fundamental and resulting second harmonic) during one complete optical maser flash. Since the ruby optical maser used for these measurements has an output that is not constant or continuous during a flash, it is necessary to make a correction to the measured values in order to use Eqs. (1a) and (1b). In order to determine the amount of the correction, it is necessary to assume a specific time dependence for the intensity of the optical maser output pulse.

A typical ruby optical maser output intensity consists of a series of triangular pulses. The height, width, and spacing of these pulses varies from pulse to pulse, but not over a large range. For the purposes of the correction mentioned above, it is assumed that the optical maser output consists of a series of uniformly spaced, equal size triangular pulses. If the relationship between the incident fundamental power $\pi(\omega)$ and the resulting second harmonic power $\pi(2\omega)$ is

$$\pi(2\omega) = \alpha\,\pi^2(\omega) \quad (2)$$

*These equations are presented here in correct form, since they were incorrect in *SPS 37-31*, Vol. IV, p. 97.

then the relationship between the total energies is

$$W(2\omega) = \frac{8}{3}\frac{\alpha}{nt_0} W^2(\omega) \tag{3}$$

where

$$W(\omega) = \int_{one\ shot} \pi(\omega)\, dt$$

$$W(2\omega) = \int_{one\ shot} \pi(2\omega)\, dt$$

and where n is the average number of pulses per shot, and t_0 the average base width of one pulse. The power in a beam is related to the electric field strength in the beam (neglecting diffraction effects) by

$$\pi(\omega) = \frac{\epsilon c A}{8\pi} E^2(\omega) \tag{4}$$

where A is the beam area. Substitution of Eqs. (1a) or (1b) into Eqs. (4), (2) and (3) gives the following expressions for the two energy components of the second-harmonic beam, where an additional factor of $\sqrt{2}$ is included to account for multimoding:

$$W_\perp(2\omega) = \frac{2^{15}\pi^3}{3\epsilon c A n t_0} \cdot \left[\frac{\sin^6\theta_s \cos^4\theta_i \sin^4\theta_t}{\sin^4(\theta_i+\theta_s)\sin^2(\theta_r+\theta_t)\sin^2(\theta_s+\theta_t)\sin^2\theta_r}\right] \times \left[(d_{15}+d_{31})\sin\phi\cos^2\phi + d_{33}\sin^3\phi\right]^2 W_\perp^2(\omega) \tag{5a}$$

$$W_\parallel(2\omega) = \frac{2^{15}\pi^3}{3\epsilon c A n t_0} \cdot \left[\frac{\sin^6\theta_s \cos^4\theta_i \sin^4\theta_t \cos^2\theta_t}{\sin^4(\theta_i+\theta_s)\sin^2(\theta_t+\theta_s)\sin^2(\theta_t+\theta_r)\cos^2(\theta_t-\theta_r)\sin^2\theta_s}\right] \times \left[(d_{33}-d_{15})\sin^2\phi\cos\phi + d_{31}\cos^3\phi\right]^2 W_\perp^2(\omega) \tag{5b}$$

2. Experimental Apparatus

The apparatus for performing the measurements, as shown in Fig. 6, is completely enclosed in light-tight boxes to prevent the possibility of scattered optical maser radiation from giving spurious output readings. The ruby optical maser used for these measurements was of the normal (non-Q-spoiled) type, with a typical output beam having a wavelength of 6943 Å and consisting of a series of short-duration pulses (~0.4 μsec wide) lasting for a few hundred microseconds during each flash. The ruby crystal used was a 90-deg oriented type, and was positioned so that the plane of polarization of the emitted radiation was perpendicular to the plane of incidence of the crystal sample (the plane of Fig. 6).

The optical maser beam enters the apparatus through a deep red glass filter to eliminate as much flash-lamp light as possible. An interference type beam splitter, which transmits approximately 0.1% into a phototube used as a monitor for the optical maser radiation, reflects approximately 99% of the beam onto the crystal sample under test. The sample is mounted in such a way that it can be rotated about the normal to its front surface. The light beams which leave the surface of the sample, the fundamental and the second harmonic, are passed through a liquid filter of saturated copper sulfate solution which absorbs the fundamental but transmits the second harmonic. This is followed by a Glan-Thompson prism used to analyze the polarization of the second harmonic, by a narrow-bandpass interference filter whose transmission is centered at 3471 Å, and by a low dark-current type ultraviolet photomultiplier tube. The overall sensitivity of the filter-photomultiplier tube combination is sufficient to give one photoelectron for every 100 second-harmonic photons leaving the sample. As mentioned previously, the output of both monitor phototubes was integrated during each burst from the optical maser, giving output records which are proportional to the total energy in each beam. The sensitivities of both the fundamental monitor phototube and the second-harmonic monitor photomultiplier tube were carefully measured at their respective wavelengths, 6943 Å and 3471 Å.

The crystal samples used are all transparent to the ruby optical maser beam, so there is a fundamental component reflected from the rear surface which returns to the front surface of the crystal sample; if the sample faces are parallel, it proceeds in the same direction as the fundamental component which is reflected from the front surface. As this returning fundamental component passes back through the crystal, the second harmonic is produced, and will hence interfere with the measurement of the reflected second harmonic emanating from the front surface. The effects of this back-reflected component can be almost completely eliminated by either using antireflection coatings on the sample or by using a wedge-shaped sample.

3. Results

The method outlined above has been used to determine the second-harmonic generation coefficients of two different crystals: potassium dihydrogen phosphate (KDP) and barium titanate ($BaTiO_3$).

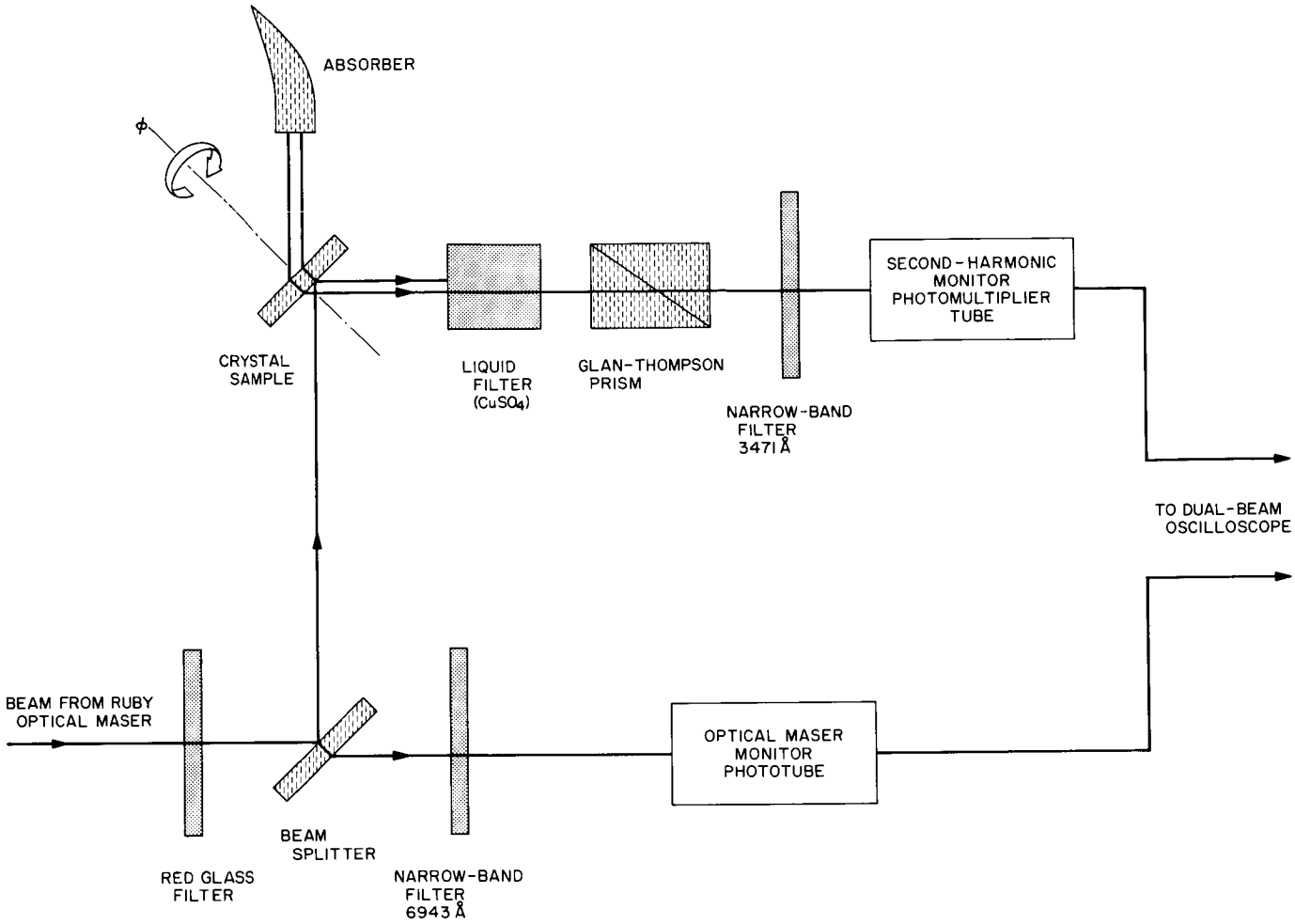


Fig. 6. Experimental apparatus for measuring second-harmonic generation coefficients

The SHG coefficient of KDP has been very carefully determined by Ashkin, Boyd, and Dziedzic (Ref. 2) using a single-mode gas optical maser. The purpose of using the present apparatus to determine the SHG coefficient of KDP was to check the accuracy of the calibration and to verify the assumptions regarding multimoding and averaging. The KDP crystals used for the present measurements were z-cut plates 15 × 15 × 3 mm. In order to eliminate the back-reflected ray, a prism of polymethyl methacrylate was contacted to the back face of the KDP crystal using a contact oil. The index of refraction of the oil and the prism closely matched that of KDP, so that very little reflection occurred at the back face of the KDP crystal. The angle of the prism was chosen so that the return beam went in a direction where it would not enter the second-

Table 1. Potassium dihydrogen phosphate result

Optical maser wavelength, μ	d_{36}
0.6943	$2 \pm 2 \times 10^{-9}$
1.15[a]	$3 \pm 1 \times 10^{-9}$

[a]Ref. 2.

Table 2. Barium titanate results at a temperature of 23°C

Optical maser wavelength, μ	d_{15}	d_{31}	d_{33}
0.6943	$1.6 \pm 0.5 \times 10^{-7}$	$6.0 \pm 2 \times 10^{-8}$	$5.8 \pm 1 \times 10^{-8}$
1.06[a]	$1.05 \pm 0.09 \times 10^{-7}$	$1.11 \pm 0.09 \times 10^{-7}$	$4.2 \pm 0.3 \times 10^{-8}$

[a]Ref. 4.

harmonic detection apparatus. The result of this measurement is summarized in Table 1. The agreement with Ashkin, et al. (Ref. 2), is quite good.

The three SHG coefficients of $BaTiO_3$ have been determined for a temperature of 23°C, using Eqs. (5a) and (5b). The results are presented in Table 2, and are compared with Miller's measurements (Ref. 4). The intensity of the reflected second harmonic exhibits the expected dependence on the angle ϕ.

There are some significant differences in the SHG coefficients of $BaTiO_3$ as measured at the two different wavelengths. This is due to the presence of an absorption edge at 4000 Å in crystalline $BaTiO_3$. The crystal is transparent to the second harmonic of 1.06 μ, but not to 0.6943 μ. It is generally expected that the SHG coefficients will increase when the second-harmonic wavelength passes a band edge (Refs. 5, 6). It should also be noted that Kleinman's symmetry condition (Ref. 7), $d_{15} = d_{31}$ for barium titanate, is not satisfied as expected due to the presence of the band edge.

Further measurements in the temperature range from 23 to 100°C are being performed, and the results will be reported in the near future.

References

1. Ducuing, J., and Bloembergen, N., *Physical Review Letters, 10*: 474, 1963.
2. Ashkin, A., Boyd, G. D., Dziedzic, J. M., *Physical Review Letters, 11*: 14, 1963.
3. Maker, P. D., and Terhune, R. W., *Physical Review, 137*: A801, 1965.
4. Miller, R. C., *Physical Review Letters, 11*: 146, 1963.
5. Soref, R. A., and Moos, H. W., *Journal of Applied Physics, 35*: 2152, 1964.
6. Chang, R. K., Ducuing, J., and Bloembergen, N., *Physical Review Letters, 15*: 415, 1965.
7. Kleinman, D. A., *Physical Review, 126*: 1977, 1962.

VI. Flight Computers and Sequencers

A. Reliability Techniques for Spacecraft Computer Systems

A. Avizienis

1. Aspects of Reliability

The principal aspects of computer reliability which require consideration are: (1) Selection of reliable components, quality control in production, and reliable packaging techniques. (2) Error-free logic design, construction, system programs, and mission programs.

The first aspect of computer reliability includes the ever-present requirements for quality control and reliable packaging, as well as the continued study of the technologies available for the construction of spacecraft computers. At the present time integrated microelectronic and magnetic logic circuits are the prevalent choices for implementation of advanced guidance computer systems. Several other technologies also merit consideration and may be preferable under special conditions; e.g., fluid logic elements for high-temperature environments.

The second aspect, error-free design, is attained by computer simulation of the logic design and by special checkout programs for the computer system. The verification of the system programs and mission programs is the second important problem in the preflight reliability effort. The second aspect also includes the final prelaunch checkout of the computer system in its operational environment.

Given a reliable system at mission launch time, the remaining problem is the control of malfunctions in the computing system during the mission; that is, the emergency backup for the above discussed efforts. Component or interconnection failures are controllable by the introduction of redundancy of hardware and/or software in the computing system. Previously-undetected design and/or program errors may also occur; redundancy is also applicable for the detection and remedial of such errors.

2. Application of Protective Redundancy

A computer system contains *protective redundancy* if effects of component failures (or program errors) can be tolerated because of the use of additional components or

programs or the use of more time for the computational tasks. These additional components, programs, and time are not required to execute the specified tasks as long as no failures or transient malfunctions occur.

The *direct symptoms* of computer failures and malfunctions are of two kinds:

(1) In logic circuits the symptom is a *logic fault;* e.g., a logic signal being "stuck on zero," or "stuck on one."

(2) In analog circuits and in some digital circuits such as clocks, the symptom is a *change of parameters:* e.g., a power supply variation, a change of frequency or loss of synchronism.

The *indirect symptoms* which lead to incorrect computation also are of two kinds:

(1) Damaged data and results.

(2) Damaged programs and/or their faulty execution.

The function of protective redundancy is to supplement the available checkout and maintenance procedures in the prevention of incorrect computation. Specifically, redundancy is essential when checkout and maintenance are not available either because of inaccessibility (e.g., during unmanned missions) or because of real-time computing requirements (e.g., during injection or approach guidance). The methods of introducing protective redundancy fall into two categories: *masking* or *massive* redundancy and *standby* or *selective* redundancy.

3. Masking or Massive Redundancy

In this approach the effect of a faulty computer component, circuit, subsystem, or system is *masked* instantaneously by permanently connected and concurrently operating replicas of the faulty element. The principal methods are:

(1) Replication of components; e.g., "quadded" diodes, resistors, and transistors, or duplicated connections (Refs. 1, 2).

(2) Replication of logic signals; use of voting elements, recursive nets, interwoven logic, and variation-tolerant threshold element nets (Refs. 3, 4, 5, 6).

(3) Adaptive logic elements; e.g., voters with variable weight inputs (Ref. 7).

(4) Immediate error correction; application of coding with combinational error-correction circuits (Refs. 8, 9).

The general advantages and disadvantages of the massive redundancy approach are presented, in comparison to selective redundancy, in Table 1.

Table 1. Comparison of massive and selective redundancy

Property	Massive redundancy	Selective redundancy
Occurrence of corrective action	Immediate	Delayed
Requirements for diagnosis	None during mission; difficult during maintenance	Required continuously
Variable interconnections	None	Switching system for replacement or reorganization
"Hard core" (unprotected) elements	May exist at interfaces	Always required
Component count increase	High; uniform throughout the entire system	Number of spares is adjustable to the mission
Design flexibility	Multiplexing is applied to the entire system	Subsystems may be individually changed and improved
Power requirements	Increased uniformly to supply all redundant components	Standby spares do not require power
Utilization of redundant elements	System operative until faulty elements are in majority	System operative until last spare fails
Catastrophic (non-independent) failures	May occur by channel interaction at connection points	Switching array provides isolation between subsystems
Conversion of non-redundant system to redundant form	Relatively straightforward	Not directly evident; new organizations may offer advantages
Experience in design and operation of systems	Several systems exist (Refs. 2, 4)	Theory and development is in an early stage

The important problem areas that are still unsolved, or only partially solved, in the application of massive redundancy are as follows:

(1) The generally employed assumption of independent failures of logic circuits may be invalidated in new technologies. The principal causes are: (a) the same defect may occur in all circuits of the same batch in batch-fabricated hardware; (b) isolation of individual logic circuits cannot be guaranteed due to physical proximity of micro-electronic components; e.g., mechanical, electrical, thermal damage may affect the entire can or pack. Catastrophic non-independent failures also require consideration.

(2) Optimum location of voting elements is to be sought, considering the relative reliability of voters and other logic circuits, the isolation of channels to minimize effects of catastrophic (non-independent) failures, and the optimal complexity of logic blocks between voters. Results of work in this area are rather incomplete; excessive computational time is required for larger logic nets. Heuristic programs are a promising area for investigation.

(3) The initial checkout and maintenance of multiplexed masking systems require special circuits such as "disagreement detectors" as well as special techniques, since initial or pre-mission faults are also masked.

(4) The increased number of interconnections increases the interface (fan-out, fan-in) requirements on logic circuits, especially in interwoven logic and component replication. In the latter case, circuit parameter margins and operation times may deteriorate, and initial checkout is difficult to implement.

(5) In sequential logic circuits which employ voting techniques, synchronization of memory elements in the individual channels is an important problem, since the signals must reach the voter at the same time in order to perform the voting correctly.

(6) Protection of interfaces with other systems, such as actuators and inputs, requires systematic consideration; redundancy cannot simply be terminated at the interface.

(7) Effective methods are needed for the assessment of expected gains in life expectancy of the protected system. Both analytic methods and the computer simulation of fault masking systems require further development. As in (2) above, heuristic programs and Monte Carlo methods must be explored.

4. Standby or Selective Redundancy

In this approach, the presence of a faulty computer element is detected by observing the symptom of the failure, and then the fault is made harmless by a corrective action. The principal methods are:

(1) Delayed error correction using sequential error-correcting circuits for coded words, or using a properly functioning duplicate of the faulty subsystem (Ref.10).

(2) Replacement of a faulty element or system by a standby spare to provide automatic maintenance or self-repair; error correction is attained by recomputation, possibly retracing several steps in the program to a "rollback" point (Refs. 11, 12; *SPS 37-22*, Vol. IV, pp. 9–12, 1963).

(3) Reorganization of the multi-processor system and other "degradable" systems into a different computer configuration.

The general comparison to massive redundancy is given in Table 1. The important problem areas are:

(1) To develop the ability of the system to tolerate interruptions in computation, and to minimize their duration.

(2) To develop the ability of the system to tolerate delays between the occurrence, detection and elimination of faults by means of a "rollback" of programs to certain points or a choice of alternate programs.

(3) To develop effective diagnosis methods which detect the symptoms of faults causing both data and program errors at relatively low cost in time and hardware.

(4) To develop and incorporate parameter-monitoring analog circuits which cover undiagnosable catastrophic failures, such as power failure and clock stopping, and enable their repair.

(5) To define and to minimize the undiagnosable "hard core" of the system, and to develop methods for its protection, such as masking and outside diagnosis.

(6) To devise methods for systematic evaluation and optimization of the possible tradeoffs between time, hardware and program redundancies in a selectively redundant system.

(7) To utilize spare or standby elements for greater computing power during the early life of the system,

called the "multiprocessor" or "reorganization" approach, and to develop systematic optimization methods for such use.

5. The Diagnosis Problem

The pre-mission checkout and maintenance of a spacecraft computer as well as the application of selective redundancy systems during a mission both require that an effective diagnosis technique should exist which detects and locates a faulty element or subsystem.

Damaged data is detected by employing error-detecting codes or by duplication and comparison in transmission and computation. Damaged programs or errors in program execution are detectable by coding duplication and by monitoring of critical control functions. Symptoms of program damage include erratic clocks, stuck instruction counters, endless loops, non-returning interrupts, and unexpected overflow signals.

Special diagnostic programs which test every logic element for proper operation are a very potent instrument for maintenance and checkout (Refs. 13, 14, 15). Combined with selective power application and disablement of selected logic blocks, diagnostic programs are applicable to the checkout of massive-redundancy systems as well. The use of diagnostic programs during a mission is limited by their time and storage requirements and by the need to protect the diagnostic program itself against faults.

Further developments in the theory and application of diagnostic programs are quite essential in the overall development of advanced computing systems for spacecraft guidance.

References

1. Creveling, C. J., "Increasing the Reliability of Electronic Equipment by the Use of Redundant Circuits," *Proceedings of the IRE*, Vol. 44: pp. 509–515, April 1956.
2. Lewis, T. B., "Primary Processor and Data Storage Equipment for the Orbiting Astronomical Observatory," *IEEE Transactions on Electronic Computers*, Vol. EC-12 (No. 5): pp. 677–686, December 1963.
3. Lyons, R. E., and Vanderkulk, W., "The Use of Triple-Modular Redundancy to Improve Computer Reliability," *IBM Journal of Research and Development*, Vol. 6 (No. 2): pp. 200–209, April 1962.
4. Dickinson, M. M., Jackson, J. B., and Randa, G. C., "Saturn V Launch Vehicle Digital Computer and Data Adapter," *AFIPS Conference Proceedings*, Vol. 26 (1964 FJCC): pp. 501–516, 1964.
5. Tryon, J. G., "Quadded Logic," *Redundancy Techniques for Computing Systems*, pp. 205–228, Spartan Press, Inc., Washington, D. C., 1962.
6. Pierce, W. H., "Interwoven Redundant Logic," *Journal of the Franklin Institute*, Vol. 277 (No. 1): pp. 55–85, January 1964.
7. Pierce, W. H., *Improving the Reliability of Digital Systems by Redundancy and Adaptation*, Stanford U. Electronics Laboratory, Report No. 1552–3, July 17, 1963. (Also Ph.D. Thesis, Stanford U., 1963.)
8. Peterson, W. W., *Error Correcting Codes*, The MIT Press and John Wiley & Sons, Inc., New York, 1961.
9. Kautz, W. H., "Automatic Fault Detection in Combinational Switching Networks," *Proceedings of the Second Annual Symposium on Switching Circuit Theory and Logical Design* (Detroit: October 17–20, 1961), pp. 195–214; AIEE Publication S-134, 1961.

References (Cont'd)

10. Kautz, W. H., "Codes and Coding Circuitry for Automatic Error Correction Within Digital Systems." *Redundancy Techniques for Computing Systems,* pp. 152–195, Spartan Press, Inc., Washington, D. C., 1962.

11. Griesmer, J. E., Miller, R. E., and Roth, J. P., "The Design of Digital Circuits to Eliminate Catastrophic Failures," *Redundancy Techniques for Computing Systems,* pp. 328–348, Spartan Press, Inc., Washington, D. C., 1962.

12. Avizienis, A., "A Study of the Effectiveness of Fault-Detecting Codes for Binary Arithmetic," Technical Report No. 32-711, Jet Propulsion Laboratory, Pasadena, Calif., September 1, 1965.

13. Galey, J. M., Norby, R. E., and Roth, J. P., "Techniques for the Diagnosis of Switching Circuit Failures," *Proceedings of 2nd Annual Symposium on Switching Circuit Theory and Logical Design* (Detroit: October 17–20, 1961), pp. 152–160; AIEE Publication S-134, 1961.

14. Poage, J., "Derivation of Optimum Tests to Detect Faults in Combinational Circuits," Technical Report No. 18, Dept. of Electrical Engineering, Digital Systems Laboratory, Princeton University, March, 1962. Also, *Proceedings of the Symposium on Mathematical Theory of Automata,* pp. 483–528, New York, N.Y., April 24–26, 1962; Polytechnic Press, 1963.

15. Seshu, S., and Freeman, D. N., "The Diagnosis of Asynchronous Sequential Switching Systems." *IRE Transactions on Electronic Computers,* Vol. EC-11 (No. 4): pp. 459–465, August 1962.

ENGINEERING MECHANICS DIVISION

VII. Materials

A. Ceramic Research

M. H. Leipold

1. Pure Carbide

The primary effort during this reporting period has been continued development of techniques for the production of blanks for polycrystalline tantalum carbide mechanical test specimens. Earlier work (*SPS 37-30*, Vol. IV, p. 72) has suggested that satisfactory blanks could be prepared efficiently from commercially[1] available submicron powder. However, when subsequent lots of this powder were obtained and hot pressed they did not densify to the same extent under the same processing conditions. Table 1 shows the variations obtained in the hot-pressed densities, plus the characteristics of these as-received powders. Differences noted in the characteristics of these powders are small and do not correlate with the variations in the pressing conditions. Attempts to determine the physical basis for these variations will continue. Negotiations are under way with CIBA to obtain processing information and final characteristics of the various lots of powder so that correlation may be made with the required pressing conditions.

Special components to permit mechanical evaluation of the carbide specimens in the mechanical testing facility (*SPS 37-34*, Vol. IV, p. 87) have been designed and constructed. A resistively heated graphite furnace and graphite grips have been completed and will be installed in the environmental test chamber. An available high current capacity power supply is being modified and will be installed to energize the furnace.

[1]CIBA Corp., Summit, New Jersey.

2. Pure Oxide

At the present time, no general theory for the structure of the grain boundary in a ceramic exists. Dislocation array theories, generally satisfactory for metals, appear less satisfactory for ceramics because of problems with ion-size disparity, charge difference, and directionality

Table 1. Characteristics of submicron tantalum carbide powder

As received from vendor				As hot pressed		
Lot	Particle size X-ray diffraction, A	Lattice parameter, A	Color	Pressing conditions, °C	Pressing conditions, psi	Density %, theoretical
1-1	950	4.4531	Dark gold	1700	6500	100
				1800	6500	100
2-2	640	4.4511	Dark gray	2000	8000	90+
				1800	7000	78
2-3	540	4.4510	↓	2000	8000	99
2-4	452	4.4510		2000	8000	100
2-5	452	4.4484		2000	8000	90

of bonding. Since it will be necessary to have some understanding of the atomic structure in the ceramic grain boundary in order to understand its interaction with cracks, an attempt has been made to formulate a theory of ceramic grain boundary structure based on application of random network theory. Prof. Harold T. Smyth, School of Ceramics, Rutgers University, has attempted to apply such a theory to the structure of a grain boundary in MgO. Preliminary calculations and findings, indicating that such application has merit, will be presented in a JPL technical report. Attempts will be made to develop the theory further to permit critical evaluation.

B. Basal Plane Shear in Pyrolytic Carbon[2]

W. V. Kotlensky

Observing the structure of a graphite single crystal and considering that the binding within the layer planes is covalent and extremely strong and that the binding between layers is van der Waals and quite weak, it is apparent that the most important slip system which can operate is slip along the basal planes.

For simplicity, basal slip in the single crystal can be thought of as the spreading out of a deck of cards on a table, each card representing a basal plane. The basal planes in as-deposited pyrolytic carbon are not smooth surfaces, in fact they are quite wrinkled. In the deck-of-cards analogy where the cards are still roughly parallel to each other, the individual cards are wrinkled and slip of these wrinkled cards over each other requires a greater stress than in the unwrinkled case. This is perhaps an overly simplified representation of basal slip in pyrolytic carbon because the importance of other components such as crystallite size, cross-link bonding, layer ordering, etc., have not been delineated or specified; nevertheless, this crude example serves a useful purpose in picturing basal slip in pyrolytic carbon on a macroscopic scale and in understanding the affect of dewrinkling on basal slip.

As a continuation of earlier work on the basal plane shear behavior of pyrolytic carbon (*SPS 37-28*, Vol. IV, pp. 43–44 and *SPS 37-31*, Vol. IV, pp. 107–108), the effect of dewrinkling produced by thermal annealing at 3000°C and *c*-direction compression annealing at 2900°C has been investigated as a function of temperature and strain rate using the double shear specimen configuration. Thermal annealing was carried out on machined specimens in a 1.5-in. inside diameter graphite resistance tube furnace in a helium atmosphere. Compression annealing was carried out on pyrolytic carbon 1-in. cubes in a 6-in. inside diameter graphite resistance tube furnace in an argon atmosphere. The pyrolytic carbon cubes were held in a graphite die and a stress of 1400 psi was applied perpendicular to the deposition surface for 1 hr at the annealing temperature. Double shear configuration specimens were machined from the compression annealed blocks. Basal shear tests were made in the same apparatus described in *SPS 37-31*, and basal flow stress was calculated as before.

Fig. 1 is a plot of the basal flow stress as a function of test temperature for the new double shear configuration specimens after several prior treatments, as well as

[2]Also referred to as pyrolytic graphite.

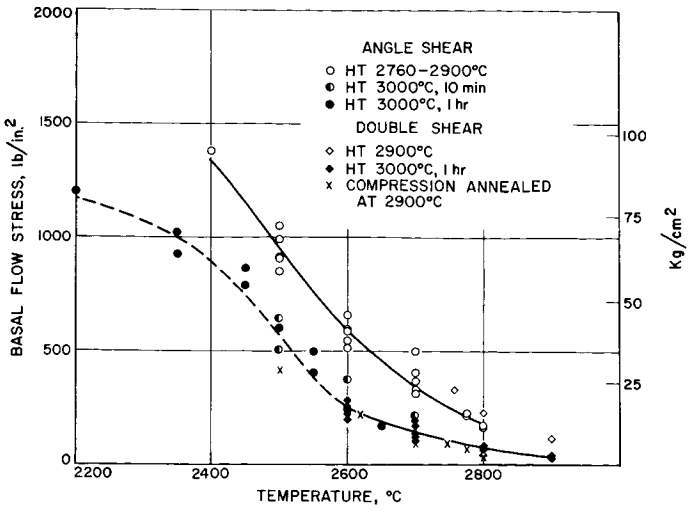


Fig. 1. Pyrolytic carbon flow stress versus test temperature

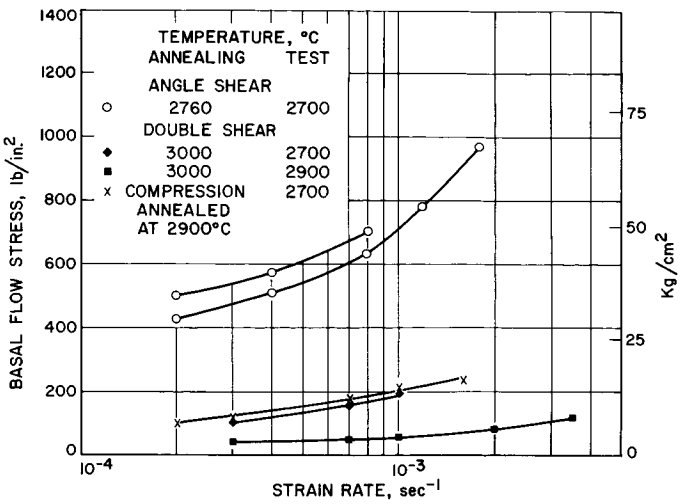


Fig. 2. Pyrolytic carbon flow stress versus strain rate at 2700 and 2900°C

data for the angle shear configuration reported previously. Two classes of behavior may be distinguished. The solid curve shows the basal flow stress–temperature dependence for the HT 2760–2900°C material; the dashed curve shows the dependence for the dewrinkled material. The significant points concerning this figure are: (1) The angle shear and double shear test data after the various similar treatments are in good agreement. (2) Prior dewrinkling, both by thermal and compression annealing, lowers the flow stress at a given temperature. For this pyrolytic carbon, the as-deposited material had a preferred orientation value, $\beta_{\frac{1}{2}}$ of 22; the 3000°C annealed material, a value of 7; and the compression annealed material a value of 2. The latter two materials had the same C_0 value of 6.72 A. The lower the $\beta_{\frac{1}{2}}$ or average misorientation angle, the higher the degree of preferred orientation and similarly the less the wrinkling along the basal planes. (3) The flow stresses for the 3000°C annealed and compression annealed materials are not markedly different. Why these flow stresses are quite similar, despite the large difference in preferred orientation, is not clearly understood. Future studies will investigate this point. (4) Attempts to measure the flow stress at lower temperatures generally resulted in specimen fracture before a constant value was reached. (5) The decrease in flow stress with increasing test temperature implies that the Davidson and Losty suggestion (Ref. 1) of an increase in basal flow strength with increasing temperature is incorrect.

The effect of strain rate on the basal flow stress for pyrolytic carbon after various treatments is shown in Fig. 2. All curves show an increase in flow stress with increase in strain rate, with the change for the 3000°C annealed material being less sensitive to strain rate at 2900 than at 2700°C. The 3000°C and the compression annealed materials at a test temperature of 2700°C, as noted before, are similar. The greater sensitivity of the 2760°C material to strain rate suggests that grain boundary processes are the major controlling factors in its temperature dependence. Moreover, the temperature and strain rate dependence implies that basal slip in pyrolytic carbon is a thermally activated process.

As a preliminary study on the measurements of basal shear fracture strength, several tests were made on as-deposited pyrolytic carbon over the range of room temperature to 2500°C. Strength values from 700 to 1300 psi were measured at room temperature. These values are in fair agreement with published shear strength data. The shear strength increased slightly with increase in test temperature.

A room temperature shear strength value of 330 psi was measured on a compression annealed pyrolytic carbon sample ($\beta_{\frac{1}{2}} = 1.5$). Other investigators (Ref. 2) have reported values at the proportional limit as low as 129 psi for compression annealed pyrolytic carbon having a preferred orientation index, $\beta_{\frac{1}{2}}$, of 0.60. L. M. Gillin, University of Cambridge, England, has found that the room temperature basal shear fracture strength in good, natural, single crystals of graphite is only 71 psi. This work on graphite single crystals was repeated by investigators at Union Carbide (Ref. 3) who obtained values of approximately 64 psi. The single crystal values are two to five times lower than the compression annealed material,

and this fact stresses the point that although a high degree of preferred orientation may be obtained in pyrolytic carbon, the compression annealed material is far from being perfect. Further studies into the nature of lattice imperfections governing basal flow will contribute to the understanding of basal slip in pyrolytic carbon.

C. Modification of Hemispherical Emissometer for Low Temperature and Automatic Data Acquisition

W. M. Hall

An apparatus for the measuring of hemispherical emittance (described in Ref. 4) uses temperature change measurements on a plane specimen facing a flat black plate which is cooled by liquid nitrogen. Since the radiation absorptance of this heat sink depends solely on the radiation absorptance of a black paint applied to a flat surface, a problem exists regarding the possibility of change of optical properties of the paint at its operating temperature (the nitrogen boiling temperature) from those at room temperature. A blackened hollow cavity-type heat sink offers the possibility of several highly absorbing reflections of the radiation from the specimen and thus would present a very black surrounding which is largely independent of changes in the paint used for blackening.

The emittance (and consequently the absorptance) of several rotationally symmetric cavity geometries has been calculated analytically and presented (Ref. 5). A conical 30-deg cavity was selected as the best and easiest to construct for this purpose.

A circular specimen geometry was adopted to permit the specimen and its radiation guard to be situated well inside the conical cavity and thus allow the specimen to be cooled well below room temperature for investigation of variation of emittance at low specimen temperatures. Measurements were made to −190°F.

In addition, a Dymec 2010A automatic data acquisition system was installed to replace manual recording of emissometer system voltages. The emissometer system requires a number of quite accurately spaced readings over a span of approximately 1 hr and the Dymec data system measures and records these data on printed tape, requiring only occasional monitoring during an experimental run.

Results of hemispherical emittance (ε_H) measurements of general interest are shown in the accompanying table. The numbers are averaged from several measurements on each material.

Material	ε_H
As-received aluminum	0.066
Polished gold plate on aluminum	0.044
ARF-2 white paint (ZnO-potassium silicate)	0.935
Laminar X-500 white paint (TiO_2 polyurethane)	0.87
3-M optical white paint	0.89
PV-100 white paint	0.82

Measurements of emittance at the lower temperature range showed a distinct decrease in the emittance value. This variation is shown in the point plot of Fig. 3 for ARF-2 white paint and a flat black paint from a commercial pressurized spray can. Similar variation of emittance of paints and dielectrics has been observed in other laboratories: U.S. Naval Ordnance Laboratory, Corona, California; the U.S. Naval Radiological Defense Laboratory, San Francisco, California; and Ref. 6.

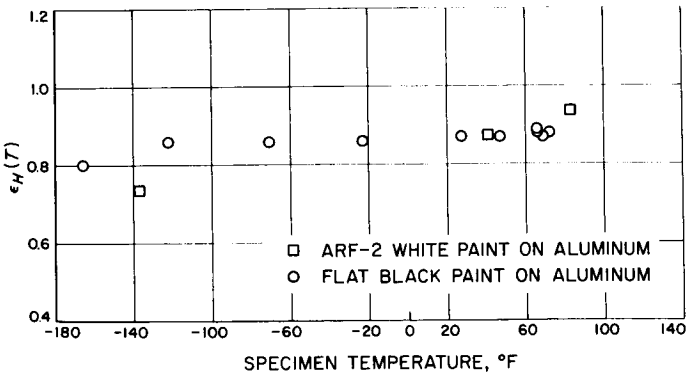


Fig. 3. Variation of hemispherical emittance with temperature of ARF-2 (Zinc oxide-potassium silicate) white paint and a flat black paint

References

1. Davidson, H. W. and Losty, H. H. W., "An Interpretation of the Mechanical Behavior of Carbons," *Proceedings of Fourth Carbon Conference*, pp. 585–591, Pergamon Press, New York, London, 1960.

2. Meers, J. T., "Effects of Neutron Damage on the Mechanical Properties of Pyrolytic and Single-Crystal Graphite," *Quarterly Progress Report No. 2*, Union Carbide Corp., Carbon Products Division, Research Laboratory, New York, July 1 to September 30, 1964.

3. Soule, D. E. and Nezbeda, C. W., "Shear Strength in Single Crystal and Pyrolytic Graphite," Paper presented at the Seventh Carbon Conference, Case Institute of Technology, Cleveland, Ohio, June 1965.

4. Hall, W. M., "The Application of Temperature Rate Measurements to the Determination of Thermal Emittance," Technical Report No. 32-596, Jet Propulsion Laboratory, Pasadena, California, April 1964.

5. Sparrow, E. M., "Radiant Emission, Absorption and Transmission Characteristics of Cavities and Passages," Symposium on Thermal Radiation of Solids, (NASA SP-55), San Francisco, California, March 1964.

6. Aronson, J. R., and McLinden, H. G., "Far Infra-Red Spectra of Solids," Symposium on Thermal Radiation of Solids, San Francisco, California, March 1964.

VIII. Applied Mechanics

A. Cavity-Type Absolute Total-Radiation Radiometer

J. A. Plamondon and J. M. Kendall, Sr.

Experiments are planned in the area of radiation-conduction-coupled heat transfer to investigate the adequacy of various idealized computational models. It is believed that the absolute measurement of local surface irradiance is highly desirable in order to obtain the maximum information from these experiments. Since there is no suitable instrument for this purpose, a radiometer is being developed, which promises to provide a means for absolute measurement of total irradiance.

Fig. 1 is a sketch of a prototype radiometer currently undergoing test. Basically, the radiometer consists of a conical cavity measuring element surrounded by a thermal guard. A cavity was selected in order to achieve spectral insensitivity and a conical geometry was picked for ease of fabrication relative to other possible cavity geometries.

The cone is made of silver foil, of 5-mil thickness. The foil serves as a thermal conduction sheath for minimizing temperature gradients resulting from nonuniform radiosity and irradiance along the cone's internal surface. Wound on the external surface of the silver cone is 1.5-mil enamel-insulated, Ballast nickel wire. This nickel wire winding is referred to as the cone resistor. The high thermal coefficient of electrical resistance of nickel makes the cone a highly sensitive resistance thermometer. The

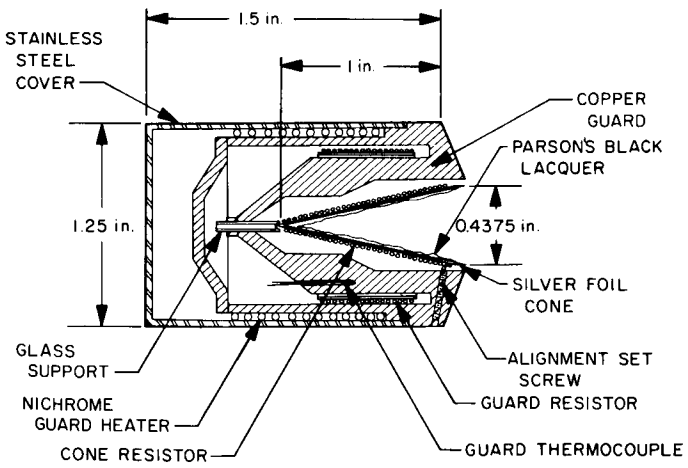


Fig. 1. Prototype radiometer sketch

cone resistor serves both as a resistance thermometer and an electrical heating element. The internal surface of the cone is coated with Parson's black lacquer. The cone is estimated to be black to better than 99.5% for all wavelengths between 0.2 and 40 μ, based on available data for the coating and on theoretical calculations in the literature (Ref. 1).

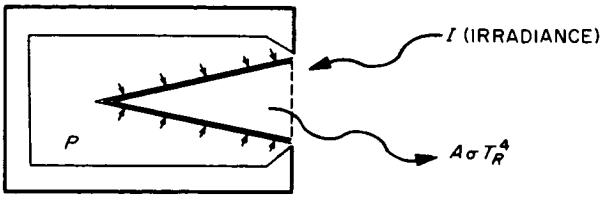
The cone is supported from its apex by a pyrex glass tube and is aligned within the guard by three glass-tipped set screws. Glass was chosen for supports because of its low thermal conductivity.

The thermal guard is machined from copper bar stock and is rather massive in order to minimize temperature gradients. It is gold plated on its interior surfaces to reduce radiant heat exchange between the guard and the cone. Located in good thermal contact with the guard is a second nickel winding made of the same materials as the cone and in an identical manner, except that the geometry is cylindrical. The second nickel winding serves as a resistance thermometer and is referred to as the guard resistor. Also, there is a nichrome wire heater winding on the external surface of the guard, and a thermocouple imbedded in the guard in the region of the guard resistor. Over the end of the guard is a gold-plated stainless steel cover.

The cone resistor and the guard resistor form two legs of a Wheatstone bridge, with the other two legs mounted externally.

The guard is maintained accurately at a preset constant temperature by electronic means, using the thermocouple and the nichrome heater. Since the guard resistor is in good thermal contact with the guard, it is at the same preset temperature. Through the use of the bridge circuit formed by the cone and guard resistors, the cone temperature is slaved to the guard temperature by using the cone resistor to heat the cone.

Fig. 2 shows an energy balance on the cone. Because the cone and guard are at the same preset temperature, no net heat exchange can occur between them. Thus, if it is a black cavity, the net radiation leaving the cone must equal the net unknown radiation entering the cone plus the electrical power required to maintain the cone at the same temperature as the guard. The net radiation leaving the cone is given by $A\sigma T_r^4$, since the cone is a black body. Thus, by knowing the cone temperature, the cone aperture area, and the electrical power supplied to the cone,



$A\sigma T_R^4 = IA+P$ WHERE: T_R = RADIOMETER TEMPERATURE
A = CONE APERTURE AREA
I = UNKNOWN IRRADIANCE
P = ELECTRICAL POWER TO CONE

Fig. 2. Heat balance on cone

the unknown radiation entering the cone can be obtained absolutely by subtracting the electrical power from $A\sigma T_r^4$. Of course, the preset temperature on the guard must be sufficiently high so that the net radiation leaving the cone is greater than that entering it.

Fig. 3 is a diagram of the control circuitry for the radiometer. Starting with the guard, the thermocouple imbedded in the guard goes to a reference junction, then to a millivolt supply. The output of the thermocouple is nulled against the preset voltage using the servo to control the amount of electrical power supplied to the nichrome guard heater. The slaving of the cone is accomplished in the following manner. DC voltage is impressed on the bridge formed by two external resistors and the cone and guard resistors. A 500-Ω potentiometer is in the bridge circuit for initial bridge adjustment. If the cone and guard are not at the same temperature, the bridge will not be in balance due to the coefficient of electrical

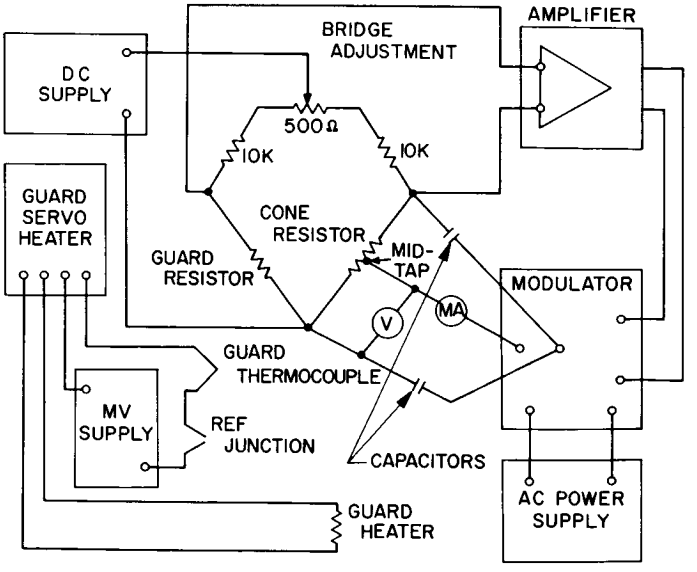


Fig. 3. Radiometer electronic circuit

resistance of the nickel windings. To balance the bridge, the DC imbalance voltage is inputed into an amplifier. The amplified signal is fed into a modulator which supplies AC power to the cone resistor through a midtap. The amplifier and modulator form a servo which nulls the bridge imbalance signal.

The accuracy of the radiometer is tested by irradiating it with a known amount of radiation and comparing this with the value as measured by the radiometer. The test setup is shown in Fig. 4. The source of known radiation is a cylindrical copper hohlraum which is heated electrically to an accurate and uniform temperature. The inside surface of the hohlraum is coated with Parson's black lacquer and the hohlraum is estimated to be black to better than 99.9%. The external surfaces of the hohlraum are wrapped in superinsulation to minimize temperature gradients in the walls of the hohlraum. To make measurements of the radiometer's accuracy, the radiometer is placed in the aperture of the hohlraum in a manner that will enable it to see only radiation originating from the surfaces of the hohlraum. Since the hohlraum is black, the known radiation entering the radiometer cone is calculated from $A\sigma T_h^4$, where A is the cone aperture area and T_h is the hohlraum temperature. The tests are performed in vacuum at a pressure of less than 5×10^{-5} torr.

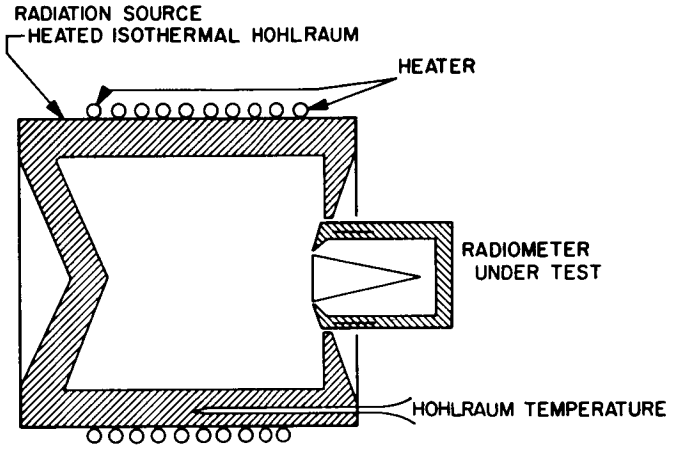


Fig. 4. Setup for radiometer test

Figs. 5 and 6 are graphs of some of the data taken with the radiometer. Fig. 5 shows a plot of the ratio of measured irradiance to the calculated hohlraum irradiance, I (meas)/I(hohl) , versus the ratio of radiometer temperature to the hohlraum temperature raised to the fourth power, T_r^4/T_h^4. The ratio, T_r^4/T_h^4, is, of course, the ratio of the radiometer radiant output to the hohlraum radiant input. The lower curve is plotted for the hohlraum temperature held constant with the radiometer temperature increased in steps. The upper curve is the radiometer temperature held constant and the hohlraum temperature increased in steps until its temperature is near radiometer temperature. Deviation of these curves from the horizontal line at I(meas)/I(hohl) equal to one is, of course, the error in measuring accuracy of the radiometer.

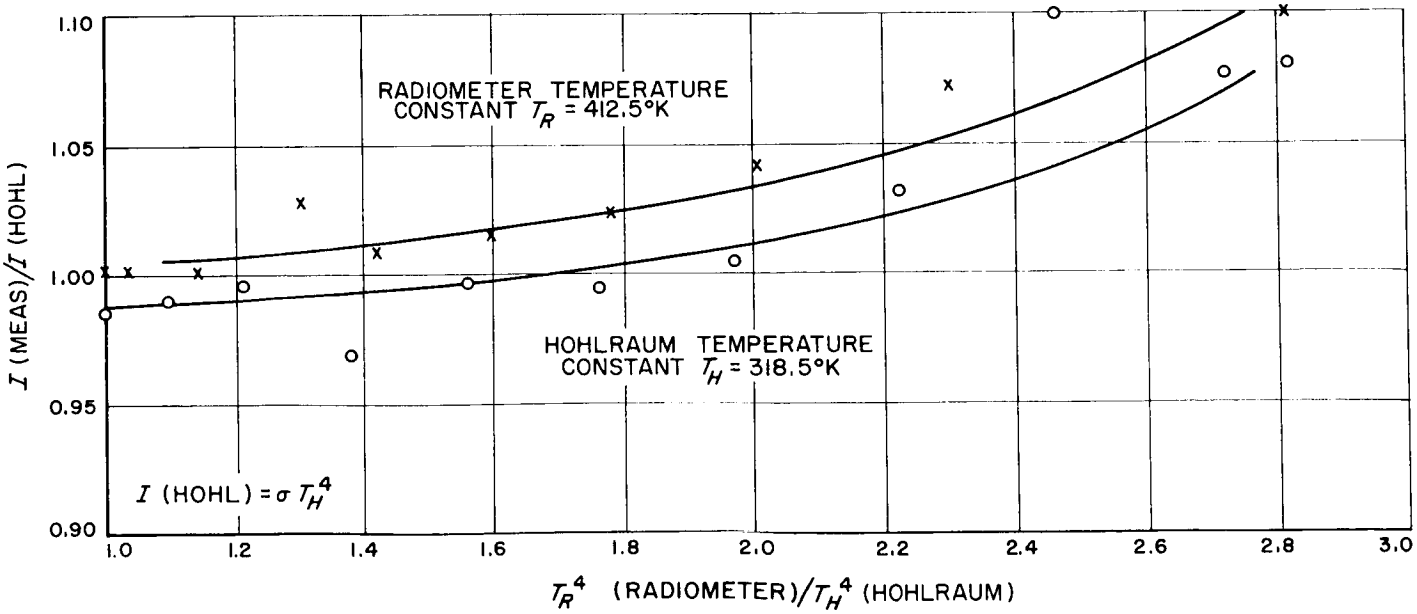


Fig. 5. Experimental data for constant radiometer and hohlraum temperatures

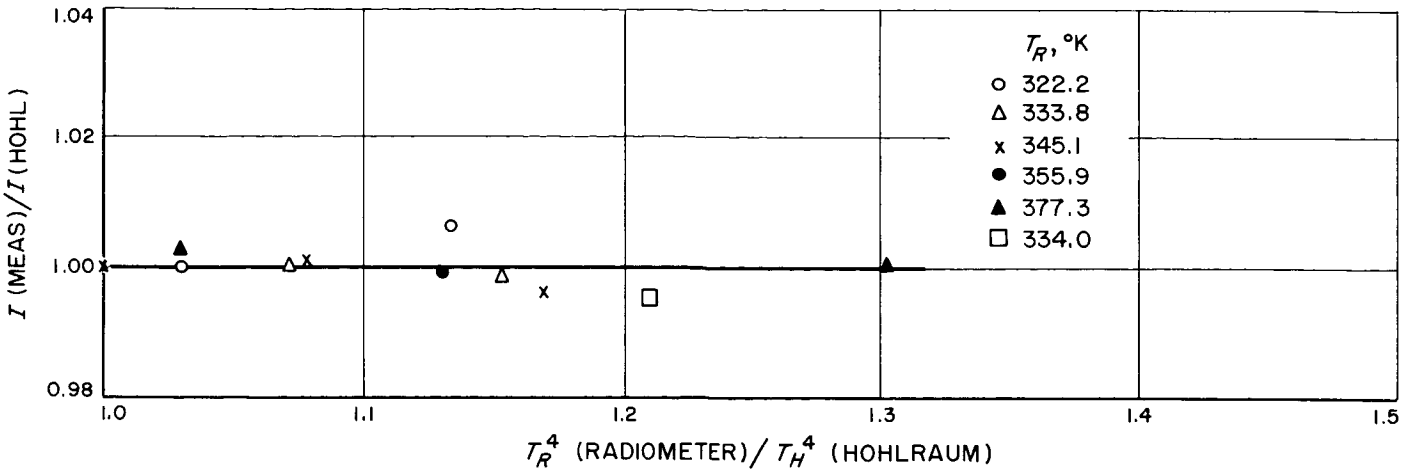


Fig. 6. Experimental data in the range of low ratio of radiometer to hohlraum temperature

As can be seen, the error in measuring accuracy increases as the ratio of radiant output to input increases. The shift in these curves results from an intentional change in the initial bridge balance setting on the 500-Ω potentiometer. For the most part, the measurement error in the radiometer is small when the radiometer temperature is near the temperature of the hohlraum; that is, the measurement error is small when the radiometer is operated near the "black" temperature of the irradiance. Recognizing this fact, additional data (plotted in Fig. 6) were taken with the radiometer operated in such a manner. In Fig. 6, the radiometer temperature was varied but the hohlraum temperature was kept near that of the radiometer. As can be seen, the measurement error is indeed small. Maximum deviation from the horizontal line at unity is 0.6%. The operation of the radiometer near the "black" temperature of the irradiance would be the normal mode of operation. On the basis of these curves, measurement accuracies of less than 1% seem possible.

B. A Study of Convective Heat Transfer Within a Mars Lander

M. Suo[1]

1. Introduction

The surface of Mars presents a unique environment for temperature control. The atmosphere has a pressure of the order of 1% of that on Earth and has a gravitational field about one-third of that on Earth. This makes it quite different from Earth, but it is also quite different from space, where there is no atmosphere and no gravitational field.

Before an attempt is made to design for temperature control, it will be advantageous to know how important convective heat transfer is with respect to radiative heat transfer. The study reported here was made to provide some of that information; it does not cover all convective heat transfer on a Mars lander, but only that used within the lander for distributing heat from a source to a radiator.

Two problems were investigated. The first was that of heat transfer from a vertical flat plate sitting within the lander. The second was the heat transfer within an enclosed loop of fluid. Both natural and forced convection were examined.

2. Heat Transfer From a Vertical Flat Plate

The heat transfer from a 4-ft-high vertical flat plate to the atmosphere and lander wall has been calculated for the geometry shown on Fig. 7.

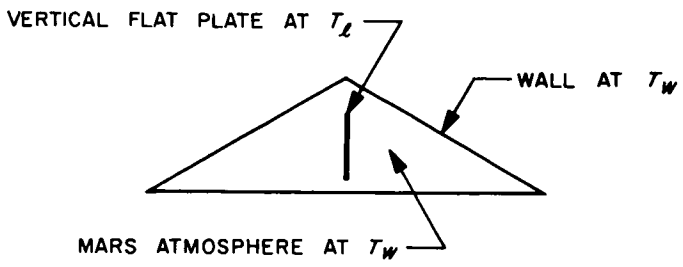


Fig. 7. Mars lander with vertical flat plate

[1]MIT Mechanical Engineering Department, Cambridge, Massachusetts.

Heat is transferred from the plate to the atmosphere by convection, and from the plate to the wall by radiation. The atmosphere does absorb a small amount of radiation, but that is ignored. The parameter which has been calculated is the ratio of convective heat transfer to the radiative heat transfer. This has been done for both natural and forced convection using the atmospheric models of Ref. 2. The parameters which affect the results are:

(1) Temperature of the plate (T_ℓ = 530°R is the only case studied)

(2) Temperature of the wall and atmosphere, T_w

(3) Model of atmosphere

(4) Total exchange shape factor ($\mathcal{F}$ of Hottel)

(5) Velocity (in forced convection)

The results of the calculations can be summarized as follows:

(1) As shown in Fig. 8, natural convective and radiative heat transfer rates are of the same order for T_w = 430 and 330°R over a wide range of values of $\mathcal{F}$ and for all three atmospheric models.

(2) As shown in Fig. 9, forced convection gives higher convective heat transfer rates and under some conditions (low $\mathcal{F}$ and high velocity) can be the controlling mechanism of heat transfer.

(3) Various atmospheric models will give different values of convective heat transfer, but within the pressure range of 10 to 20 mb and considering only CO_2 and N_2 as the constituents, the possible errors in convective heat transfer should not be greater than 75%.

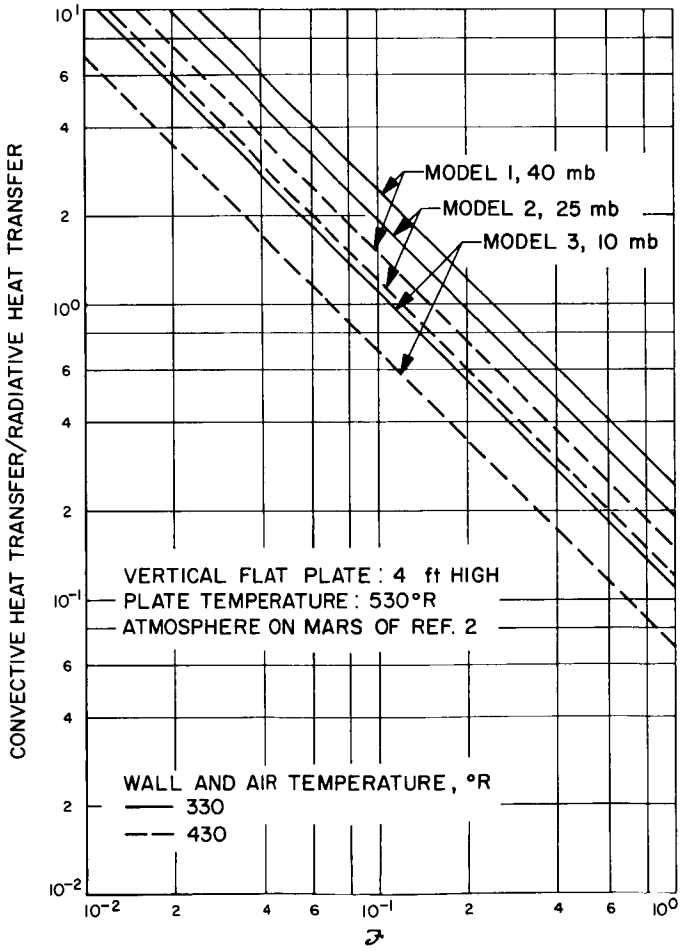


Fig. 8. Ratio of natural convective-to-radiative heat transfer

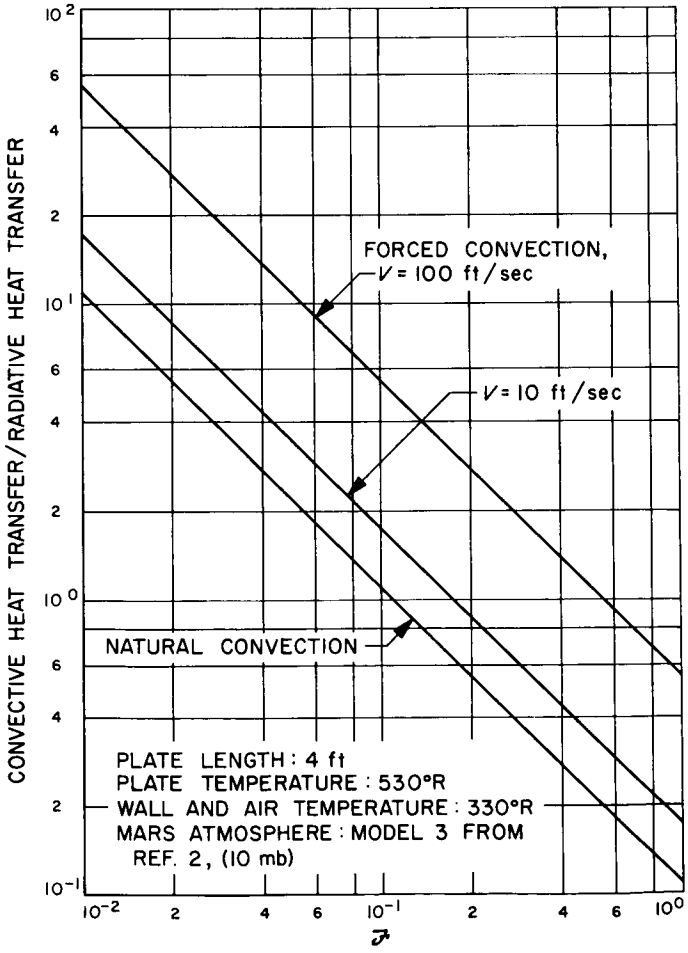


Fig. 9. Effect of forced convection on ratio of convective-to-radiative heat transfer on vertical flat plate

3. Heat Transfer in an Enclosed Loop

Another way to use convective heat transfer is to make an enclosed fluid loop as shown in Fig. 10. This is a natural circulation loop. The fluid, heated in the source heat exchanger and cooled in the radiator, is driven by the difference in density between the fluid in the down corner and the riser. The operation of a loop with the following parameters was analyzed:

Source and radiator: 1-in. diameter tubes, 2-ft long

Riser and downcomer: 1-in. diameter, variable length

Source temperature: 530°R

Radiator temperature: 430°R

Working fluid: CCl_2F_2 (Freon 12)

The analysis shows that as the riser length increases, the heat transfer for these temperatures reaches a limit of 686 Btu/hr. Furthermore, it reaches 90% of the maximum with a riser length of approximately 5 ft. This maximum limit is not very high, but it compares favorably with radiation. To transfer the same amount of heat with radiative transfer alone, between the same two temperatures, would require a total exchange area $A\mathcal{F}$, between two heat exchangers of 9 ft.[2]

If the system were to be operated in forced convection by inserting a pump, the amount of heat transfer would be essentially limitless. The heat transfer increases with the fluid circulation rate and this is only limited by the amount of available pumping power.

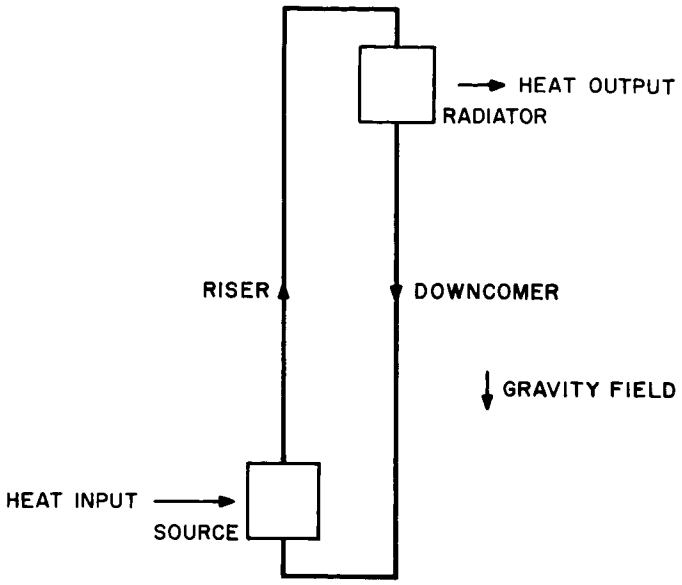


Fig. 10. Natural circulation loop

4. Conclusions

(1) Natural convection can transfer a significant amount of heat either around a flat plate or in an enclosed loop. This is important both as a means of control and as an inherent problem in the systems on Mars.

(2) Forced circulation can transfer a much larger amount of heat. It is limited primarily by the amount of pumping power available.

References

1. Sparrow, E. M., and Johnson, V. K., "Radiant Emission Characteristics of Diffuse Conical Cavities," *Journal of the Optical Society of America*, Vol. 53, pp. 816–821, 1963.

2. Levin, G. M., Evans, D. E., and Stevens, V., "NASA Engineering Models of the Mars Atmosphere for Entry Vehicle Design," NASA TN D2525, Washington, D. C., November 1964.

IX. Lunar Spacecraft Development

A. High Impact Technology

J. L. Adams

JPL has a continuing program to develop the technology necessary for building spacecraft-type equipment capable of surviving impacts on the order of 10,000 *g* from velocities of several hundred feet per second. Impact-resistant equipment simplifies the trade off between the retardation system, energy absorber, and payload impact environment; allows the building of rough landers and diagnostic packages for soft landers; and results in an increased ability to survive abnormal touchdowns due to retardation system anomalies.

This report summarizes accomplishments in the JPL high impact program during the reporting period.

1. Ruggedized Gas Chromatograph

A series of transistors for use in the gas chromatograph was screened by being impacted laterally (normal to the leads) at 10,000 g from an impact speed of 180 ft/sec. The transistors tested were:

Transistor No.	Manufacturer	Total number tested
2N2979	Amelco	28
2N2605	Texas Instruments	17
2N3350	Texas Instruments	8
2N930	Texas Instruments	10
2N930	Fairchild	21
2N2843	Solid State Products	6

One Fairchild 2N930 and one Amelco 2N2979 failed. Both failures were at the bond between the transistor chip and the lead wire.

Fig. 1 shows the gas chromatograph pressure regulator which regulates the 3000-psi carrier gas supply to 65 psi.

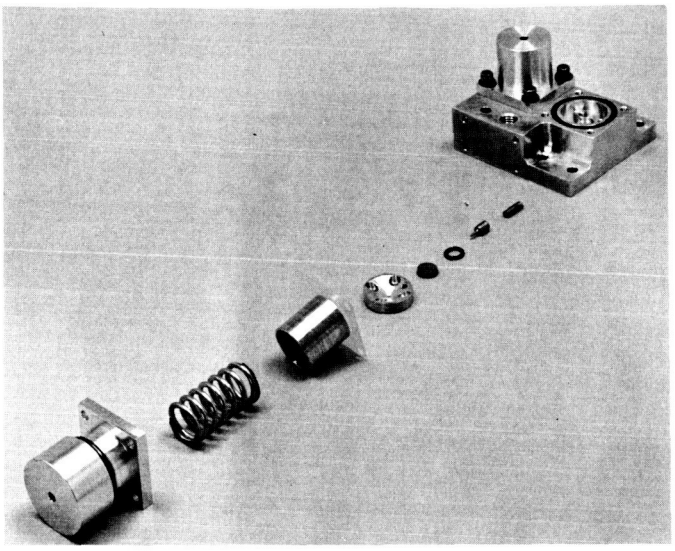


Fig. 1. High impact gas chromatograph pressure regulator

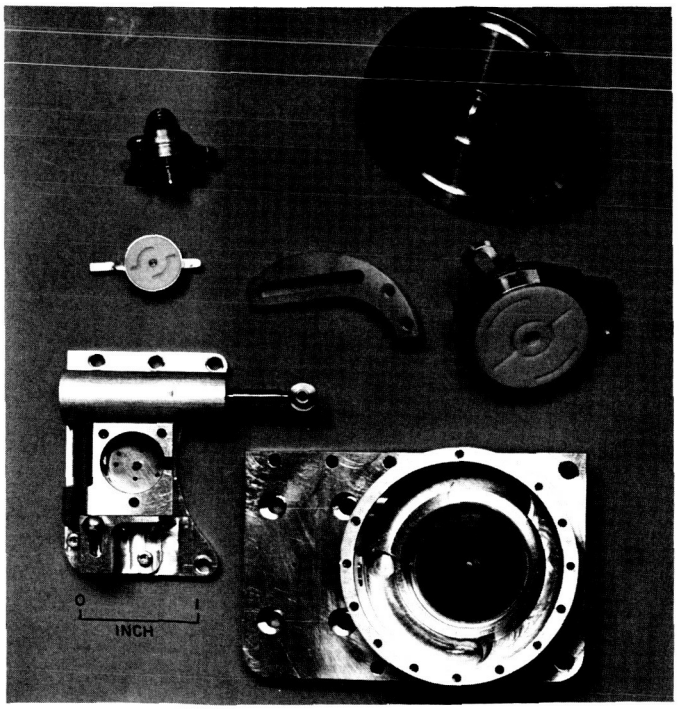


Fig. 2. High impact gas chromatograph sample valve

The regulator is a two-stage device with an adjustment in the low pressure stage, and weighs 5 oz. Fig. 2 is a view of the gas chromatograph sample valve. This device is a three-position rotary gas valve used to interconnect the sample source, carrier gas, and vacuum line from the jet pump, sample loop, and the analysis line in various combinations. The valve is powered by regulated carrier gas which in turn is controlled by a solenoid actuated valve. The sample valve weighs 8 oz; it has not yet been impact-tested, but should survive since it contains many features tested on a previous sample valve model.

2. *Ruggedized S-Band Transmitter*

Three crystal units produced under contract with Valpey-Fisher Corp., Holliston, Massachusetts, were impacted. The units did not meet the performance specifications before impact. However, they were typical enough to be representative of crystal behavior under impact. Two of the crystals were mounted by clamping the mounting ring under the heads of three screws which were threaded into the carriage of the horizontal testing machine. The crystals were impacted parallel to the mounting ring and in both axial directions. They survived tests up to and including 10,000 *g* from 200 ft sec in all directions, with the exception of the direction which tended to pull the mounting ring from the crystal. In this direction, one mounting ring peeled loose due to the bending which was allowed by the mounting technique. This crystal broke upon hitting the carriage. The other crystal did not leave the fixture, even though the mounting ring was bent and partially torn loose. The second crystal suffered a frequency shift of approximately 300 cps (nominal frequency, 23.1 Mc).

The third crystal was mounted to the fixture by means of a continuous clamping ring. This unit survived 10,000 *g* from 200 ft/sec in all directions. The frequency shift during testing was approximately 200 cps. The second module (power transmitter) of the S-band transmitter was impacted at various levels up to 9,000 *g* from 175 ft/sec in all principal directions without failures occurring. It survived a 12,000-*g* impact from 175 ft/sec in one direction, but failed a similar impact parallel to the longitudinal axis. The failure was found to be due to a loose nut on a transistor. After repairing the unit, it survived impacts of more than 10,000 *g* from 175 ft/sec in all principal directions.

Fig. 3 is a ruggedized turnstile S-band antenna which has survived impacts of 9,000 *g* from 170 ft/sec with no significant mechanical or electrical damage. The antenna weighs 3½ oz. A Sanders Associates Model AS-63 spiral antenna has successfully survived impacts up to 5,000 *g* from 100 ft/sec in all principal directions. This antenna was not designed for impact use. Failure occurred because of deflections in the spiral element, which

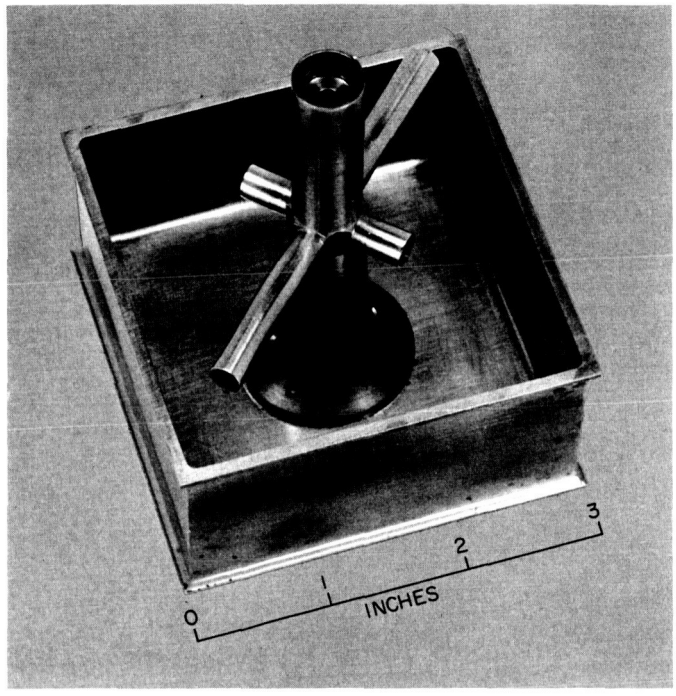


Fig. 3. High impact S-band antenna

is 1/16-in. thick epoxy glass laminate. By properly supporting this element, the antenna should be capable of surviving 10,000 *g*.

A Hughes Model 216 H TWT was impacted at various *g* levels. This tube was not ruggedized and was impacted solely to determine the problems involved with ruggedization. Impacts ranged up to 1000 *g*. During all impacts with the acceleration along the tube and away from the gun, the tube performance was degraded. The tube failed during the 1,000-*g* impact in this direction. In tests previous to the 1,000-*g* impact, the degradation could be removed by an impact in the opposite direction.

3. *Batteries*

Table 1 summarizes various tests performed on battery cells during the reporting period. In all cases the cells were completely supported in a rigid (typically 3/8-in. wall aluminum) fixture. The impact directions referred to are shown in Fig. 4; arrows represent the direction of accelerations (opposite to the direction of preimpact velocity). The plates are parallel to the *BA* plane. In general, the direction *A* is most severe and *D*, *B*, and *C* follow in order of decreasing severity. The Whittaker Power Sources Division CD-3 silver cadmium cells and the Yardney HR-1 silver zinc cells have withstood impacts of approximately 10,000 *g* in all directions with no short-term electrical degradation. The cell cases cracked and the plates were deformed, but charge–discharge behavior did not reflect the mechanical damage. The Gould nickel-cadmium button cells survived impacts in all principal directions with no detectable damage. However, voltages were intermittent during certain impacts. This was because the cell plates were held in contact with the case by means of a spring contact. When the cells were impacted so that the acceleration was toward the spring, the plate mass collapsed the spring and broke the electrical contact.

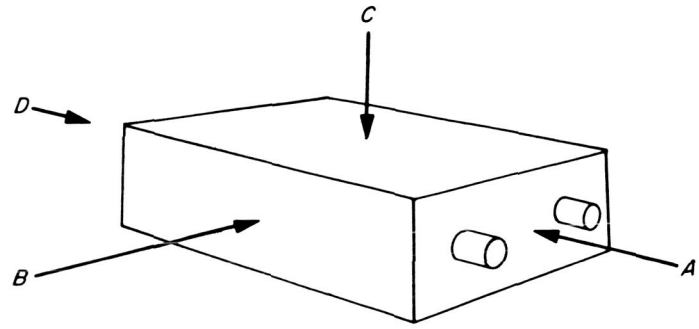


Fig. 4. Battery impact acceleration directions

4. *Mechanisms*

Fig. 5 shows an air turbine, constructed to test the feasibility of operating a rotating device during an impact. This turbine consists of an 8-oz rotor supported between spring-centered bearings. The spring-mount spring constants were approximately 25,000 lb/in. axially

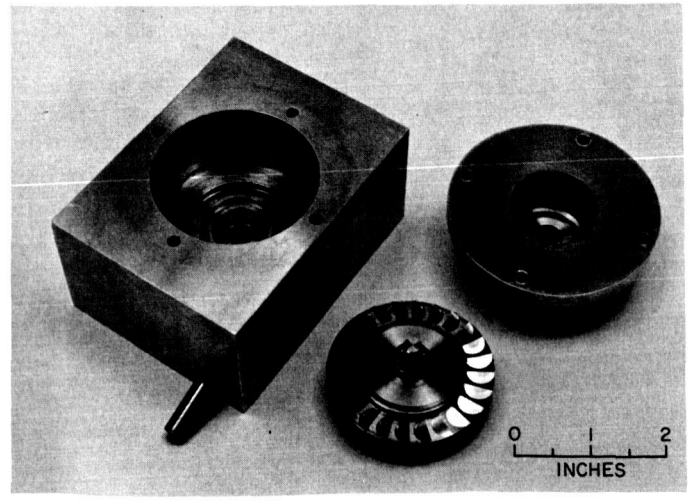


Fig. 5. High impact air turbine

Table 1. Battery cell test summary

Battery	Condition	Potting	Direction	g level	ΔV, ft/sec	Comments
Whittaker CD-3 sealed silver cad cells, 2 potted in fixture	Charged	Epon 815	A	10,000	164	No electrical degradation; potting and one case cracked
Ranger ESB cell	Charged	Epon 815	D	6000	165	Case cracked; internal shorting
	Charged	Epon 815	A	6800	170	Potting and case cracked; plates shifted
Whittaker CD-3 sealed silver cad cells, 2 potted in fixture	Charged	Epon 815	A	9800	176	Potting and case cracked; no internal damage
Plates in bag, potted in structure	Charged	Epon 815	Edgewise	11,600	135	Battery was destroyed
ESB 6-cell battery	Charged	Epon 815	C	5000	107	Survived
	↓	↓	D	5000	107	↓
			B	5000	108	
			C	8000	115	
			D	7200	117	Failed
Yardney HR-1, 2 cells potted in fixture	Charged	Solithane 8	A	8000	122	Survived
	↓	↓	A	6300	122	↓
			C	6920	148	
			B	6910	148	Case cracked; vacuum potting sucked out some KOH
			B	9000	166	
			C	9000	179	
			D	10,800	167	
			A	12,000	171	
Gould nickel-cadmium button cells − 12 v + 6 v + M N + 6 v − + O P	Charged	828/D	M	5800	106	
	↓	↓	M	5900	109	↓
			P	5900	109	
			N	5600	108	Intermittents when impacted axially with acceleration in the positive (+) direction
			O	5600	109	
			M	8000	165	
			N	8000	168	
			O	8000	167	
			P	8500	170	
Impacted under load: 12 v, 500 BH, 40 Ω 6 v, 100 B, 75 Ω						

and radially, and the bearings were lightweight (Microtech Mc814SSR 25L10) ball bearings with welded retainers. The turbine was impacted at 10,000 *g* from 170 ft/sec in the axial and radial directions while operating at approximately 30,000 rpm. Air was continuously supplied to the turbine before, during, and after impact through flexible tubing. Although the turbine speed varied slightly during impact, the device regained its preimpact rotating speed and operated successfully for several hours before it was shut down. Examination showed that negligible bearing damage had taken place (the major loads, of course, were absorbed by bottoming between the rotating mass and the housing). Damage to the turbine and the housing was small.

Fig. 6 shows a *Ranger* Block I titanium pressure tank which has been modified and tested under high impact. A 0.200-in. thick by 0.300-in. wide titanium ring was welded to the circumference of the tank, which was then clamped in a fixture by this ring and impacted both normal and parallel to the ring. The tank was tested empty and unpressurized at *g* levels up to and including 10,000 from 150 ft/sec with no apparent damage. It was then pressurized to 3.000 psi with water and impacted in both

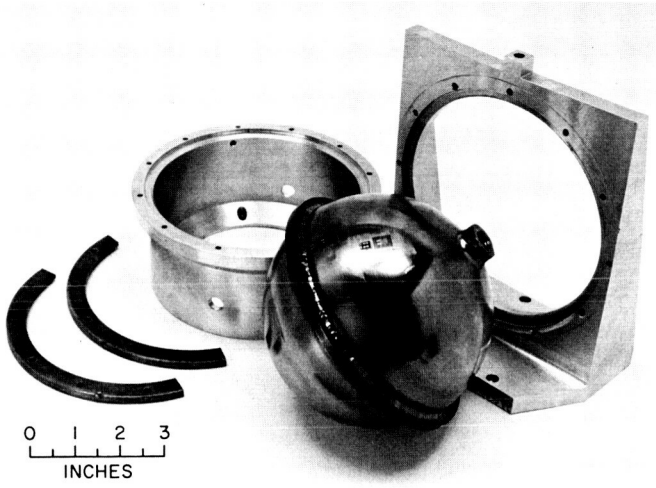


Fig. 6. Pressure vessel and testing fixtures

directions at 10,000 g from 140 ft/sec with no apparent damage.

A strain gage accelerometer (Fig. 7) was developed and tested in conjunction with an effort to gain strain gage data on structural elements during impact. This device consists of a thin-walled tube with a lead weight in the upper end and a mounting flange at the bottom.

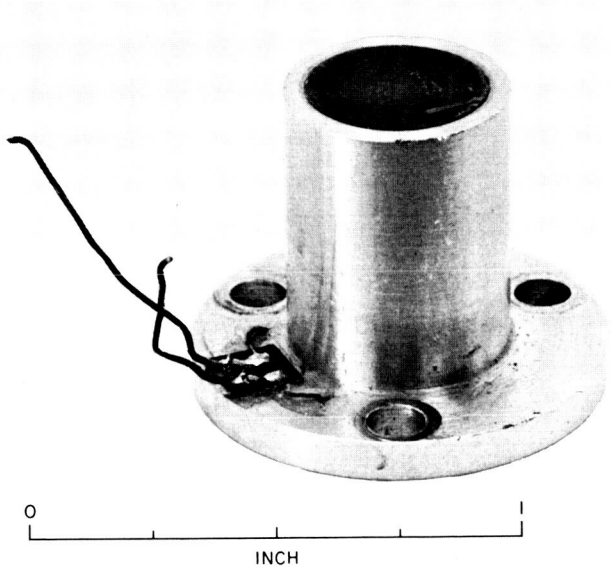


Fig. 7. Strain gage accelerometer

Strain gages are placed on the inside of the tube below the weight so that they measure compressive strain due to axial inertial forces. Acceleration versus time curves obtained from this device compare favorably with those obtained from the standard piezoelectric accelerometers.

ENGINEERING FACILITIES DIVISION

X. Aerodynamic Facilities

A. Wind Tunnels

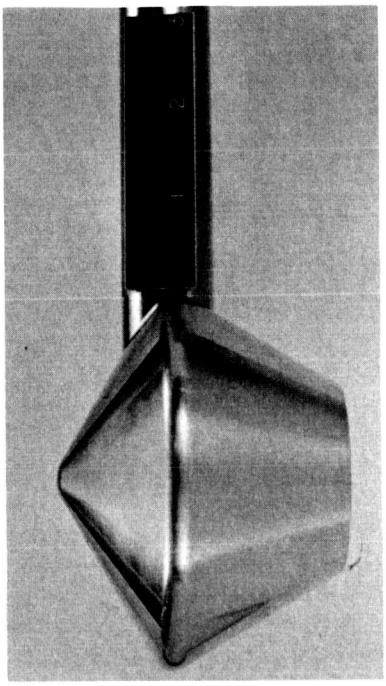
W. Marko, G. M. Goranson, B. Walker, and J. Welton

1. *Entry Capsule Afterbody Configuration Study,* W. Marko

Wind Tunnel Test 21-191 was conducted in order to obtain the static aerodynamic characteristics of possible afterbody configurations which would assure capsule orientation with heat shield forward during atmospheric entry. Normal and axial forces and pitching moments were measured with a six-component internal strain gage balance. Centers of pressure on the body longitudinal axis of symmetry were calculated, and afterbody surface static pressures were also recorded. Modified Newtonian aerodynamic coefficients were generated for comparison with test data.

All tests were conducted in the JPL 21-in. Hypersonic Wind Tunnel at a nominal Mach number of 6.0 and Reynolds numbers/in. of 0.06×10^6, 0.15×10^6 and 0.26×10^6. Test data were recorded for the body axes angles of attack, 0 to 180 deg.

The M-100 model configuration, shown in Fig. 1, was chosen on the basis of Newtonian aerodynamic estimates and past experimental investigations. All other configurations tested are afterbody perturbations of the M-100. The blunted 60-deg cone forebody was not changed. To perform a parametric study, model shapes were selected, varying one M-100 afterbody variable at a time. Model afterbody length was held constant while changing half-cone angle (Fig. 2), and the amount of afterbody truncation was varied while retaining the afterbody half-cone angle of 30 deg (Fig. 3). Additional combinations of afterbody half-cone angle and length (M-111 and M-118) were selected on the basis of Newtonian aerodynamic estimates. Special interest afterbody configurations illustrated in Figs. 4–6 are: afterbody cavity (M-114, -119 and -120); truncated cone,cone cap (M-112); truncated cone, hemisphere cap (M-113); double truncated cone (M-121); and maximum diameter corner edge radius models (M-115 and



α = 90 deg
SGB6-2 BALANCE USED FOR
ALL MODELS

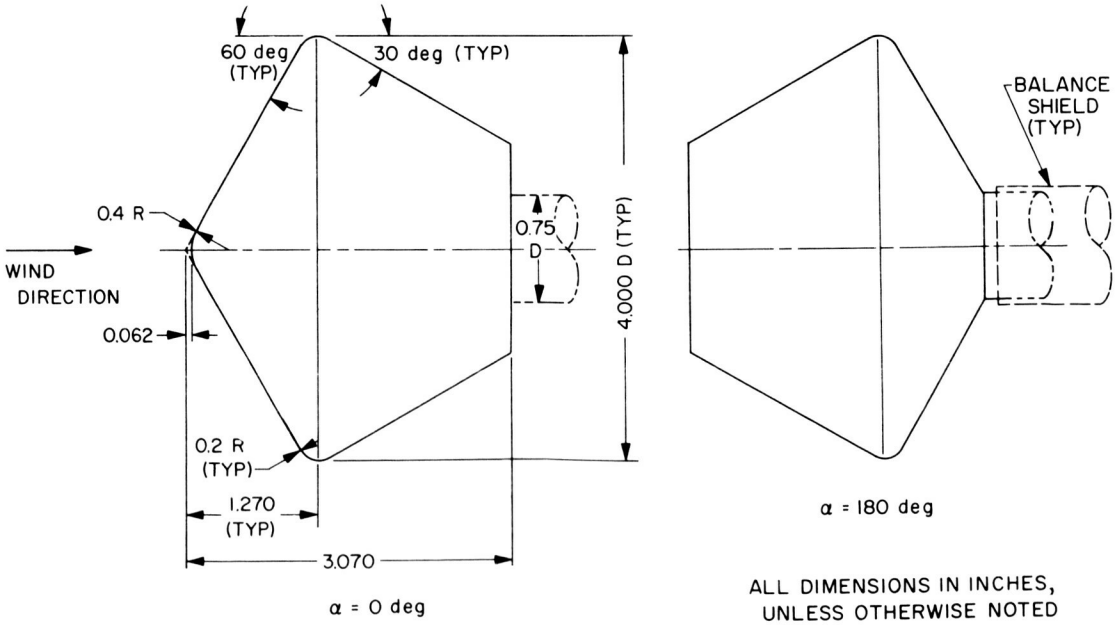
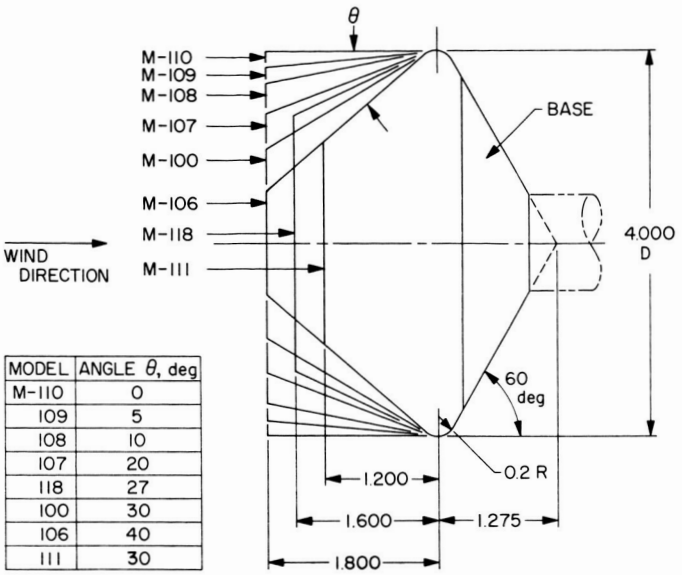


Fig. 1. Model M-100



MODEL	ANGLE θ, deg
M-110	0
109	5
108	10
107	20
118	27
100	30
106	40
111	30

ALL DIMENSIONS IN INCHES, UNLESS OTHERWISE NOTED

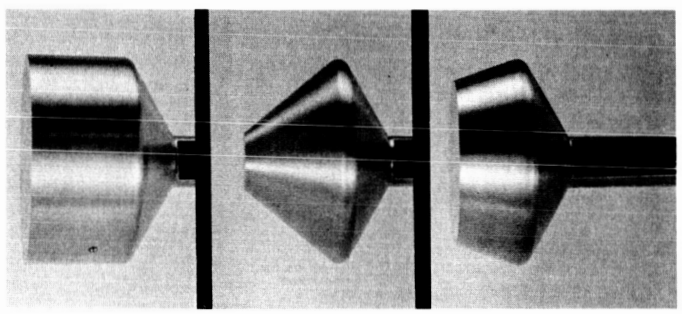
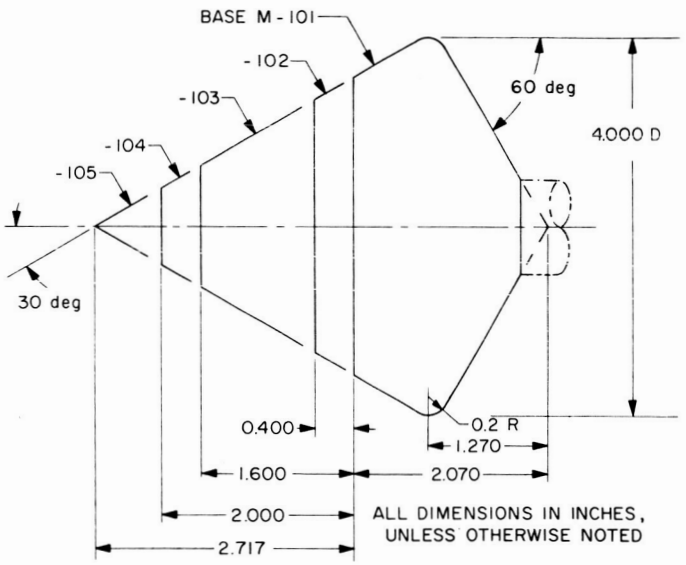


Fig. 2. Cone angle models



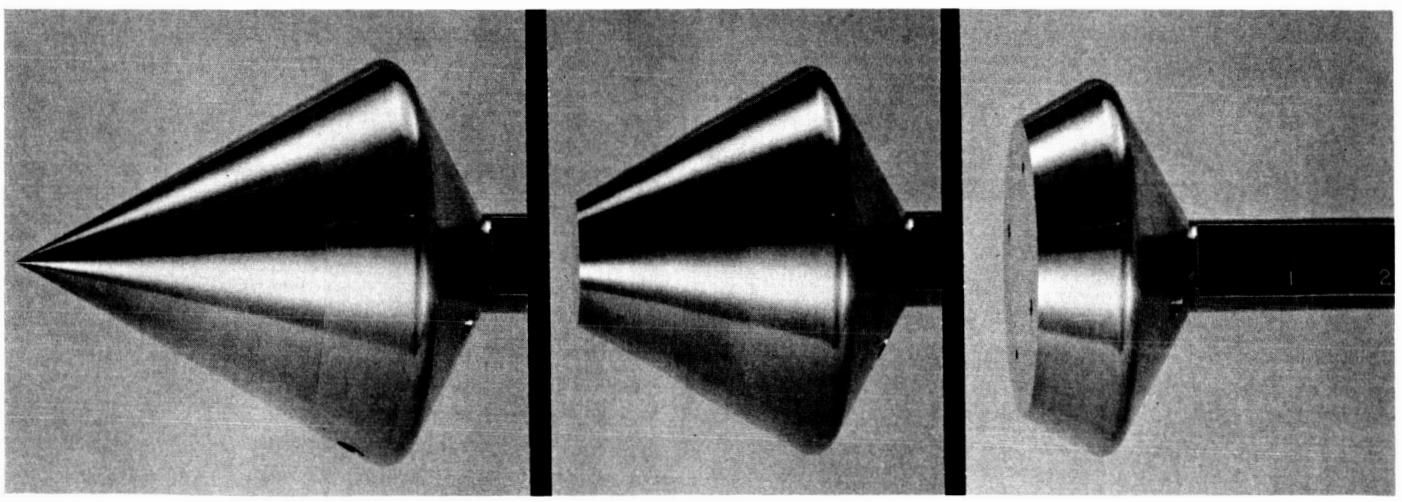
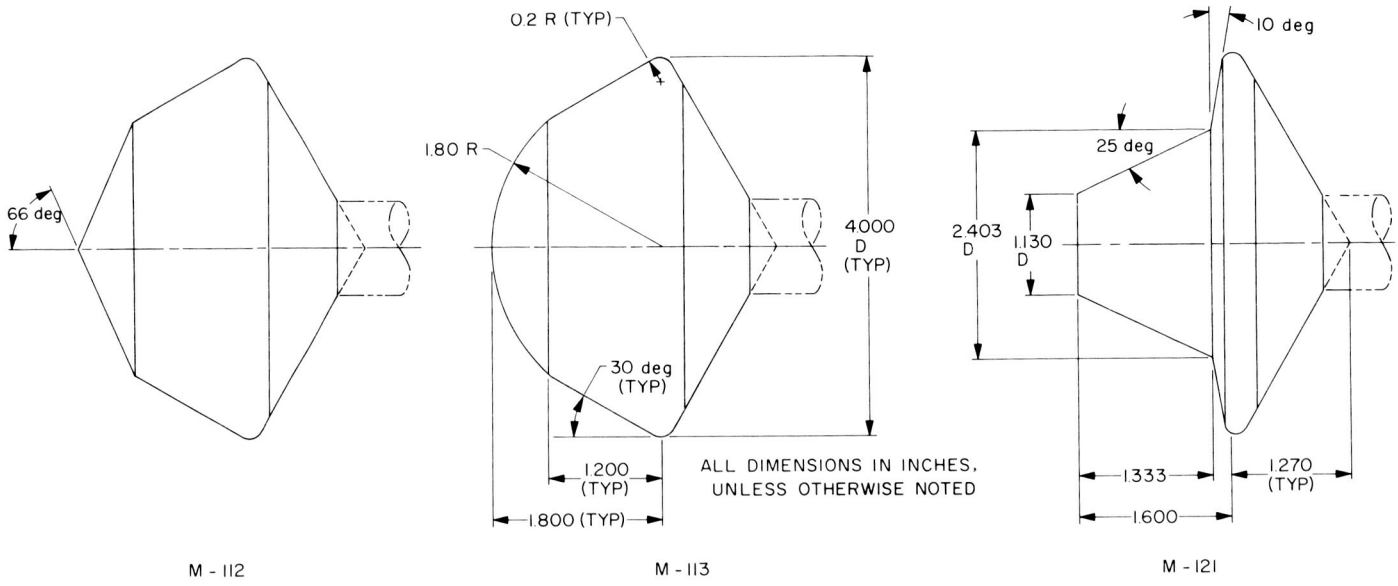


Fig. 3. Truncated cone models



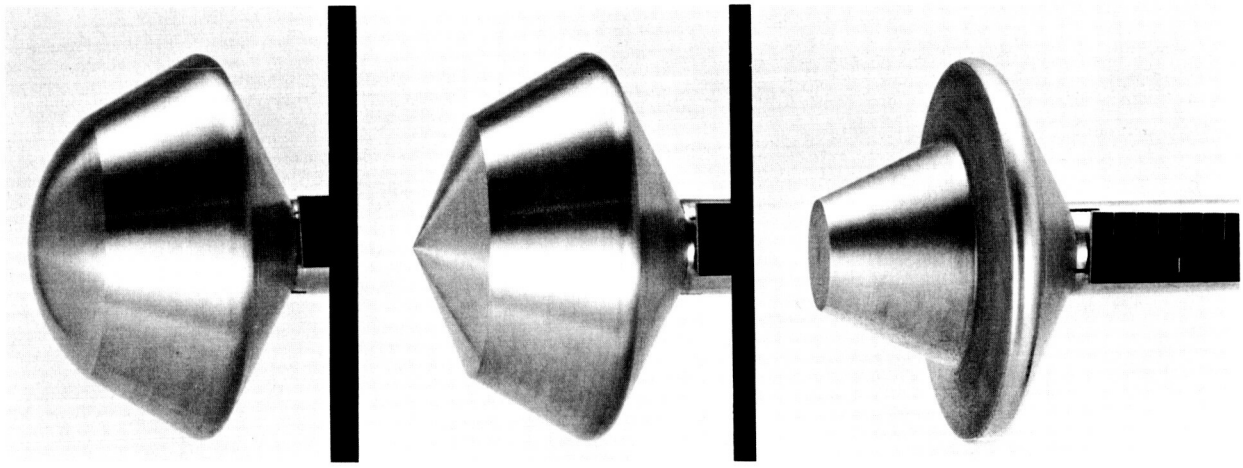
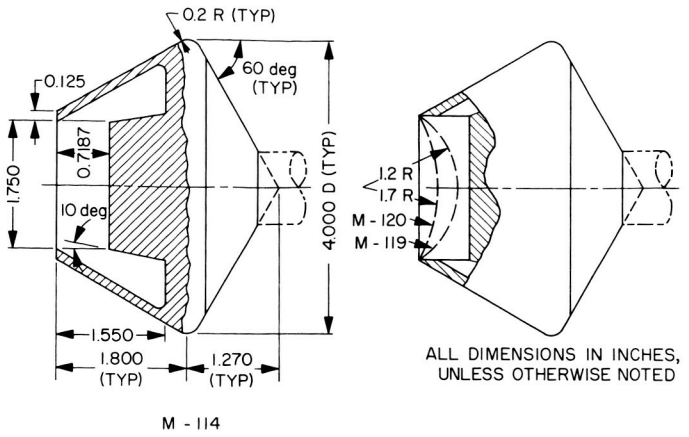


Fig. 4. Special afterbody shapes



M - 114

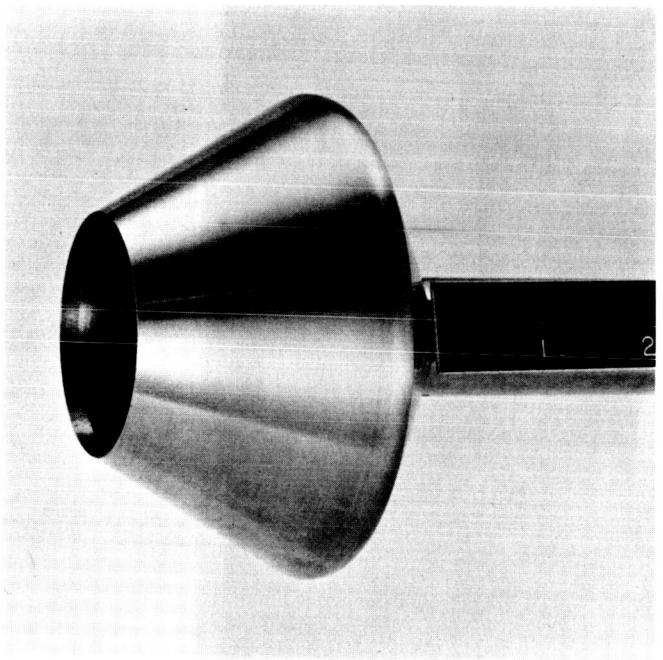
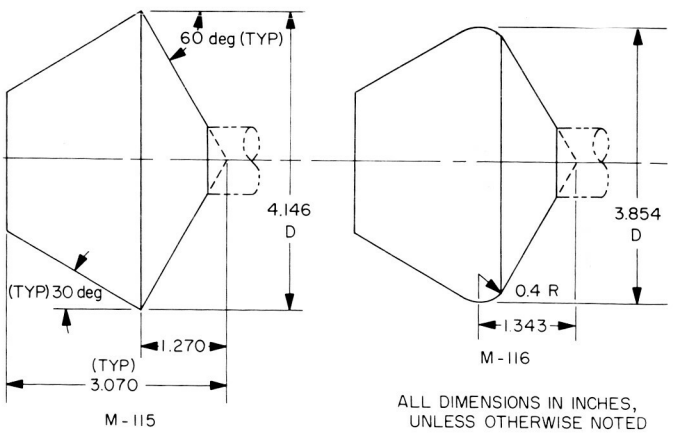


Fig. 5. Cavity models



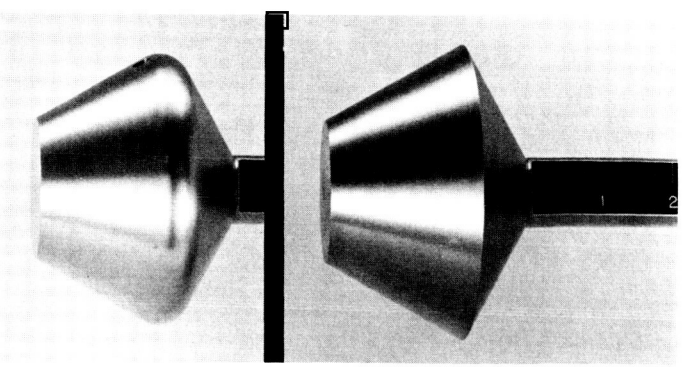
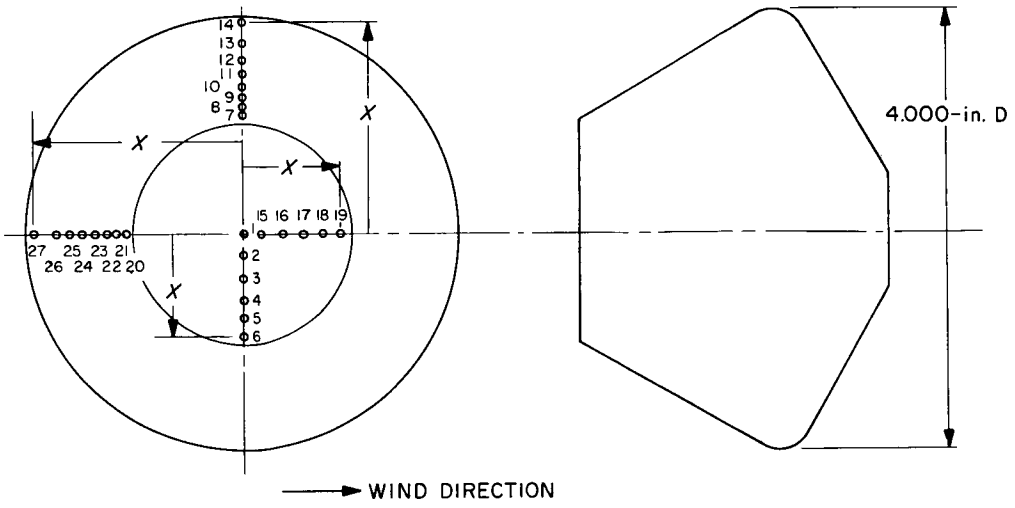


Fig. 6. Maximum diameter radii models



ORIFICE NO.	X DISTANCE FROM CENTER, in.
1	0.000
2	0.200
3	0.415
4	0.605
5	0.775
6	0.925
7	-1.041
8	-1.116
9	-1.201
10	-1.291
11	-1.404
12	-1.429
13	-1.574
14	-1.759
15	-0.200
16	-0.415
17	-0.605
18	-0.775
19	-0.925
20	1.041
21	1.116
22	1.201
23	1.291
24	1.404
25	1.429
26	1.574
27	1.759

Fig. 7. Pressure model M-117

-116). An afterbody surface static pressure model M-117, having the same dimensions as the M-100, is illustrated in Fig. 7.

All forces and moments were transferred to the nose of the 60-deg forebody (Fig. 8) and were reduced to body

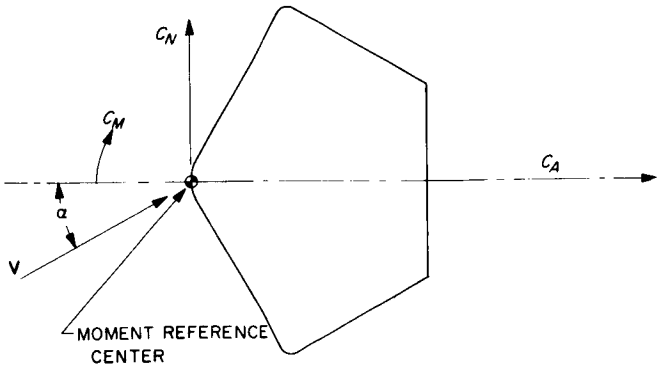


Fig. 8. Sign conventions

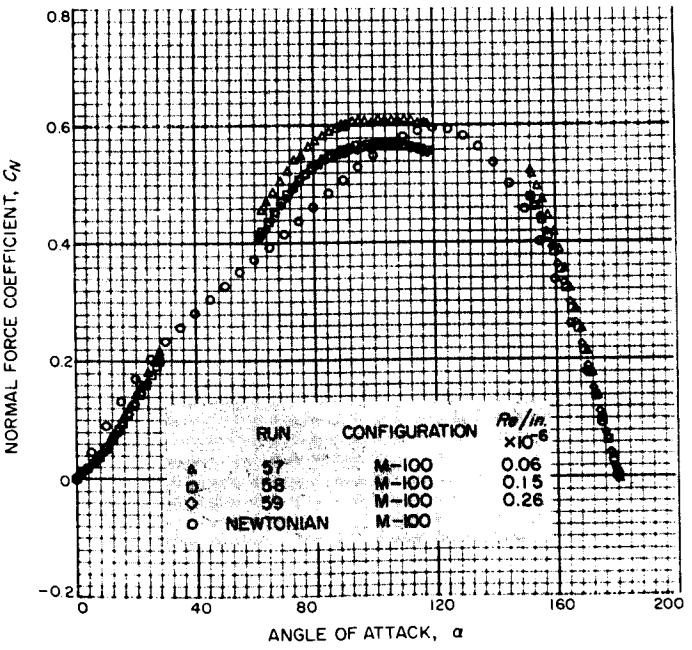


Fig. 9. Normal force coefficient versus angle of attack

axes coefficients using a reference area and length of 12.56 in.^{2} and 4.0 in., respectively.

Only a sample of the data obtained can be shown here. Body axes static aerodynamic coefficients, normal force coefficient C_N, axial force coefficient C_A, pitching moment coefficient C_M, and center of pressure $X_{cp}/D(C_M/C_N)$ for M-100 are presented in Figs. 9–12 for three Reynolds numbers. Modified Newtonian coefficients ($C_{p_{max}} = 1.83$) are shown to indicate the quality of aerodynamic estimates using this theory.

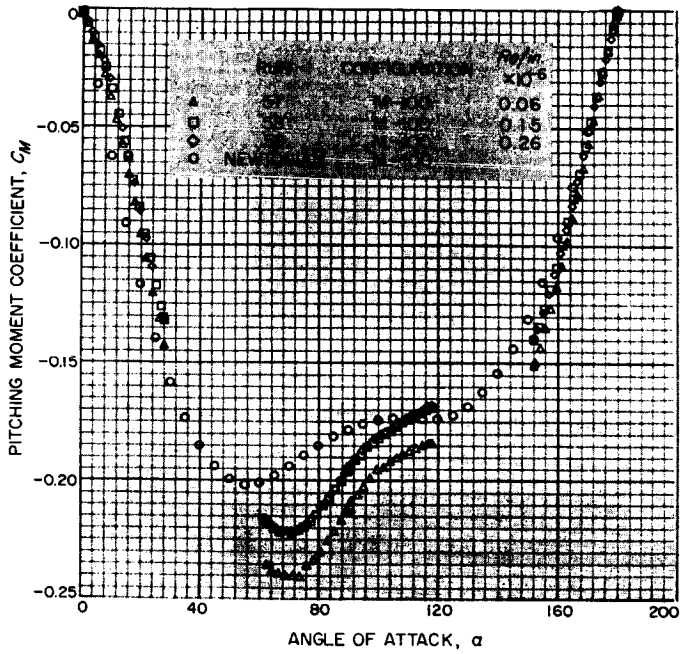


Fig. 10. Pitching moment coefficient versus angle of attack

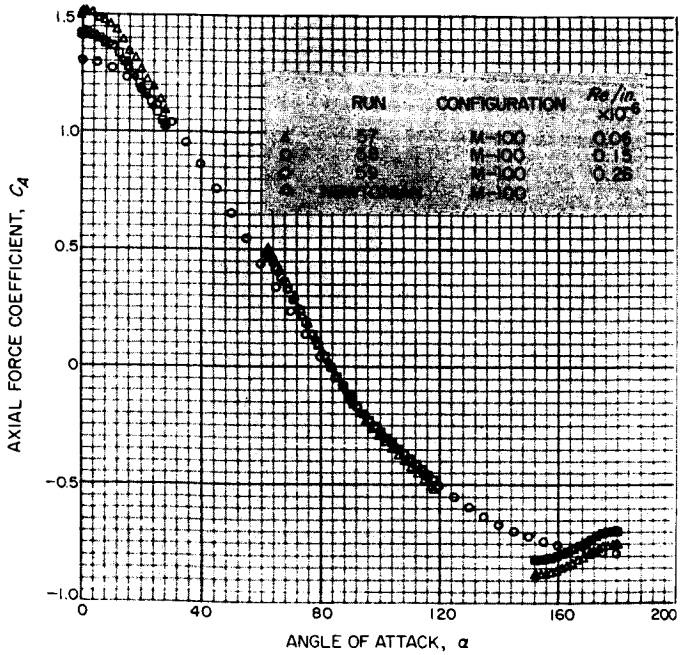


Fig. 11. Axial force coefficient versus angle of attack

Analysis to determine the most promising afterbody, in terms of rearward instability, is in progress.

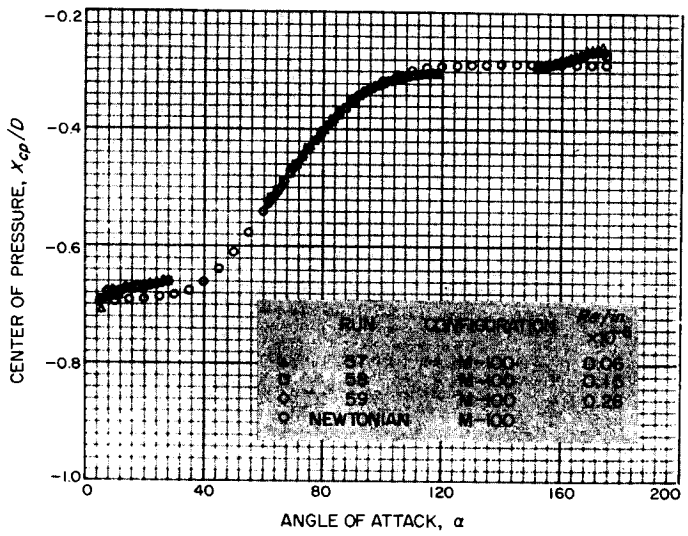


Fig. 12. Center of pressure versus angle of attack

2. *Free-Flight Wind Tunnel Body-Drogue Studies,* G. M. Goranson

The aeromechanics of the body-drogue system is the subject of investigation in a continuing research program in the JPL 20-in. Supersonic Wind Tunnel (*SPS 37-26*, Vol. IV, p. 78). Both drag and dynamic behavior data are being obtained via the free-flight model technique. Data reduction has been completed recently on an earlier test for drag data. The drogue drag coefficient, C_{D_d}, is defined as the body-drogue system total drag coefficient minus the forebody free-stream drag coefficient, converted to the drogue's reference area, i.e.:

$$C_{D_d} = \left(\frac{\text{System total drag}}{q_\infty A_D} - \frac{\text{Forebody free-stream drag}}{q_\infty A_D}\right) A_D/A_d$$

where

q_∞ = free-stream dynamic pressure

A_D = forebody reference area

A_d = drogue reference area

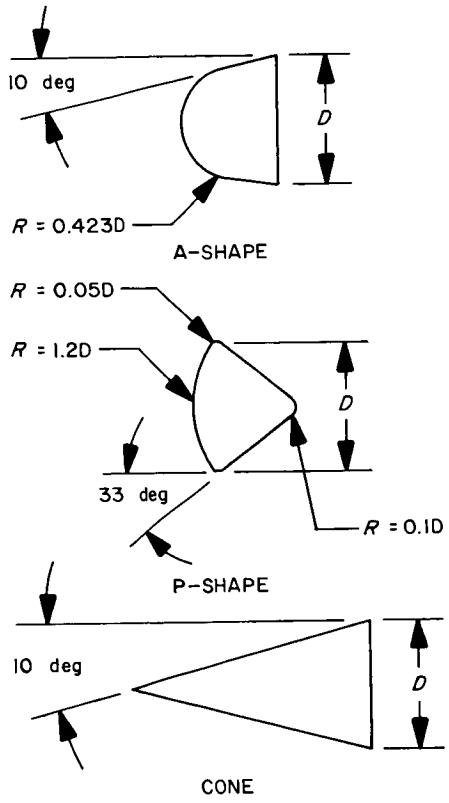


Fig. 13. Forebody shapes tested

Fig. 13 shows the forebody shapes tested. The flow conditions were Mach number 4.0, Reynolds number (based on forebody diameter, D) 85×10^3.

The data show a significant variation in drogue drag, for a spherical drogue, with three geometric factors:

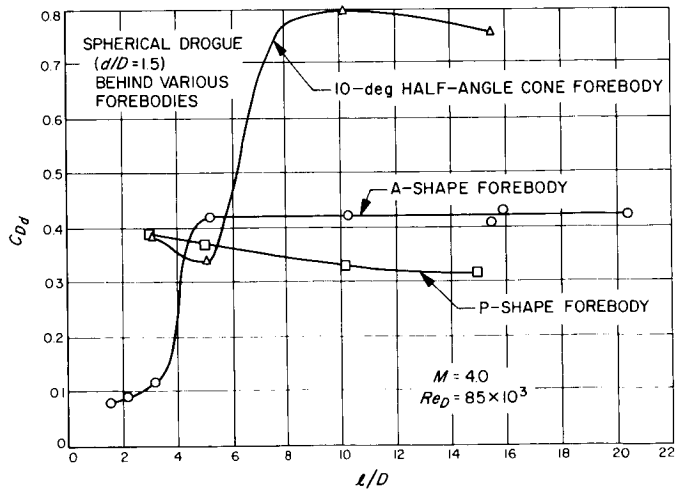


Fig. 14. Variation of drogue drag coefficient with length to forebody diameter ratio

(1) Forebody shape.

(2) Drogue trailing length to forebody diameter ratio, ℓ/D.

(3) Drogue diameter to forebody diameter, d/D.

In general, complex interactions exist between the above noted factors. Fig. 14 shows the variation, with drogue trailing length, of C_{D_d} for a spherical drogue trailing various forebodies. The effect of various drogue shapes on C_{D_d} as a function of trailing length behind the A-shape forebody is shown in Fig. 15.

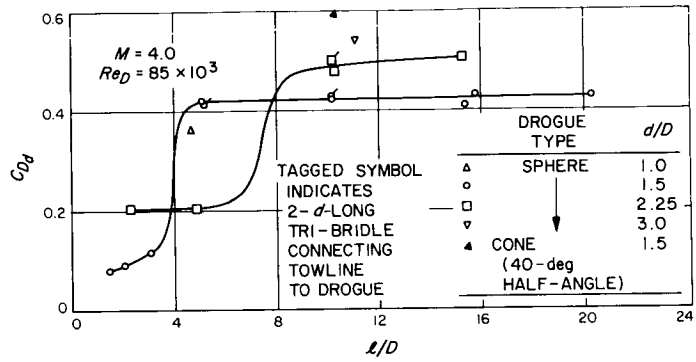


Fig. 15. Various drogues behind the A-shape forebody

3. *Aspect Ratio Effects,* B. Walker

Wind Tunnel Test 20-615 was an investigation of the effects of aspect ratio on two-dimensional flow separation. The approximate aerodynamic test parameters were Mach numbers 4.0 and 5.0 nominally, and Reynolds numbers/in. from 0.022×10^6 to 0.298×10^6. The test variable was aspect ratio (distance between inner fins/distance from leading edge to step) which was varied from 0.33 to 6.00.

Two models were tested. The first was an axisymmetrical annular configuration, 10 in. in diameter, with a 10-deg flare (Fig. 16). The chord length of the model was 7 in., with the flare beginning 3 in. aft of the leading edge. Effectively, the model was a two-dimensional flat plate with a 10-deg ramp wrapped around a diameter which allowed internal supersonic flow. Fig. 17 shows the results of tests with the annular model at Mach = 4.0. It indicates that for $AR \geq 1$ configurations, the centerline pressure distribution is indistinguishable from the "clean" (no fins) configuration.

The second model, a two-dimensional configuration, is shown in Fig. 18. It consists of a 7-in. flat plate with a 10-deg. wedge leading edge. A 10-deg wedge is located on top of the plate, beginning 3 in. aft of the leading edge.

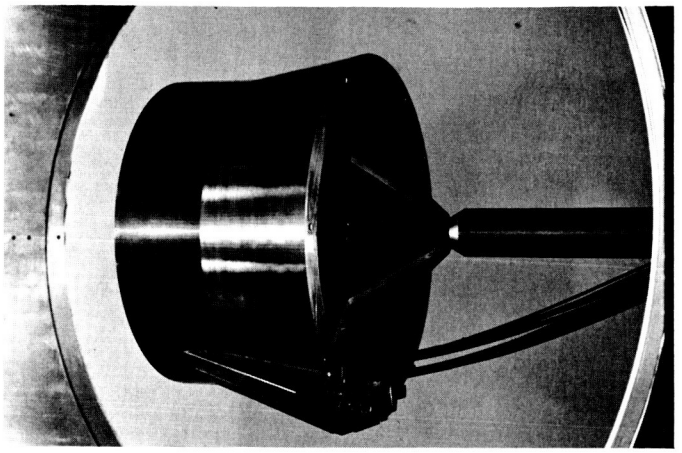


Fig. 16. Installation of annular configuration

The purpose of this configuration was to approach two-dimensional flow without possible end effects introduced by the fins, and thus the end effects, if any, could be determined. This model was also fitted with fins to simulate varying aspect ratios, and again the two-dimensional

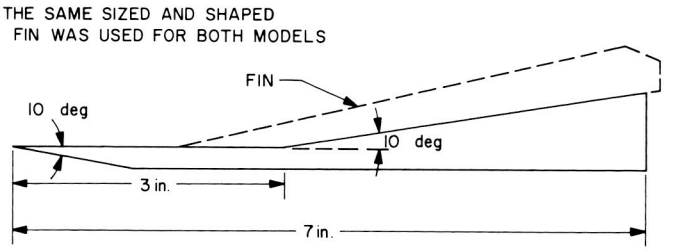


Fig. 18. Two-dimensional model configuration

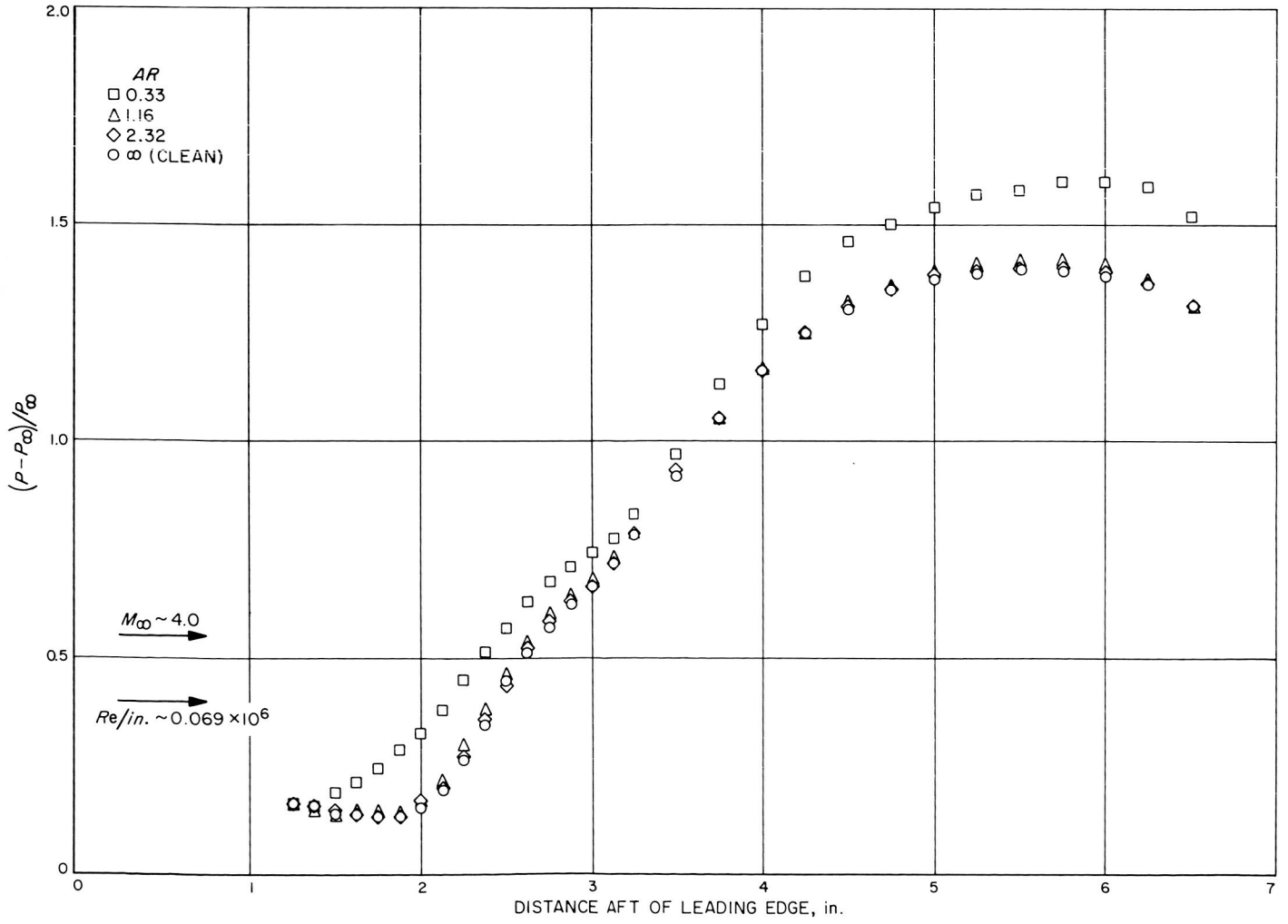


Fig. 17. Aspect ratio variation (annular configuration)

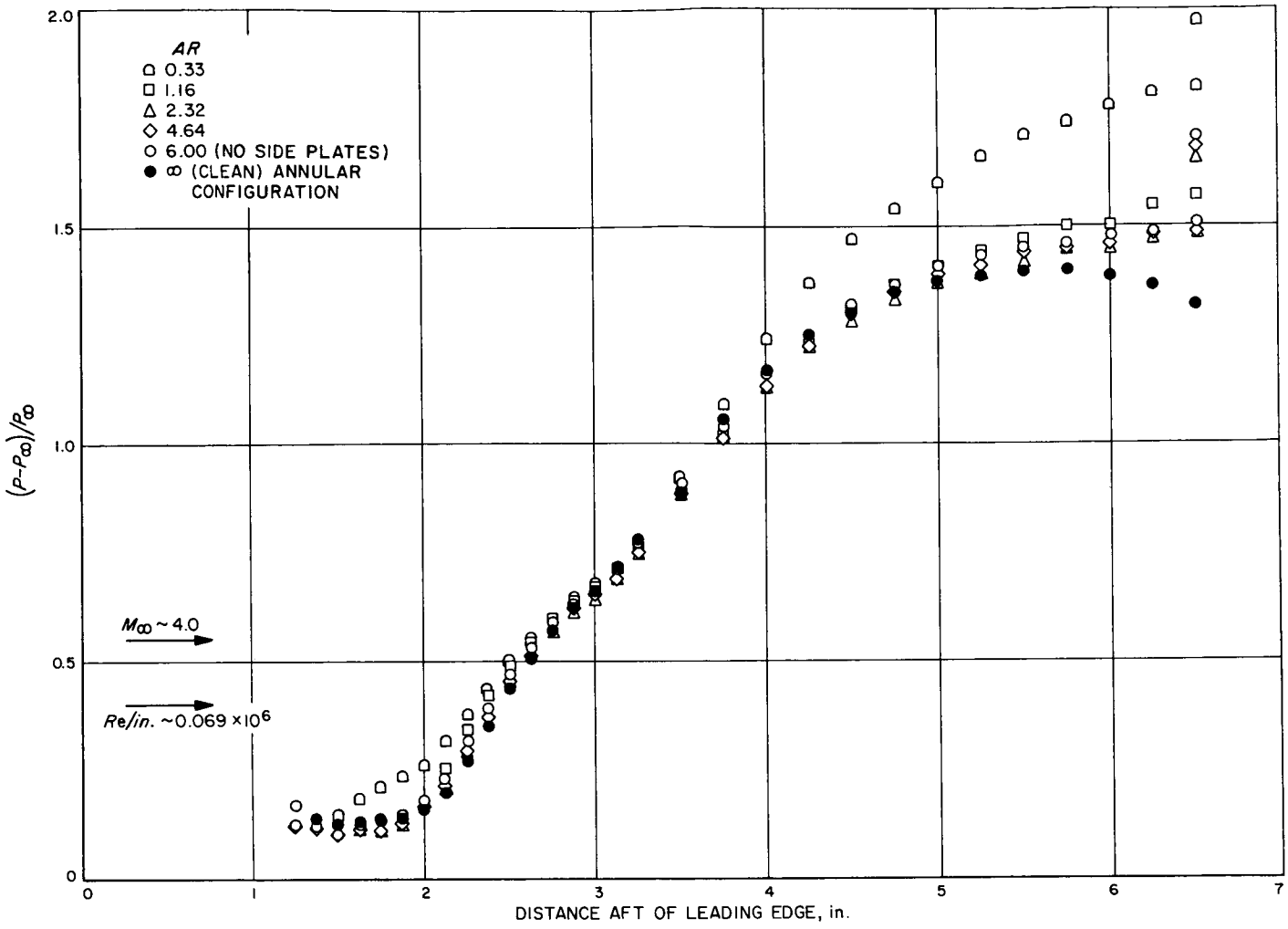


Fig. 19. Comparison of two-dimensional aspect ratio variation with annular configuration

flow limit was approached for $AR \geq 1$. This pressure distribution corresponds almost exactly to the annular configuration until approximately 2 in. along the ramp where the effect of the axisymmetrical configuration begins to appear (Fig. 19).

In addition to determining the aspect ratio limit for two-dimensional flow, the test demonstrated that the fins could be used to vary aspect ratio without introducing serious end effects.

4. Cone Wake Study, B. Walker

Pitot pressure, static pressure, and recovery temperature surveys at various stations behind a wire-supported cone were obtained during Test 20-613, performed in cooperation with the Space Technology Laboratories, Redondo Beach, California. Model configurations for this test consist of a series of blunted 10-deg half-angle cones with a base diameter of 1.5 in. Nose-to-base radii bluntness ratios for these models were 0, 0.104, 0.268, and 0.4. The noses of the models were aluminum and the aft portions were made of plastic.

The models were supported by a 0.010-in. diameter vertical wire. In order to minimize the support interference 0.020-in. diameter boundary layer wire trips were placed on the model. They were located 0.38 to 0.50 in. aft of the nose of the $RN/RB = 0$ model, and one-fourth of the length of the model aft of the nose on the remaining three models.

Pitot pressure surveys were taken through the wake with a probe 0.032 in. in diameter, using a close-coupled transducer. The probe was mounted on the vertical traverse assembly with the X, Y, Z, attachment, (Fig. 20). A

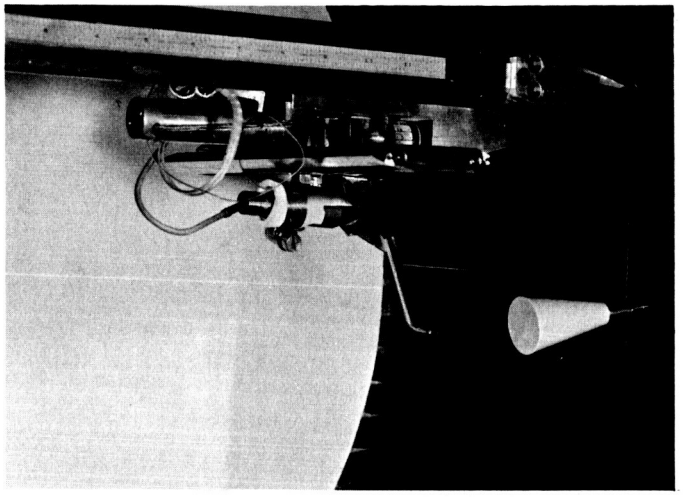


Fig. 20. Installation of cone, pressure probe, and thermocouple in the supersonic wind tunnel

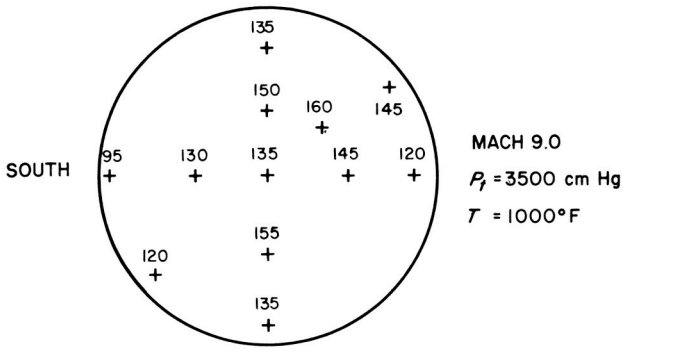
thermocouple mounted on the same assembly 1.5 in. above the pitot probe was used to make the temperature surveys. A static pressure probe with three 0.018-in. diameter pressure taps was mounted on the vertical traverse assembly for the static pressure measurements.

A pretest check was made to determine the effect of size of wire on the wakes of models with and without trips. The purpose was to find a wire which caused minimum interference and yet was strong enough to support the model in flight. Wires of 0.015, 0.010, and 0.007 in. in diameter were tested. The wakes of the models supported by the different size wires were compared to the wake of a free-flight model. The necks of the wakes of the models supported by the 0.010 and 0.007-in. diameter wires were approximately the same distance aft of the model as the neck in the wake of the free-flight model. Consequently, the 0.010-in. diameter wire was used because of its greater strength.

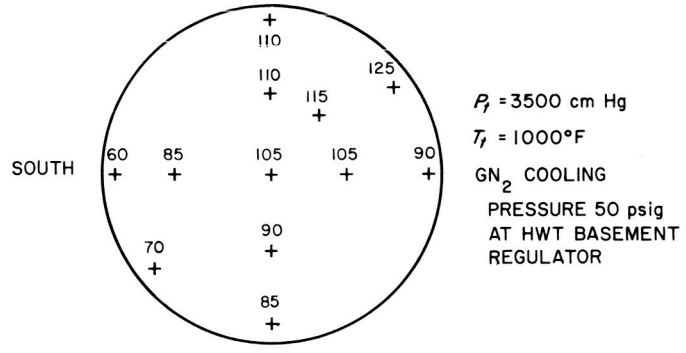
The specific test conditions for this test were Mach 4.0 at maximum Reynolds number, nominally 3.5×10^5/in., at a supply pressure of 240 cm Hg. No free-flight data were obtained to confirm the wire-supported data.

5. Operation of the 21-in. Hypersonic Wind Tunnel Schlieren System with Quartz Windows Removed, *J. Welton*

The JPL 21-in. Hypersonic Wind Tunnel (HWT) nozzle flex plates, side wall and window frames are water cooled to approximately 60°F. When operating the facility at supply temperatures greater than 500°F the quartz schlieren windows are heated quasi-uniformly by the boundary layer and cooled conductively by the window frames. This uneven cooling creates density gradients in the quartz and is observed as uneven darkening of the schlieren image. The temperature distribution over the outside of one 18-in. diameter quartz window is shown in Fig. 21(a). In order to reduce the schlieren degradation during higher temperature operation, methods for reducing the window temperature gradient were investigated.



(a) NORMAL TEMPERATURE READINGS, °F



(b) COOLED TEMPERATURE READINGS, °F

Fig. 21. Window temperature distribution

The first approach to the window distortion problem was an attempt to cool the air-wetted surface of the windows. A manifold was constructed to discharge gaseous nitrogen over the window surface (Fig. 22). This method had a general cooling effect; however, no significant improvement to the surface temperature distribution was apparent [Fig. 21(b)]. No attempt was made to improve on this method as the hardware was bulky, and favorable results were not expected.

Another method to eliminate the window problem was investigated: to remove the windows from the optical system. The HWT schlieren utilizes a completely enclosed

light path which, during operation, is evacuated to the test section static pressure. The tunnel windows are, therefore, not necessary under this mode of operation.

Removal of the windows greatly improved the quality of the schlieren image, yet the effect on the establishment of flow or the continuous operation of the wind tunnel was negligible.

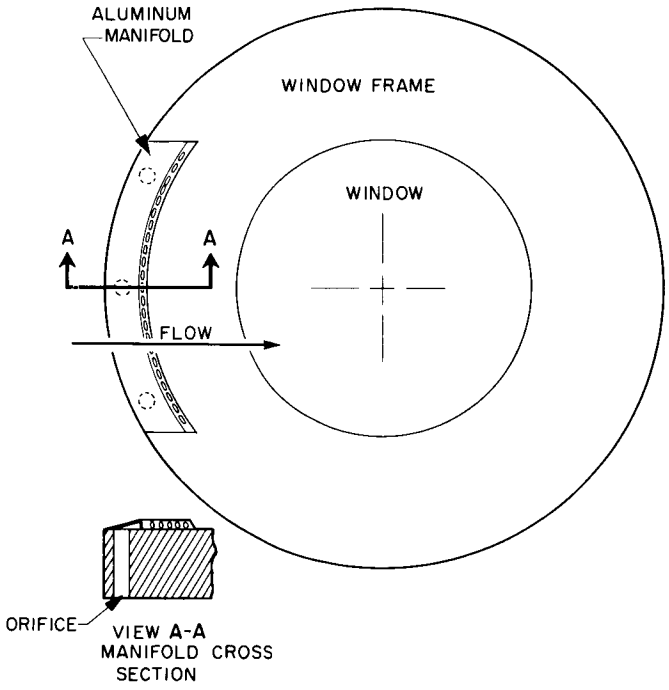


Fig. 22. Cooling manifold configuration

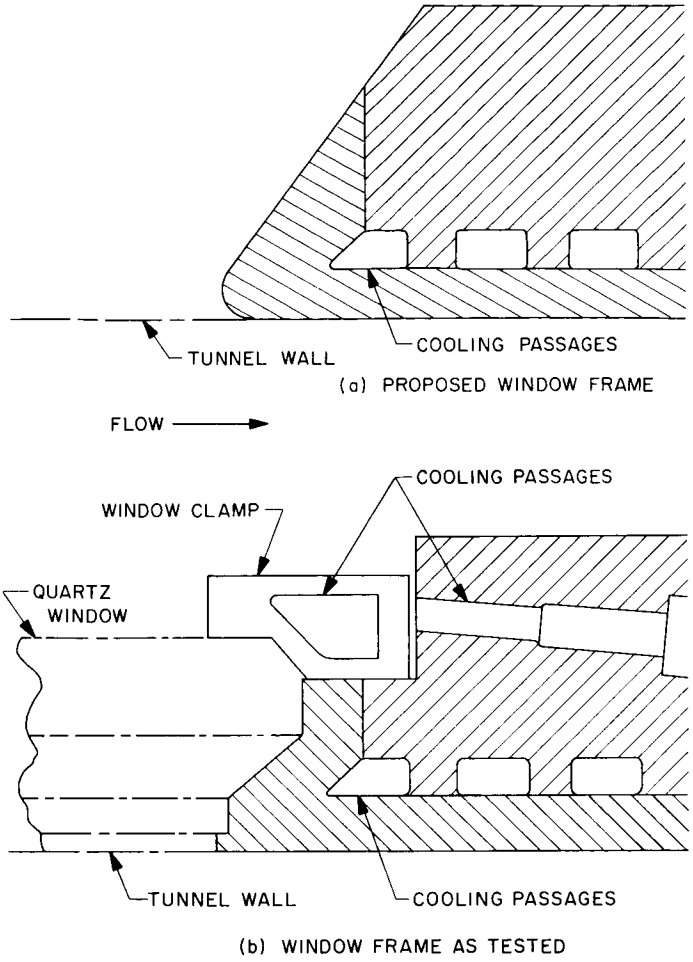


Fig. 24. Window frame comparison

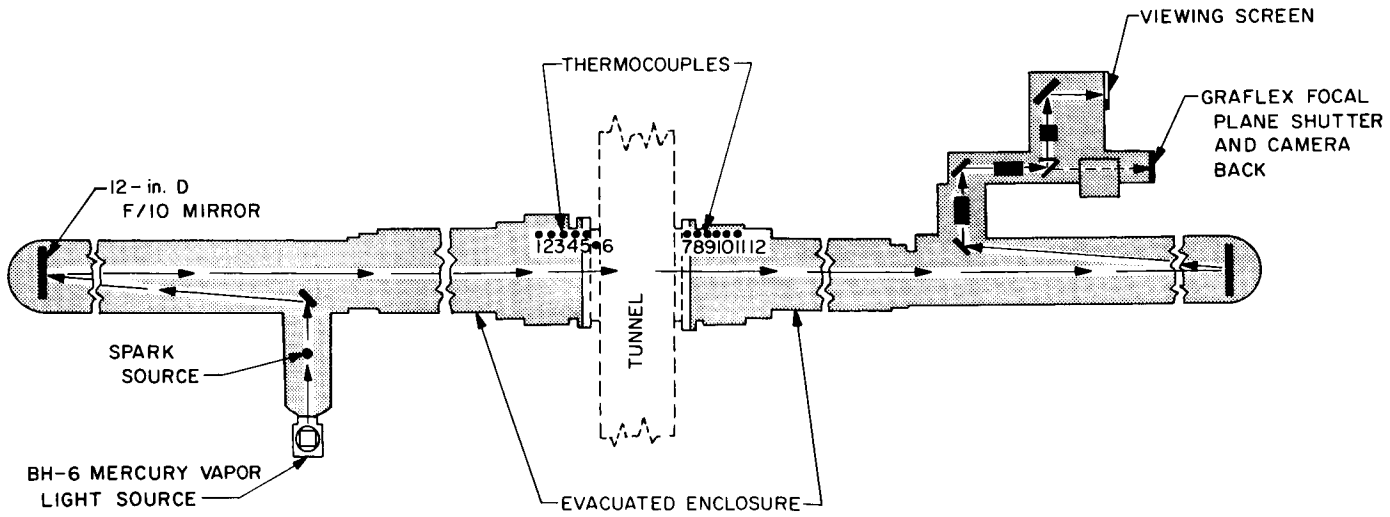


Fig. 23. Tunnel and schlieren configuration

Twelve thermocouples were attached to the schlieren enclosure (Fig. 23) in order to detect any extreme local heating from the air. The steel schlieren enclosure is not water cooled. The highest indicated temperature was at the downstream edge of the window frame opening (TC No. 6). At Mach 9, with a supply pressure of 3500 cm Hg and a supply temperature of 1000°F, this temperature was 350°F. The thermocouples were located ½ in. away from the surface of the enclosure, and the enclosure surface temperature at the above given conditions and in the area of the thermocouples, was approximately 90°F, and no hot spots were detected.

The only detrimental effect observed while operating the HWT without windows was the unusual amount of dust that collected on the exposed optical surfaces. As the tunnel was returned to atmospheric pressure, the enclosure was exposed to air flows not experienced during normal operation.

Except for the dust problem mentioned, the tunnel operates in a normal manner throughout the pressure and temperature range at Mach 9. With a minor improvement to the window frame, this method of operation may be considered a standard technique, particularly when high-quality photographs are required. The proposed improvement would include a window frame constructed just for this use; i.e., a 12-in. diameter opening with no provision for mounting a window. The opening need not be as large as 18 in. because the schlieren mirrors, and hence the light path, are 12 in. in diameter. A comparison of the existing and proposed window frames is made in Fig. 24.

B. Hypervelocity Laboratory

B. Riale and T. Babineaux

1. Performance Evaluation of a 12-in. Diameter Free Piston Shock Tube

Recent shakedown runs on the 12-in. diameter free piston shock tube (Fig. 25) show that it is now operational.

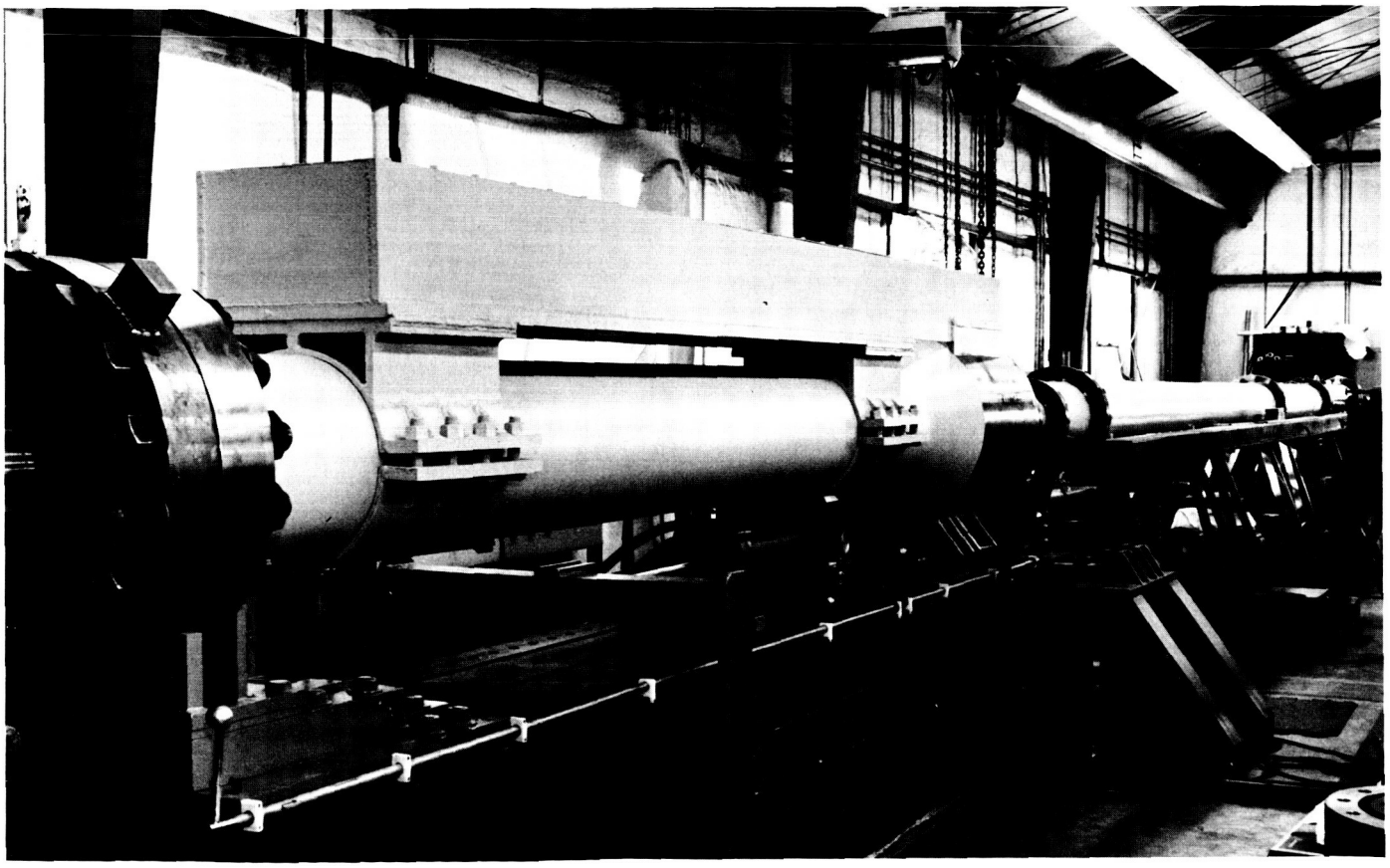


Fig. 25. JPL free piston shock tube

Measurements of shock speed were made using 6 thin film gages mounted on the shock tube. Helium was used as the piston driving gas and also as the driver gas. Air was used as the driven gas.

The initial pressures behind and ahead of the piston were 260 and 2 psia, respectively. This pressure ratio caused the piston to travel to the end of the tube and stop because of the 1100 psia developed in the 6-in. diameter driver insert. At the end of the stroke, the spike caused the scribed brass diaphragm to rupture. The diaphragms used had a static burst pressure of 1300 psia.

A problem was incurred when the aluminum piston repeatedly impacted the driver volume after the diaphragm opened, causing the piston to become deformed and to have a slightly larger diameter. This problem was solved by using a flat piece of silicone rubber in the breech of the tube to cushion the impact of the piston. At the low pressure ratios now used, this silicone rubber does not burn or melt. At the present time, it is not known if the silicone rubber can withstand higher pressures and associated temperatures planned for future shakedown runs.

The response of the thin film gages to passage of the shock wave was photographically recorded from oscilloscope traces. The average shock speed between the gages was calculated from the measured times and known distances between the gages.

Fourteen runs were made into initial driven tube pressures (P_1) of 0.25, 0.10, and 0.05 mm Hg. The shock speeds are given in Table 1.

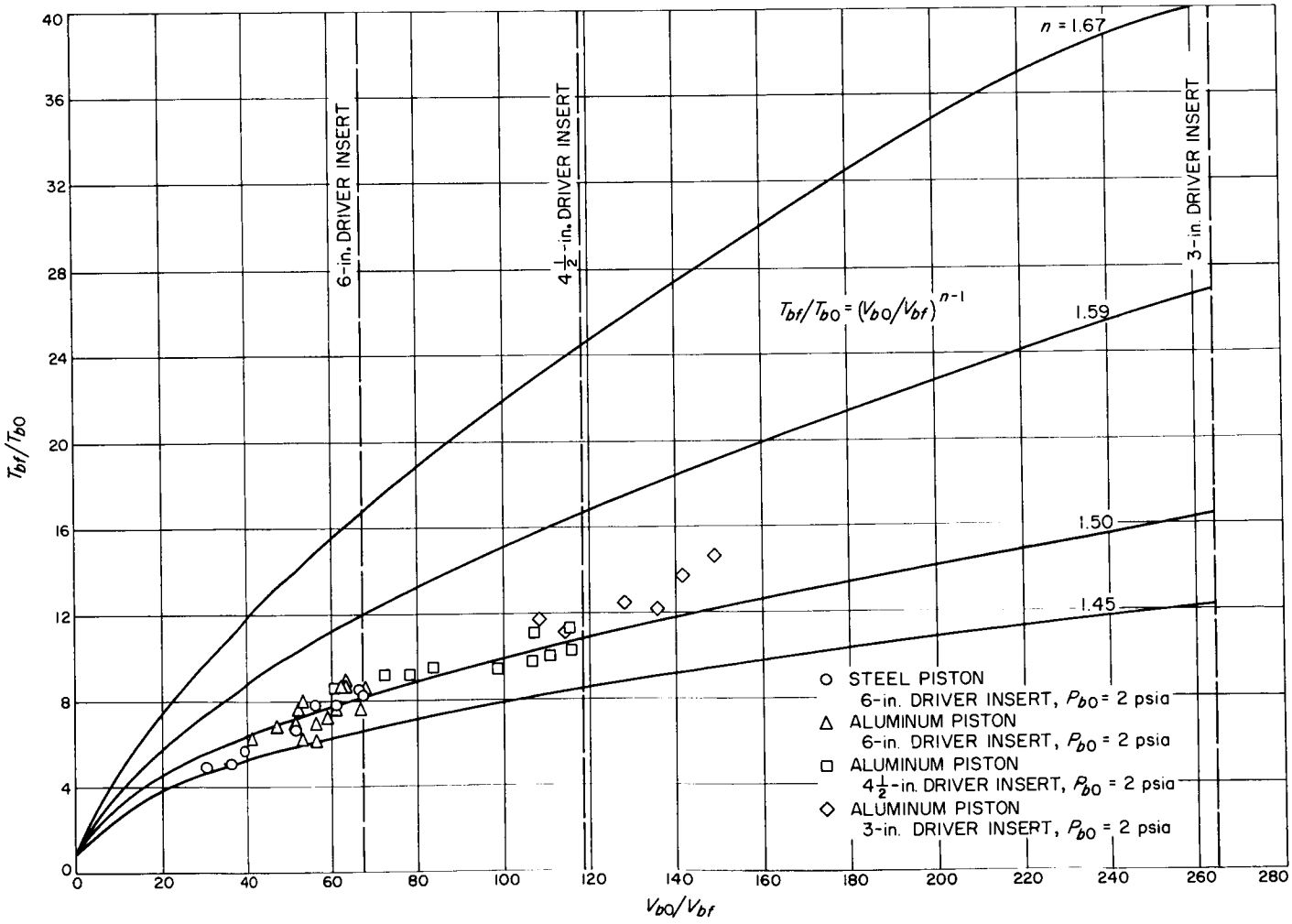


Fig. 26. Final temperature in the driver as a function of volume showing the effects of different piston weights

Table 1. Summary of shock tube performance

Run No.	P, mm Hg	Average shock speed between gages, ft/sec		
		6 and 7	8 and 9	10 and 11
279	0.25	18,990	19,290	—
281	0.25	16,860	—	—
283	0.25	17,580	17,410	16,720
284	0.10	16,090	16,090	15,190
285	0.10	22,390	25,190	20,300
286	0.10	20,730	21,190	19,720
288	0.10	19,490	19,490	19,030
289	0.10	20,700	20,950	20,120
290	0.05	22,550	23,200	22,050
291	0.05	22,610	22,870	21,350
292	0.05	22,350	22,500	21,550
293	0.05	21,950	22,390	21,100
295	0.05	22,390	22,550	21,490
296	0.05	22,150	22,320	21,030

A steel piston (315 lb) was used for 8 runs in the same manner as the aluminum piston (115 lb) in previously reported driver investigation runs (*SPS 37-32*, Vol. IV, p. 101, and *SPS 37-34*, Vol. IV, p. 111). Theory predicts that the heat loss to the wall of the tube is less using the light piston (Ref. 1). Stroke time is longer with the heavy piston, and therefore, the losses are greater since there is more time for heat loss to the wall. A comparison of the results using the light and heavy pistons in the driver investigation mode is given in Figs. 26 and 27. The difference in the data is not immediately noticeable at the lower values of piston tube pressure P_{b_0}; however, it is thought that as P_{b_0} is increased, requiring larger driving piston pressures, the effects of the heavier piston will be more pronounced, resulting in lower values of the effective polytropic exponent n, and reflected in lower values of driver pressure and temperature.

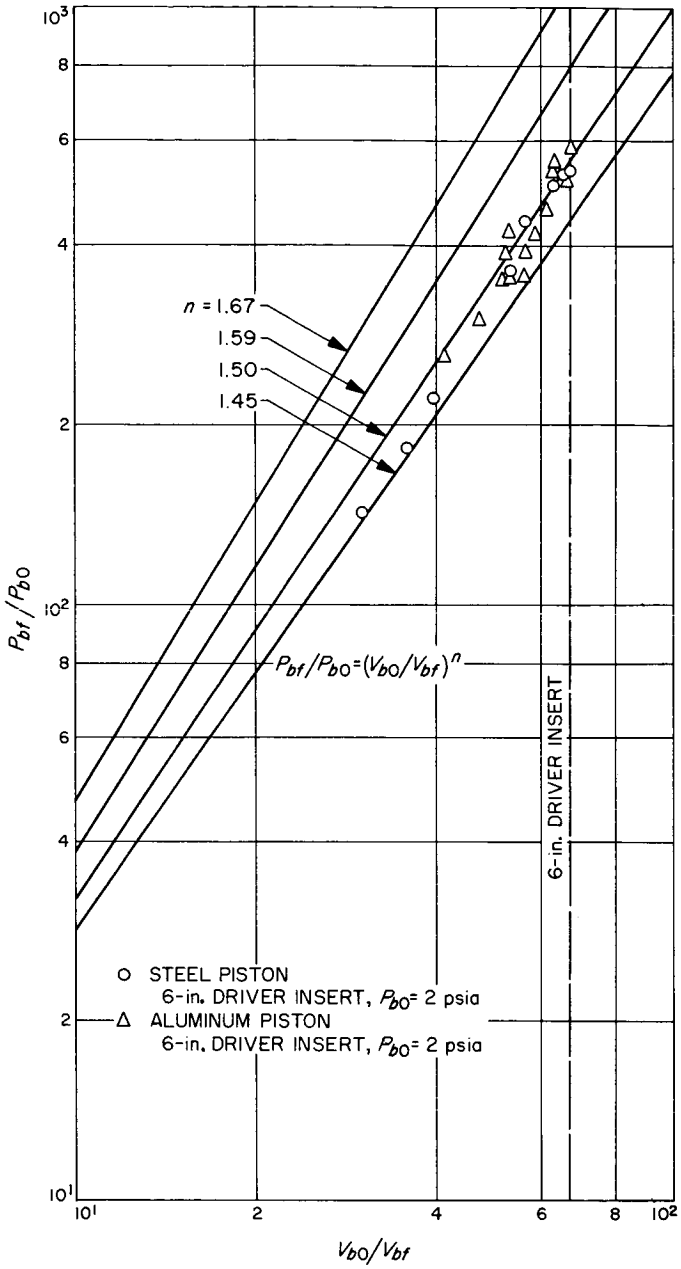


Fig. 27. Final pressure in the driver as a function of volume showing the effects of different piston weights

Reference

1. Roffe, G. A., *The Free Piston Shock Tube Driver: A Preliminary Theoretical Study*, Technical Report No. 32-560, Jet Propulsion Laboratory, Pasadena, California, December 15, 1963.

PROPULSION DIVISION

XI. Polymer Research

A. Mechanical Behavior of Styrene-Butadiene Rubber Filled with Glass Beads

R. F. Fedors and R. F. Landel

In a previous study (Ref. 1) a series of vulcanized SBR-glass bead composites containing from 0 to 0.339 volume fraction ϕ, of glass bead filler were prepared and characterized. The present purpose is to present data obtained on the mechanical response of these composites when subjected to uniaxial tension at various test rates and test temperatures.

1. Experimental

The stress-strain data were obtained with ring specimens having an inside diameter $D_i = 0.750$ in. and outside diameter $D_o = 0.875$ in. Since the stress is nonuniform across the cross-section of the ring specimen, an extrapolation technique described elsewhere (Ref. 2) was used in order to minimize or eliminate this effect. The strain ϵ, is calculated from the measured crosshead displacement ΔL, by means of

$$\epsilon = \frac{2\Delta L}{\pi D_i} \qquad (1)$$

The stress σ, is calculated by dividing the load by the initial cross-sectional area of the ring.

A summary of pertinent data from Ref. 1 is listed in Table 1. It was intended to keep the chain concentration ν_e, constant and thus study the behavior of the composites with ϕ alone varying. However, as can be seen in the table, ν_e increases slowly with ϕ from a low value of 102×10^{-6} moles/cm^3 for the unfilled material to 142×10^{-6} moles/cm^3 for the composite with the largest ϕ value. Fortunately, the total variation in ν_e is small and furthermore, the variation can be taken into account by a normalization procedure (Ref. 3).

Table 1. Some properties of SBR-glass bead composites

Sample designation	ϕ	$\nu_e \times 10^6$, moles/cm^3	Density, g/cm^3	Gel fraction
CA	0	102	0.9430	0.887
3A	0.0922	111	1.086	0.926
4A	0.132	108	1.128	0.934
5A	0.169	115	1.182	0.942
6A	0.204	124	1.240	0.947
7A	0.234	121	1.273	0.951
8A	0.270	131	1.329	0.954
9A	0.289	133	1.369	0.958
10A	0.339	142	1.439	0.960

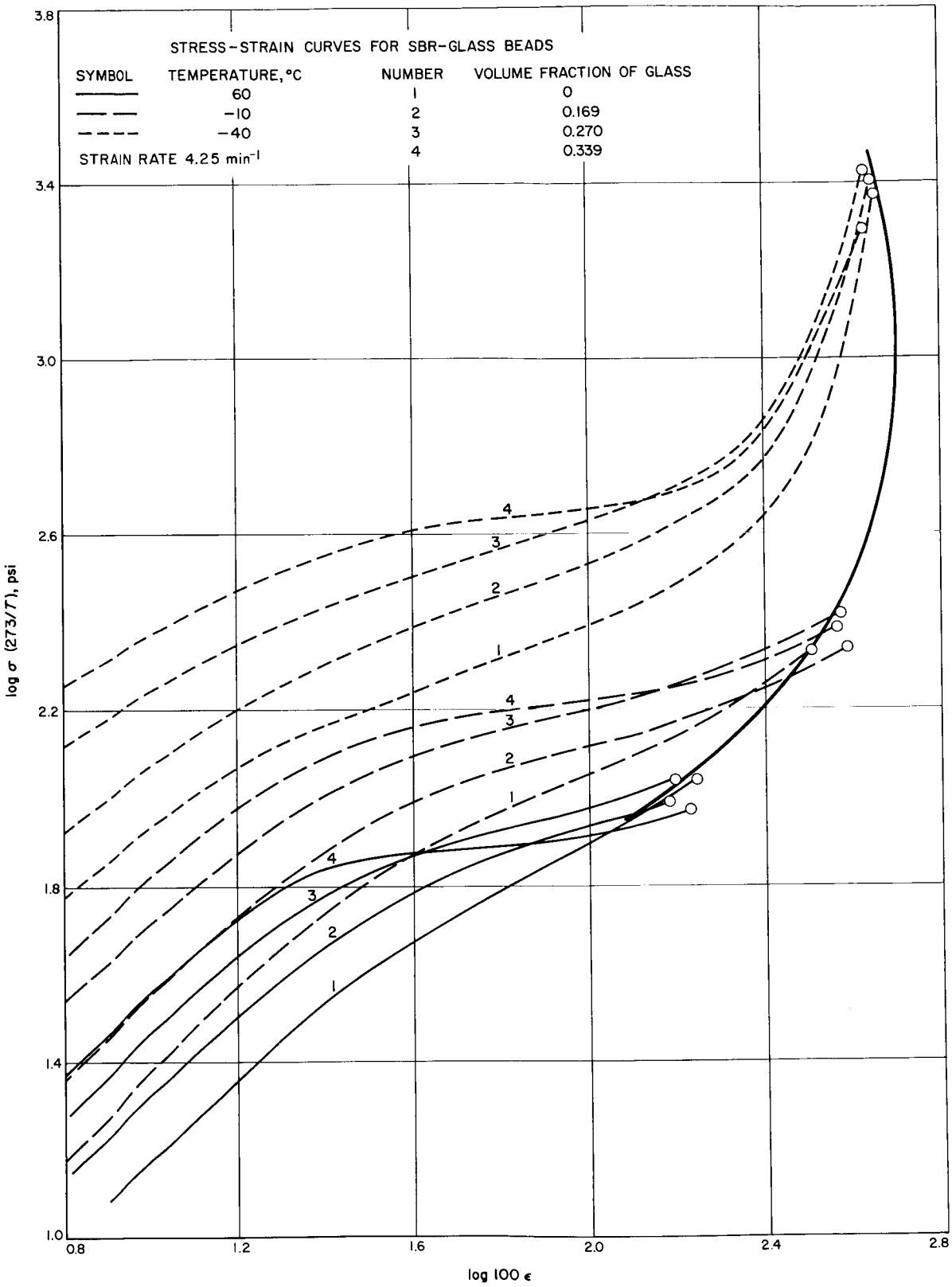


Fig. 1. Stress-strain relations as a function of temperature for selected composites. Unfilled circles represent the break points. The curve to the right represents the location and shape of the failure envelope obtained using all the data.

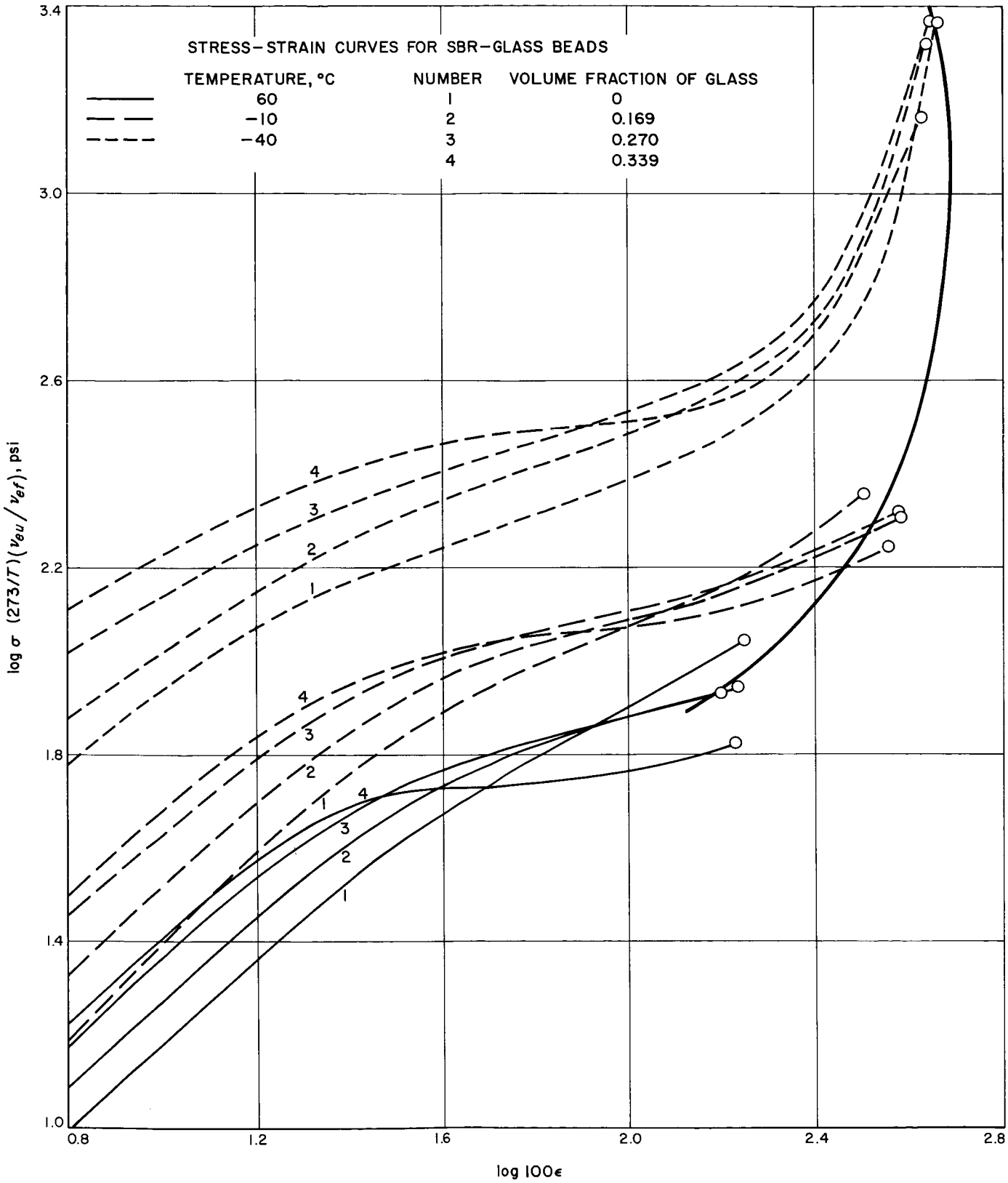


Fig. 2. The data of Fig. 1 replotted with the stress reduced by the ratio of the unfilled ν_{eu} to the filled composite ν_{ef}

2. Results

The effect of temperature and ϕ on the stress-strain curves obtained at a rate of 4.25/min are shown in Fig. 1 for values of $\phi = 0$, 0.169, 0.270 and 0.339 and for temperatures of +60, −10 and −40. Since the stress-strain data are so extensive, and, furthermore, since the data shown are typical of the curves obtained at the other temperatures, rates, and ϕ values, Fig. 1 will suffice to depict the general stress-strain behavior for these composites.

It is apparent from Fig. 1 that the stress at small strain levels increases monotonically with ϕ, as expected. In addition, Curves 2, 3, and 4 approach the stress-strain curve for the unfilled material at higher strains; this effect is presumably caused by dewetting. At break, denoted by the unfilled circles, the values of both the stress-at-break σ_b and the breaking strain ϵ_b are essentially independent of the volume fraction of non-reinforcing glass fiber for at least ϕ values up to 0.34. This behavior is supported by preliminary data obtained with silicone rubber-glass bead composites having ϕ from 0 to 0.50. The curve to the right represents the position and the shape of the failure envelope obtained when all the failure points are plotted.

As indicated in Table 1, ν_e increased slightly with ϕ. In order to eliminate this variation, the stress-strain curves shown in Fig. 1 have been replotted with σ normalized by the factor of the ratio of ν_e of the unfilled rubber to ν_e of the filled composite. These results are shown in Fig. 2. A comparison of the curves in both figures indicates that the major effect of the normalization has been to bring the stress-strain curves closer together so that the spread, for example, between the curve for $\phi = 0$ and $\phi = 0.34$ is less. The open circles again represent the occurrence of rupture and the curve to the right represents the position and shape of the failure envelope obtained using all the break data.

3. Effect of Filler on Initial Modulus

An attempt was made to assess the effect of ϕ on the small-strain modulus E. The latter was evaluated from $\log \sigma - \log \epsilon$ plots such as shown in Fig. 1. At sufficiently small strains such plots should be linear and have unit slope. The evaluation was hampered by the fact that the experimental procedure (ring-shaped specimens) did not permit an accurate determination of strains in the region below about 10%. The required unit slope was found only at 40 and 60°C, so only these data were used in the analysis of the E, ϕ relationship.

Previous work (Refs. 4, 5, 6, 7) has indicated that a simple expression which accounts for filler effects is:

$$E/E_0 = \left(1 + \frac{1.25\,\phi}{1 - \phi/\phi_m}\right)^2 \qquad (2)$$

where E_0 is the modulus of the gum rubber and ϕ_m is the maximum filler content which the system can attain. For the present system, this should be about 0.70, the values observed for glass beads in mineral oil when the glass has been coated with a surfactant. (The rubber contains residual fatty acid soaps used in the manufacture.)

In Fig. 3, the ratios of the stress on the filled composite to that on the unfilled material at $\epsilon = 0.1$ are depicted as a function of ϕ for data obtained at a strain rate of 4.25 per minute. The stress values for the filled composites were reduced to a unit ν_e in order to take account of the small variation in ν_e shown in Table 1.

The expected curve, according to Eq. (2), lies considerably above the experimental data. The discrepancy cannot be associated with the choice of ϕ_m, since a negative value (a physical impossibility) is required if Eq. (2) is to describe the results. Instead, the differences are probably due to the fact that significant dewetting had occurred in the filled composites even at the small strains used to evaluate E. Hence the composites represent ternary systems of rubber, glass and voids, so that Eq. (2) and other equations developed for two-component systems (Refs. 8, 9, 10, 11, and 12) cannot be expected to be applicable.

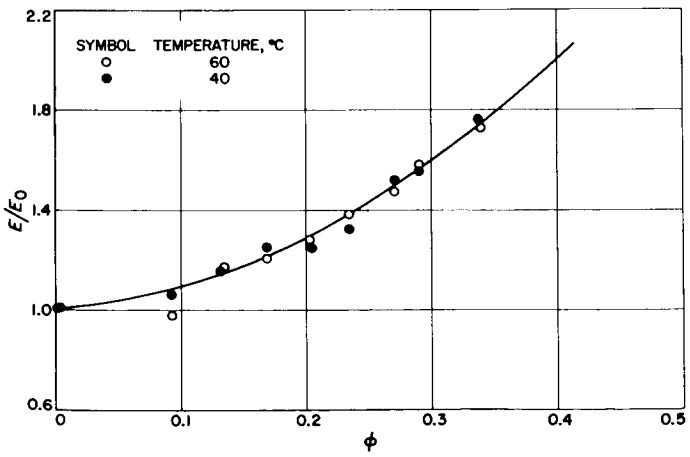


Fig. 3. Dependence of the ratio of the modulus of the filled composites E to that of the gum rubber E_0 on ϕ. The E values have been reduced in order to take account of the small variation of ν_e with ϕ shown in Table 1

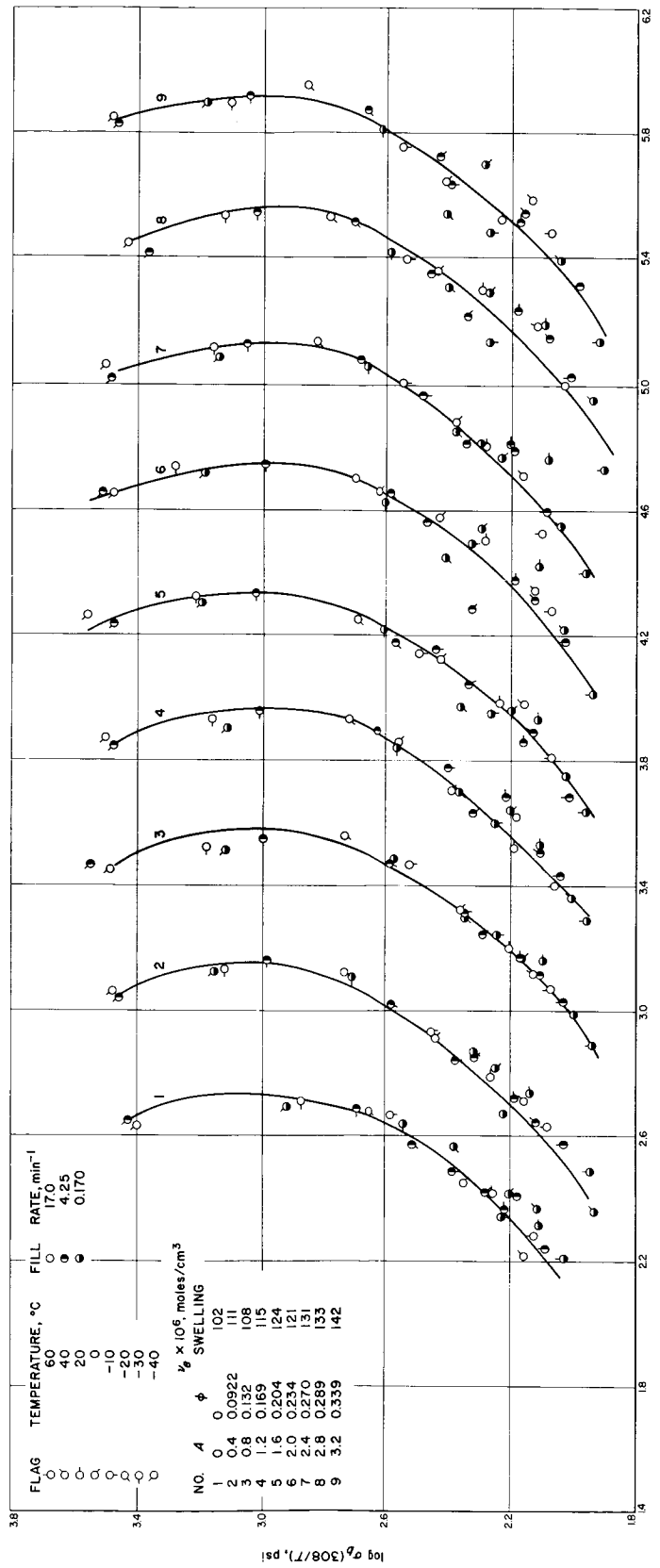


Fig. 4. Failure envelopes for the composites as a function of ϕ. Each curve, except the gum rubber, has been shifted along the abscissa for ease of presentation. All the curves shown have the same form

The data of Fig. 3, may be fitted to a quadratic in ϕ, as a convenient analytic description of the results:

$$E/E_0 = 1 + a\,\phi + b\,\phi^2 \qquad (3)$$

In simple systems (Refs. 7, 8) the values of a and b would be 2.5 and 14.1; the actual values here are considerably lower, $a = 0.5$ and $b = 5.0$.

4. Failure Properties

The values of σ_b, and ϵ_b, determined as a function of test rate and test temperature are depicted in Fig. 4 as failure envelopes. Each envelope represents the data obtained at a particular value of ϕ. For ease of presentation, each envelope, except for that for the unfilled rubber, has been shifted along the abscissa by A log units, with the value of A shown in the legend.

It is evident from the figure that the presence of glass beads as a filler has not significantly affected either the shape or location of the envelope. Even more remarkable and unexpected is that fact that presence of the filler has little or no effect (beyond experimental error) on the actual values of σ_b and ϵ_b obtained when test rate and test temperature are fixed, but ϕ is allowed to vary from 0 to 0.34; that is, the failure properties for this system are independent of ϕ. Preliminary data obtained with silicone rubber-glass beads with ϕ ranging from 0 to 0.5 support this independence.

In Fig. 5, the envelopes of Fig. 4 have been replotted with σ_b reduced to unit ν_e values. The solid line shown in the figure has the same shape as that shown in Fig. 4. Although the scatter is relatively large, most experimental points are within $\pm 15\%$ of the average curve drawn through the points. It is also readily evident from Fig. 5 that the maximum values of ϵ_b are also independent of ϕ, at least for ϕ values less than about 0.34. Also shown in the figure as the dotted curve is a portion of the common response curve of the reduced failure envelope taken from Ref. 3. The agreement is satisfactory.

5. Summary

Vulcanized SBR-glass bead composites have been characterized by density and swelling measurements and the stress-strain behavior in uniaxial tension has been determined as a function of ϕ, the test rate and the test temperature. The crosslink density determination of the composites rests in large measure on the assumption that the glass beads do not adhere to the rubber phase in the swollen gel.

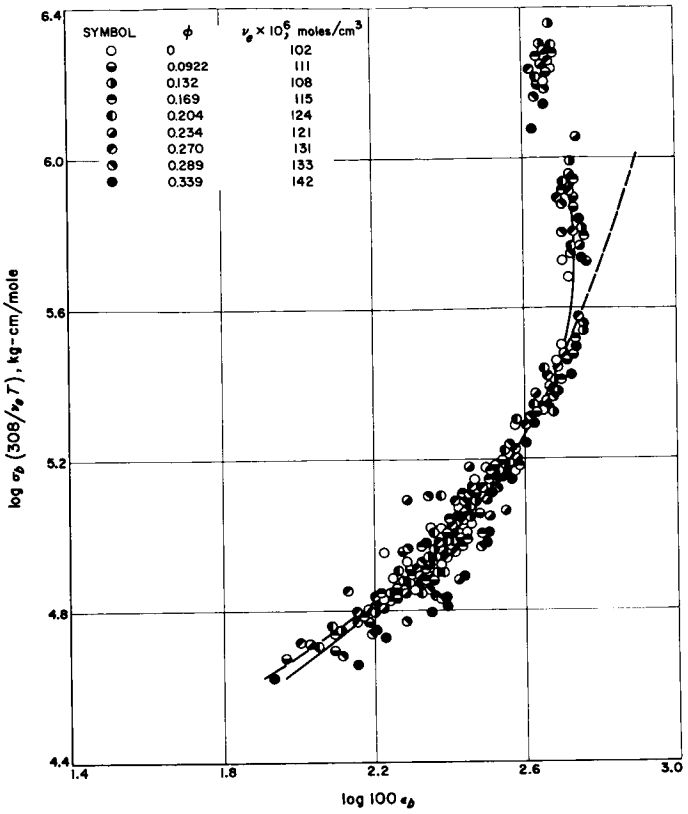


Fig. 5. Reduced failure envelope for the data shown in Fig. 4. The full curve has the same shape and form as that depicted in Fig. 4. The dashed curve represents the common response curve of the reduced failure envelope obtained with gum rubbers

Stress-strain curves to break were determined as a function of test rate and test temperature. As expected, the stress at small strain levels increases with ϕ; however, at the larger strains, significant dewetting presumably occurs, so that all curves at a given rate and temperature approach the gum rubber curve. The dependence of the stress, for very limited data, at a given small strain in the region where a plot of log σ versus log ϵ is linear with unit slope can be represented by a quadratic in ϕ; however, the constants are much smaller than a simple theoretical equation would predict. The failure envelope for all composites studied are identical within experimental error. Hence, the maximum value of ϵ_b is independent of ϕ. In addition, the values of σ_b and ϵ_b obtained at a given rate and temperature are essentially independent of ϕ. The reduced failure envelope for these composite systems agrees with the common response curve of the reduced failure envelope.

B. The Effect of Surface Modifications on the Burning Rate of a Composite Solid Propellant

B. G. Moser, R. E. Weich, Jr., and R. F. Landel

1. Introduction

In solid propellant technology certain areas have been intensively investigated, such as the chemistry of the formation and degradation of the binder, the chemistry of the oxidizer, and the size, size distribution, and shape of the oxidizer (and other filler) particles. However, one important parameter has been largely neglected — the surface characteristics of the oxidizer. The effects of surface modification of the oxidizer particles in solid rocket propellant are twofold: (1) rheological, i.e., the mix viscosity and tensile properties of the cured propellant and (2) ballistic, i.e., the burning rate. The former having been discussed in prior publications (Ref. 13), this paper will be limited primarily to the burning rate changes induced by surface modifications.

When a surfactant is added to a slurry of oxidizer and binder ingredients, a rather large change in burning rate may occur. For example, in a slurry containing 20% by weight of mineral oil and 80% of ammonium perchlorate (AP), the burning rate at 1000 psi is decreased by several hundred percent by the addition of only 0.1% of a surfactant, but in a similar slurry in which the liquid phase is polypropylene oxide (PPO), the use of a surfactant can result in an increase of 10%.

It is possible that these changes in rate originate either from effects on the decomposition of the AP or from changes in the degree of dispersion. With normal burning rate accelerators such as MgO or copper chromite, it is known that both the predecomposition temperature and the main decomposition temperature, as observed in differential thermal analysis test, are decreased. On the other hand, it is known from rheological studies that the solids are more uniformly dispersed in the presence of an effective surfactant.[1] In the poorly dispersed system it might be expected that the resulting material structure of close contacts will burn as though the oxidizer concentration were much higher than it actually is.

We have adopted these two facts (changed decomposition and degree of dispersion) as working hypotheses for the explanation of the changed burning rate, and in this report test them against experimental data. When a surfactant is used to modify the oxidizer surface it is very difficult to assess quantatively which factor — dispersion or decomposition — is the dominant burning rate modifier, especially in polyurethane systems where the effects are not too large. However, the use of a surfactant in a mineral oil–AP system results in a much greater change in the degree of dispersion and an accompanying great decrease in the burning rate. One of the paradoxes in the use of surfactants is that an effective surfactant can also be used to increase the burning rate by permitting a higher solids loading and attendant higher flame temperature (Ref. 13).

2. Dispersion Effects on the Burning Rate

When a surfactant is added to a polyurethane propellant slurry, there is an optimum surfactant concentration for maximum rheological effectiveness, e.g., minimum viscosity, a phenomenon attributed to bimolecular leaflet or micelle formation (described in more detail in Refs. 13 and 14). This optimum concentration can be readily assessed by measuring the radius of a given volume of slurry after the latter has been freely compressed between parallel glass plates under a constant load for a given length of time and plotting this radius value versus the surfactant concentration. We call such plots "effectiveness curves," and the concentration at their maximum is the desired optimum surfactant concentration. At this effective concentration, the slurry has a minimum viscosity at fixed volume percent solids loading ϕ, or the highest possible solids content at a fixed viscosity.

The surfactant or dispersing agent used was asolectin, a mixture of phosphatidyl cholines, ethanolamines and inositols (Refs. 13, 14). To prepare solids with a known degree of surfactant coverage, the method of deposition was generally that of evaporation of an ether solution of surfactant of known concentration directly onto the solid particles. In addition, straightforward adsorption isotherm techniques and direct injection of surfactant solution into the slurry were also employed. In the latter case, the ether solvent was evaporated under a vacuum. The end result in all three cases was found to be identical.

Burning rates of AP/PPO slurries were assessed by extruding the slurry into standard paper soda straws (5×0.25 in. diam), capping the ends with plastic caps and firing in a Crawford bomb under a nitrogen atmosphere. This procedure has been used for quality control of production propellant and works well, although the

[1] An effective surfactant is here described as one which not only can change the surface characteristics, but also is being employed at the proper concentration level.

scatter is greater than for standard strands that were burned in the course of the study. However, a large number of straws were burned, and a good average value for each firing pressure was ascertained. This value is the average of a scatter amounting to ±4%.

Figs. 6 and 7 show the effectiveness curves, using PPO as the liquid phase, for 53 glass spheres and bimodal AP, respectively. Fig. 3 shows the corresponding curve for bimodal AP in USP mineral oil. The percent monolayer coverage in each figure is calculated from the measured surface areas (krypton adsorption), the known weight of surfactant adsorbed, and the assumption that the mean molecular weight of the phospholipid mixture is 770 with a cross-sectional area of 40A, and that the films are close-packed (Refs. 15, 16, 17).

The viscosity of the glass bead system was also determined and is shown in Fig. 6 to illustrate that the radius

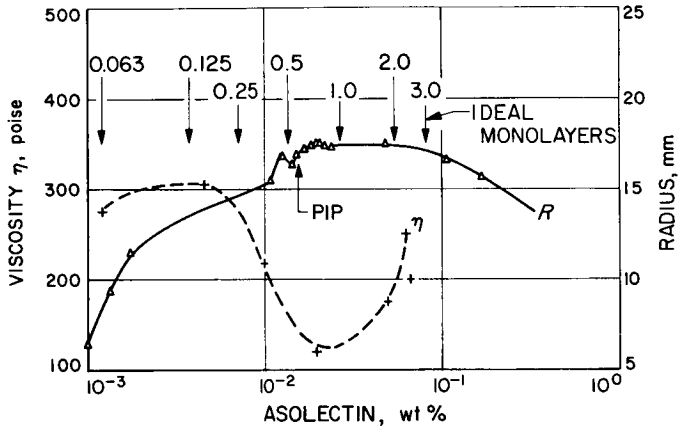


Fig. 6. Viscosity of poise versus surfactant concentration and weight percent

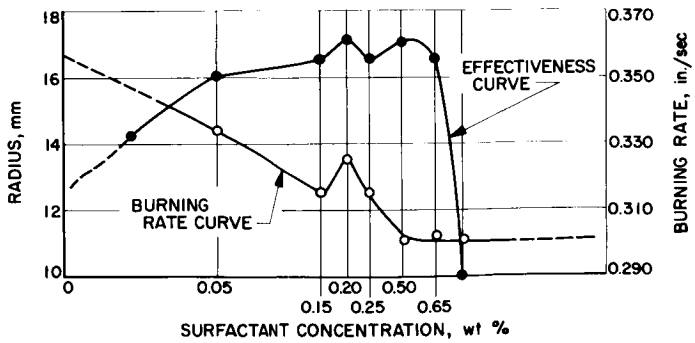


Fig. 7. Effectiveness curve of 75 wt% AP (0/100) and burning rate of 80 wt % AP (30/70) at 1000 psi

employed for the effectiveness curves is, qualitatively, inversely proportional to the viscosity. That is, the effectiveness curve maximum coincides with the viscosity curve minimum and is, therefore, related to the degree of dispersion (Refs. 13, 14).

It should be noted that the effectiveness curve for all types of solids in this media are of the same general shape, showing a secondary maximum or pip at lower concentration than the true maximum. The effectiveness curve is similar to the shape assumed by anomalous adsorption isotherms of surface active substances adsorbed onto solid surfaces.

At present, the most convincing interpretation of the shape of anomalous isotherms has been presented by Vold (Ref. 18). His interpretation is that monomer is uniformly adsorbed onto the surface for concentrations below that at which the pip in the isotherm occurs. At this point, micelles begin to form in the solution, and hence the name "critical micelle concentration" or CMC. Since the adsorbed layer is now thermodynamically unstable, it partially desorbs, with an attendant change in slurry viscosity. At still higher concentrations, there is a rearrangement of the adsorbed molecules into "surface micelles" in the concentration range where the effectiveness maximum occurs. Note that the AP mineral oil curve shown in Fig. 8 does not turn down toward the initial value after reaching a maximum, but remains at equilibrium at a constant viscosity. It was determined that the asymptotic drop in the poly(propylene oxide) system was due to jelling of the medium by the surfactant after the AP surface had saturated and hence the viscosity change was not due to decreased dispersion of the particles as it might at first appear.

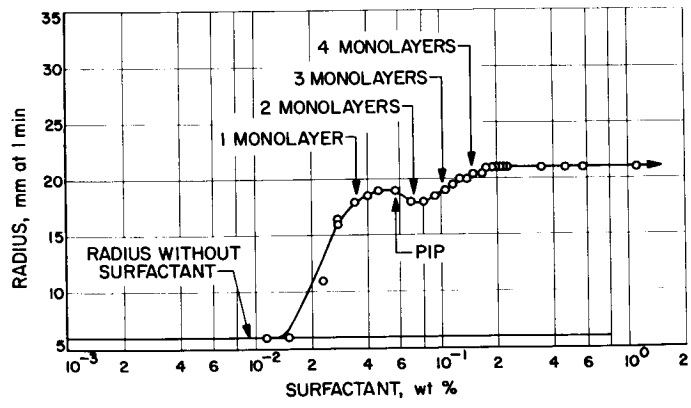


Fig. 8. Bimodal AP 30%-12μ and 70%-100μ with 2% aluminum in mineral oil (82 wt% solids)

If the degree of dispersion is contributing to the burning rate, then a plot of burning rate vs. surfactant concentration should resemble an effectiveness curve. Fig. 7 (AP in PPO, at 1000 psi) shows that this is indeed the case. The shape is repeated even to the inclusion of the pip, up to the optimum surfactant concentration of about 0.35 wt%. At this concentration, the burning rate reaches an equilibrium and does not return asymptotically toward the initial value as does the viscosity.

Fig. 9 depicts burning rate versus surfactant concentration for 82 wt% AP in a USP mineral oil. The change in the burning rate with surfactant concentration in this system, when compared with the PPO system, is quite startling. It is not too surprising, however, when the dispersing effectiveness of the surfactant in the two different media is considered. It will be shown below that PPO is a weak surfactant and that the particle-particle structure is considerably less in PPO than in mineral oil. Since the particles are partially dispersed to begin with, even in the absence of surfactant, the net result of the added surfactant effectiveness is diminished, and the attendant burning rate will not be decreased as much in the PPO system as in the mineral oil system.

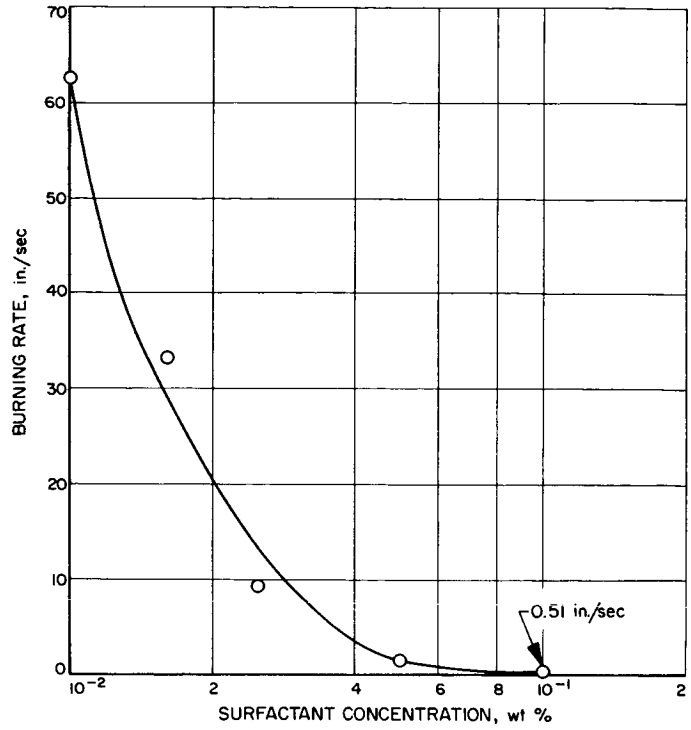


Fig. 9. Burning rate versus surfactant concentration for bimodal AP in USP mineral oil (82 wt% solids)

3. *Evaluation of Media Effects on Dispersion*

It has been shown earlier (Refs. 13 and 14) that the apparent relative viscosity of concentrated slurries as a function of the volumetric loading ϕ, is described by

$$\eta_r = \left(1 - \frac{\phi}{\phi_m}\right)^{-2.5} \qquad (1)$$

where η_r is the ratio of the viscosity of the slurry η, to that of the liquid phase η_0, ϕ is the volumetric loading of the slurry under consideration, and ϕ_m is the maximum possible volumetric loading. The latter may be determined independently in a sedimentation experiment and so is also called the sedimentation volume. From Eq. (1) it is obvious that ϕ_m is the controlling parameter of the viscosity of a slurry (degree of dispersion or particle-particle structuring). Concurrent studies (Ref. 19) have shown that ϕ_m can be described by

$$\phi_m = 0.37\left[1 + 0.707 \exp\left(\frac{-\beta S}{d}\right)\right] \qquad (2)$$

where S is a measure of the attractive force between particles, d is the particle diameter, β is a proportionality constant. The quantity βS is called the stickiness of the particles. Fig. 10 is a generalized plot for a stickiness of $\beta S = 1$. Quite obviously, with decreasing particle size

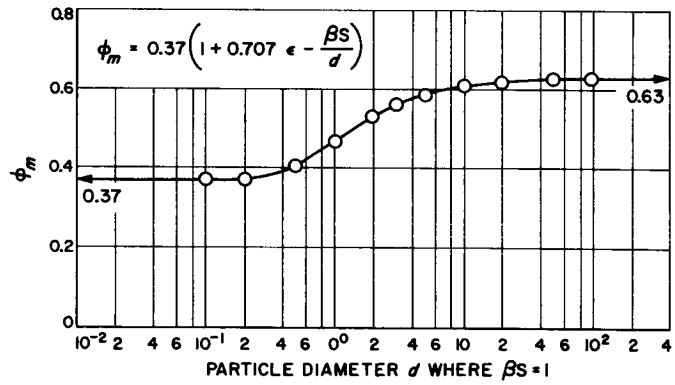


Fig. 10. Maximum volumetric loading with particle diameter d where $\beta S = 1$

through the transition region, where ϕ_m is decreasing from 0.63 to 0.37, particle-particle structure increases with attendant increased slurry viscosity and burning rate. This increase is of course observed in composite solid propellants.

Eq. (2) would say that there are no media effects on the dispersion of particles—it should be independent of the medium. However, as mentioned above, this is not strictly true—the sedimentation volume ϕ_m was measured for

Md 105 aluminum powder (Metals Disintegrating Co. aluminum, mean weight diameter 11.5μ) slurried in a variety of liquids. A summary of these determinations is presented in Fig. 11. There appear to be two classes in the liquids studied.

The first class, including such materials as acetone, benzene, chloroform, ether, freon 113, *n*-hexane, isopentane, and USP mineral oil, had no effect on sedimentation volume. That is, ϕ_m in these liquids was identical with that measured in air. On the other hand, the second-class liquids, including methanol, ethanol, octanol, decanol, and PPO (the backbone of the binder in many polyurethane composite propellants) were effective in changing ϕ_m. The mineral oil and the PPO are the same liquids used in the burning rate studies.

Comparison of the two classes of liquids shows that the primary difference is that the second substances all contain hydroxyl groups, while those of the first class are inert in this sense. This suggested that the Class II substances react with the particle surface in some way and should therefore be surface-active agents. As a test, an Md 105 aluminum-ether system was prepared ($\phi_m = 0.45$). Into this system small amounts of PPO, which has been shown in Fig. 11 to be a Class II substance, were introduced. The results, shown in Fig. 12, tell two important

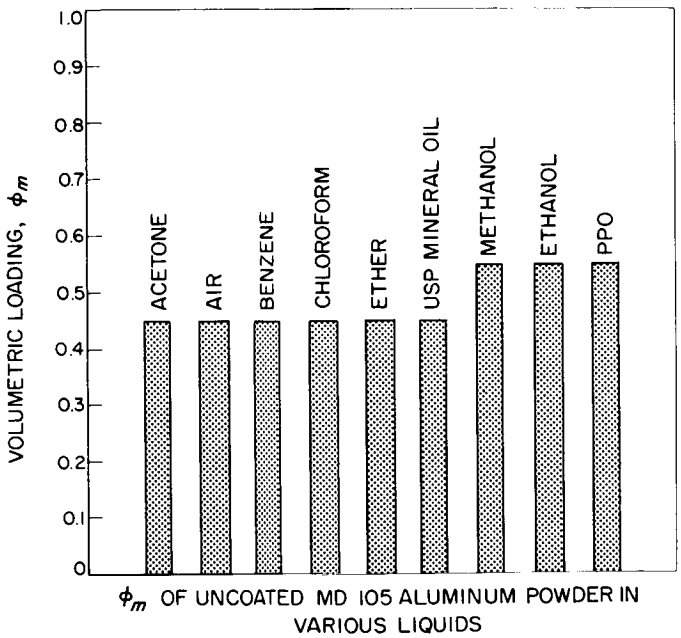


Fig. 11. Volumetric loading versus maximum possible volumetric loading of uncoated MD 105 aluminum powder in various liquids

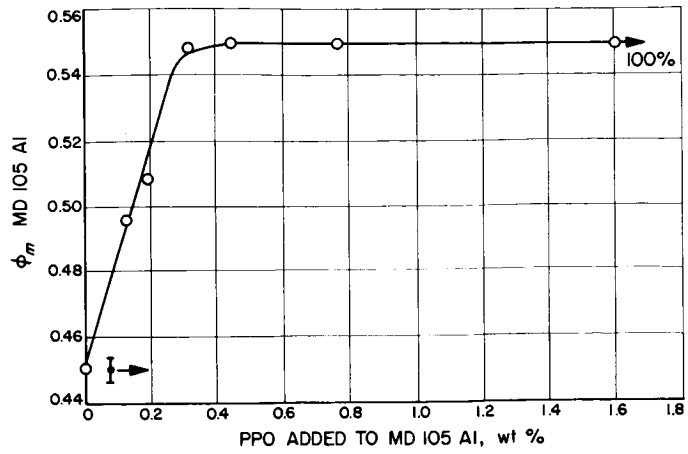


Fig. 12. ϕ_m MD 105 Al versus PPO added to MD 105 Al, weight percent

things. First, PPO is a weakly effective surfactant (not as effective as asolectin and others). Secondly, after more than a few tenths of 1% of PPO are present, ϕ_m is independent of the PPO concentration. This would confirm that ϕ_m (for a given particle system) is controlled by the surface condition, i.e., energy, of the particle only and is independent of the medium between the particles, if the surface of the particles is in a lower energy state than that of the continuous phase. That is, ϕ_m is controlled by particle-particle interaction and not by particle-medium interaction, unless the media is also an effective surfactant. Direct measurements of heats of adsorption confirm this argument.

4. Decomposition Effects on the Burning Rate

When a surfactant or burning rate modifier such as magnesium oxide or copper chromite is used to modify the AP surface or is dry-blended in the latter two cases, there is a change in the thermal decomposition characteristics of the AP as determined by differential thermal analysis.

Fig. 13 shows differential thermal analysis curves for production run AP, such as used in solid propellant and for the same AP that has been coated with 0.2 wt% asolectin by the evaporation technique. As can be clearly seen in the curves, the predecomposition exotherm at 290°C has been shifted up in temperature approximately 30°, to 320°C. Furthermore, the amplitude of the predecomposition exotherm has been considerably reduced. The additional endotherm at 95°C is interpreted to represent volatile constituents of the surfactant phospholipid mixture. Although it is extremely difficult to separate

decomposition effects from dispersion effects, it appears that the progressively reduced burning rate of surfactant treated propellant slurry at a given solids loading is in a small part due to progressively greater amounts of surfactant adsorbed onto the AP particles with a concomitant change in the predecomposition behavior and, of course, in degree of dispersion. At the pip or CMC there is a partial desorption, and consequently a slight increase in burning rate, until the surface micelles form. Further increase of the surfactant concentration results in additional micelle or other structural formation in the media with no further adsorption onto the particle surfaces. With no further change in surface characteristics, the predecomposition behavior, and thus the burning rate, become independent of the surfactant concentration.

Figs. 14 and 15 are DTA curves of magnesium oxide and copper chromite blended AP, respectively. Note that in both cases the major decomposition peak has been shifted down temperature, and that the amount of shift in this particular case would predict the amount of change in the burning rate. That is, the burning rate of the MgO-treated slurry was increased to a greater degree than that treated with copper chromite (at 1000 psi burning rate, increased from approximately 0.32 in./sec to 0.48 in./sec for magnesium oxide and from 0.32 in./sec to approximately 0.46 in./sec for copper chromite).

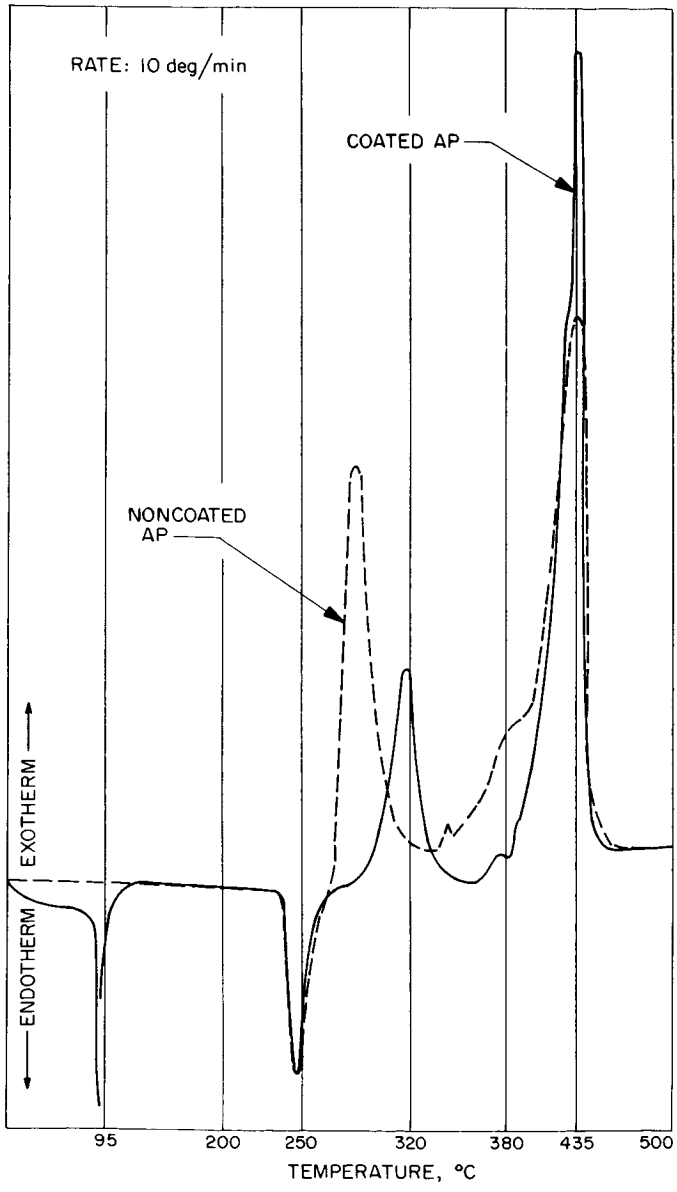


Fig. 13. The differential thermal analysis of AP

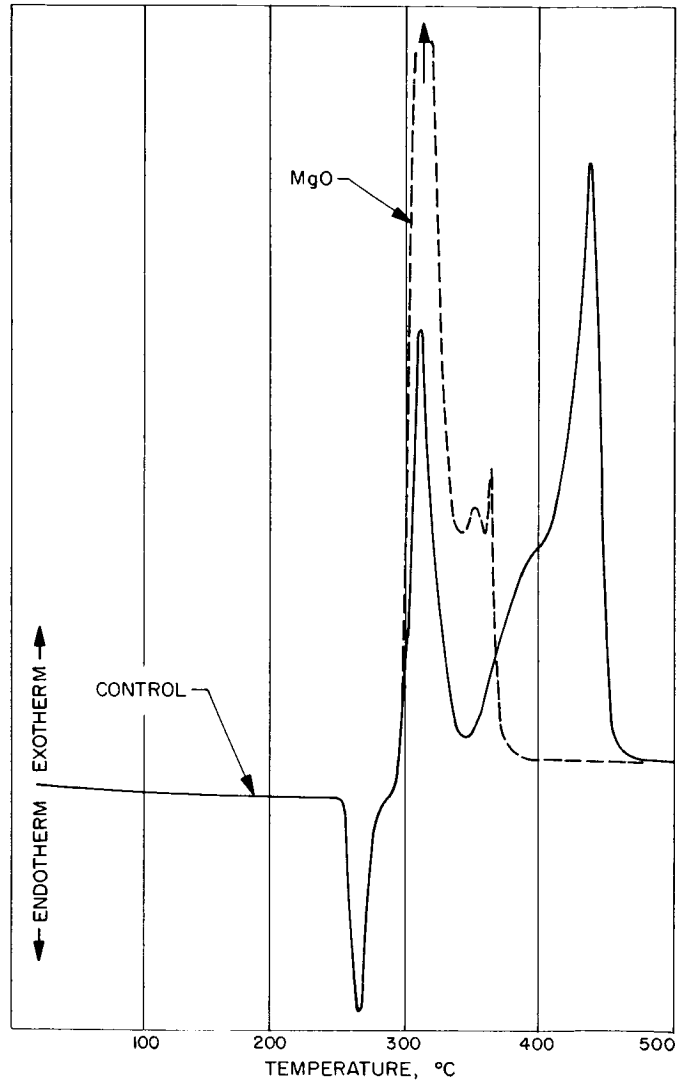


Fig. 14. DTA thermogram of AP surface modified with magnesium oxide by dry mixing

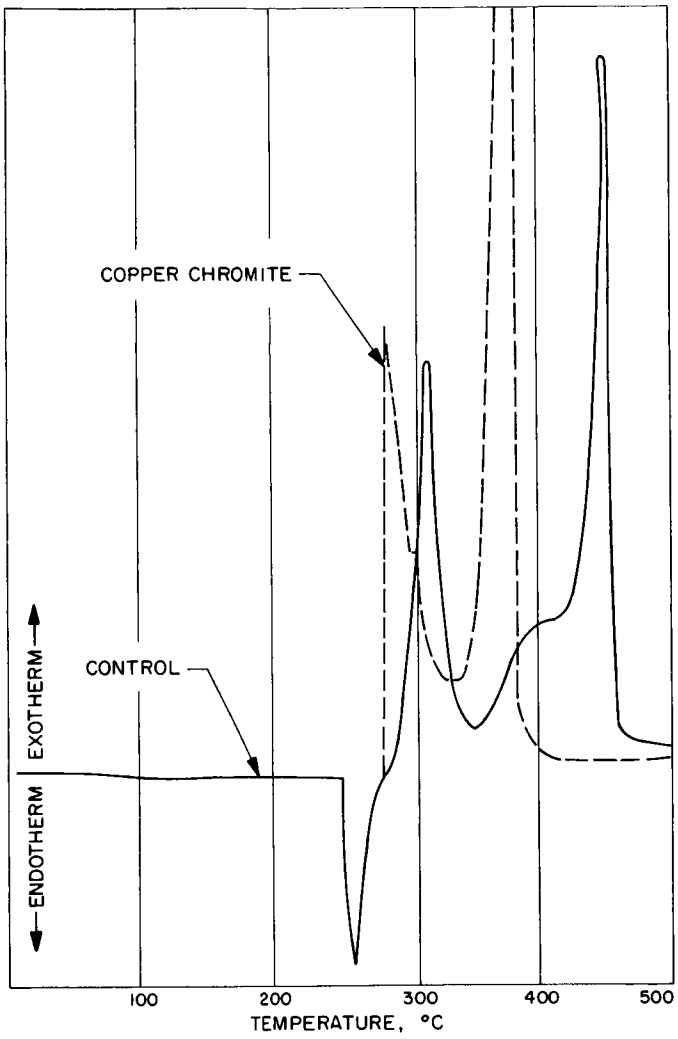


Fig. 15. The DTA thermogram of regular production of AP

In the case of magnesium oxide and copper chromite modified propellants increased or decreased dispersion play no part in modifying the burning rates. Sedimentation volume and viscosity measurements on these modified slurries showed that there was no difference in the rheological properties between the control and the magnesium oxide or copper chromite-treated slurries.

5. Conclusions

(1) The effects of surface modification on propellant oxidizer is twofold. First, there is a change in the rheological properties. Secondly, there is an effect on the burning rate.

(2) When a surfactant is used to modify the oxidizer surface, the change in the burning rate may be large. This change is primarily due to changed degree of dispersion, but second-order effects may arise from changed thermal decomposition characteristics of the oxidizer (AP).

(3) When burning rate accelerators such as magnesium oxide and copper chromite are blended onto the oxidizer, there appears to be an effect on the thermal decomposition in the opposite direction from that observed with an effective surfactant. This change in decomposition qualitatively appears to be of the same magnitude as changes in the burning rate of the propellant. There is no change in degree of dispersion.

C. Polymer Degradation Mechanisms: C^{14}-Labeled Isocyanate Adducts of Poly (Propylene Oxide)

E. F. Kopka and J. D. Ingham

1. Introduction

Since relatively large changes in mechanical properties can result from small chemical changes that occur in the early stages of polymer degradation, sensitive means of observing chemical reactions that produce bond scission are required. One method is to include C^{14} in the polymer chains at links that will be expected to be involved in the degradation reactions and then to monitor the formation of a C^{14}-labeled product. For poly(propylene oxide)-2,4-tolylene diisocyanate (PPO-TDI) urethanes it has been established that the urethane linkages undergo some scissions to evolve carbon dioxide (*SPS 37-25* and *37-23*, Vol. IV). As a means of isolating reactions at the 2- and 4-positions of TDI-urethanes and to correlate small extents of bond scission with $C^{14}O_2$ evolution, studies of the thermal degradation at moderate temperatures ($\leqq$150°C) of PPO reacted with C^{14}-labeled *p*-tolyl isocyanate (PTI) were initiated and are presently in progress.

2. Discussion of Results

Some initial measurements of C^{14} evolution from PPO-PTI at the refluxing temperatures of toluene (109°C), isoamyl alcohol (128°C) and 2-heptanone (150°C) have

been completed. Two limitations of the method are that radioactive carbon can be evolved in at least two forms, either as $C^{14}O_2$ or as *p*-tolyl isocyanate, and that any *p*-tolyl isocyanate evolved may have originated either from unreacted isocyanate or from urethane scission reactions. In the experimental procedure used here, the *total* radioactivity from $C^{14}O_2$ and *p*-tolyl isocyanate was obtained, so that incomplete urethane formation and degradation are not readily distinguishable initially. However, the data given in Table 2 indicate the conditions for nearly complete reaction. It should be noted that after a reaction time of greater than 24 hr at 150°C and greater than 48 hr at 128°C, the incremental fraction of C^{14} evolved exceeds that for the reaction at 109°C; therefore, complete curing must require a minimum of 24 hr at 150°C and 48 hr at 128°C. From results given in *SPS 37-25*, Vol. IV, the greatest chain extension was obtained at a temperature near 75°C and below 110°C for PPO-TDI, and at the latter temperature some bond scission was observed. Thus, qualitative information about the reaction time-temperature relationship for urethane formation with PPO is available, and a reasonably optimum reaction schedule can be followed. Therefore, in future experiments, after mixing the PPO and isocyanate, the reactants will be heated for 72 hr at 105°C to promote complete reaction and minimize the presence and subsequent evolution of unreacted isocyanate.

Results obtained for PPO-PTI at the three temperatures mentioned are shown in Fig. 16, in which the fraction of initial C^{14} evolved is plotted versus time. It can be seen that the C^{14} evolved at 109°C in less than 50 hr exceeds the amounts at 128 and 150°C. This is apparently due to the fact that more unreacted isocyanate is present and therefore evolved initially at 109°C because of the longer cure cycles required at lower temperatures. One other unexpected feature of the curves is that the linear rates at times greater than 500 hr are essentially the same at either 128 or 150°C. Since only one experiment was run at each temperature, further tests are being conducted to determine whether the correspondence is real or whether it was caused by some inadvertent artifact. The linear rates are tabulated in Table 3 along with data

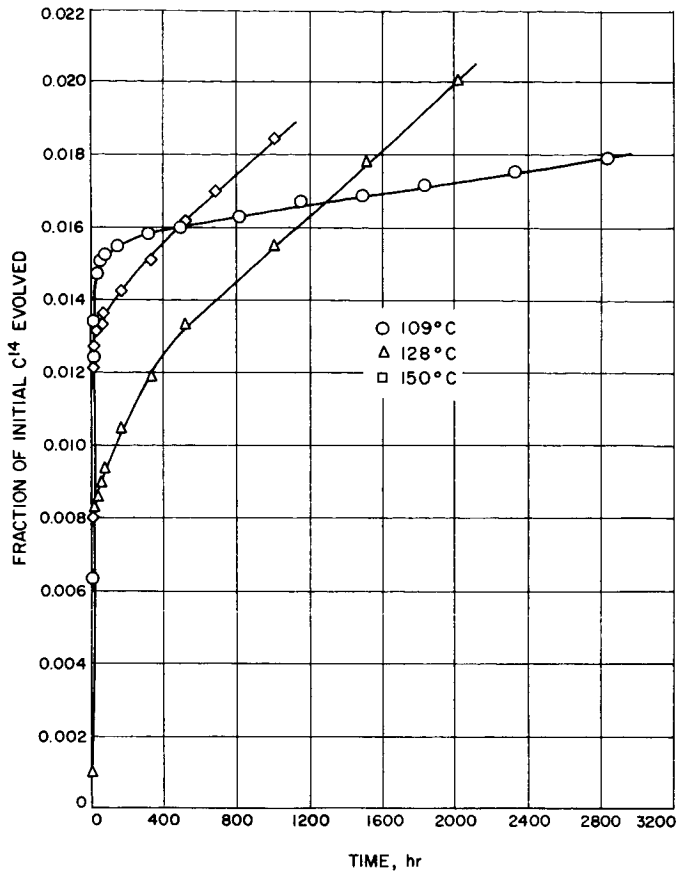


Fig. 16. Fraction of C^{14} evolved from poly(propylene oxide)–*p*-tolyl isocyanate urethane versus time

Table 2. Evolution of C^{14} from a poly(propylene oxide)–*p*-tolyl isocyanate adduct

Reaction time, hr	Incremental fraction of C^{14} evolved at 109°C	Incremental fraction of C^{14} evolved at 128°C	Incremental fraction of C^{14} evolved at 150°C
0.5	0.0262	0.0334	0.00397
6	0.0252	0.0169	0.0293
12	0.00388 >	0.00230 >	0.000986
24	0.00458 >	0.00169 >	0.000982
48	0.00152 >	0.000775 <	0.00157
72	0.000659 <	0.000895 <	0.00148
168	0.000300[a] <	0.00213 <	0.00362

[a] Not measured; this figure is an interpolated estimate from Fig. 16.

Table 3. Linear rates of the fraction of C^{14} evolved from the poly(propylene oxide)–*p*-tolyl isocyanate adduct (PPO-PTI) and PPO-TDI polyurethane

Product	Temperature, °C	Rate of C^{14} evolution, %/hr × 10^5
PPO-PTI	109	8.4
PPO-PTI	128	44.5
PPO-PTI	150	46.0
PPO-TDI	110	8.8
PPO-TDI	150	51.6

obtained for PPO-TDI at 110 and 150°C in earlier work (*SPS 37-25*, Vol. IV). The remarkable agreement at these temperatures appears significant. Although a high degree of reliability cannot be ascribed to the data without further work, there is an indication that for PPO-TDI, the urethane bond *para* to the methyl group undergoes scission to produce CO_2 and that the rate of scission for the *ortho*-urethane bond is nearly negligible at these temperatures. If true, these observations will greatly simplify the interpretation of results of degradative studies at moderate temperatures of polyurethanes containing 2,4-tolylene diisocyanate.

An Arrhenius plot of the data is shown in Fig. 17 which gives a relatively low nominal value for the activation energy of 13.7 kcal/mole for the evolution of C^{14} from either PPO-TDI or PPO-PTI at 110 to 150°C. On the basis of the 128°C data, the activation energy would be 26.3 kcal/mole; however, the rate data at this temperature may not have been extended to sufficiently long times to achieve a truly linear rate.

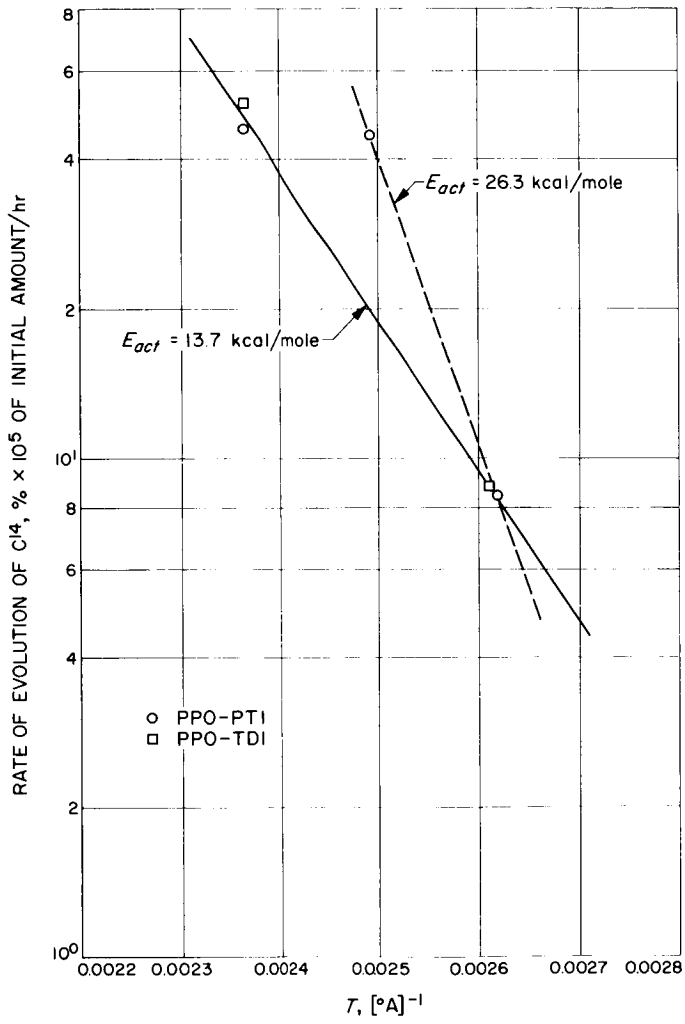


Fig. 17. Activation energy plot for C^{14} evolution rates for poly(propylene oxide)–*p*-tolyl isocyanate adduct and PPO-TDI polyurethane

Infrared and intrinsic viscosity data were also obtained for the PPO-PTI adduct. Although they are not presented here, these results indicated gradual elimination of free isocyanate and showed no change in molecular weight of the residual polymer, and therefore, neither polyether scission nor chain extension.

3. Experimental

The poly(propylene oxide)–*p*-tolyl isocyanate adduct was prepared in a dry box by weighing out 30.6 g of degassed PPO (mol wt 2040) and adding 4.0 ml (4.2 g) of C^{14}-labeled *p*-tolyl isocyanate with an activity of 0.0036 mC/cm^3. The NCO/OH ratio was 1.05. The PPO was stirred for 30 min after addition of the isocyanate, and 11.0 g of the product was added to each of three tubes fitted with a dispersion tube and inlet and outlet stopcocks.

The degradation experiments were performed by immersing the polymer tubes in each of three oil baths heated by refluxing toluene (109°C), isoamyl alcohol (128°C) and 2-heptanone (150°C). During the experiments volatiles containing C^{14} were trapped at intervals by passing dry nitrogen through the polymer tube and out through a syringe needle immersed in a mixture of 1.0 ml of methanol, 1.0 ml of phenethylamine, and 15 ml of scintillator solution that was cooled with an ice-bath. Ordinarily, 20 min was allowed for flushing radioactive products into the vial. When a polymer sample was required for infrared, intrinsic viscosity, or radioactivity determinations, it was removed, after flushing the tube by increasing the nitrogen flow rate, by raising the top of the polymer tube and inserting a weighed Pasteur pipet with which the sample was withdrawn.

D. Structural Integrity of Solid Retrorockets

A. San Miguel

A study program is under way to evaluate the structural integrity of the *Surveyor* spacecraft solid propellant

retropropulsion motor within the environment to which it will be subjected in fabrication and flight use. This program consists of four tasks. Task 1 is to construct four ¼-scale *Surveyor* retromotors. Two of these retromotors will be cast with *Surveyor* propellant at JPL using a proper *Surveyor* formulation (JPL-EDI) and casting schedule. The other two cases will be cast with a photoelastic stress-freezing epoxy. All four retromotors will be appropriately instrumented. Task 2 is to employ a linear viscoelastic analysis which includes temperature-time effects. This complex analysis is being undertaken by the Douglas Co. under Contract NAS7-392. Task 3 is to measure all the temperature-time mechanical properties of the *Surveyor* propellant for use in the viscoelastic theory. Task 4 will be an experimental evaluation of the stress and temperature state of the *Surveyor* retromotor by direct measurement, employing miniature stress rosettes, thermocouples, strain rosettes, and other necessary instrumentation.

1. Task 1 Progress

Six ¼-scale rocket cases have been constructed (Fig. 18) having the dimensions shown in Fig. 19. The manufacture of the cases consisted of hydroforming stainless steel, welding, heat-treating, and machining. A seven-point star mandrel and associated casting hardware were also machined. The seven-point bore configuration was a primary candidate over the eight-point bore configuration at the initiation date of this program.

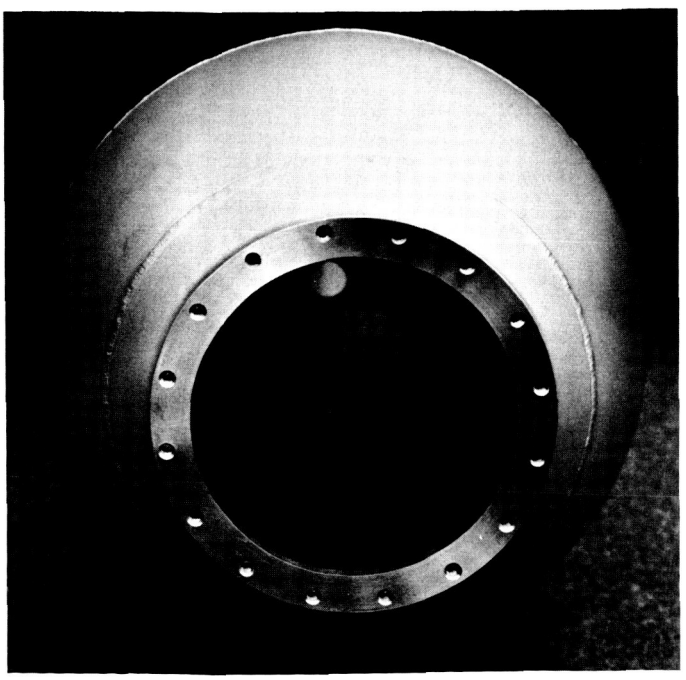


Fig. 18. *Surveyor* retrorocket casing (1/4-scale)

2. Task 2 Progress

The structural analysis of the ¼-scale retrorocket grain is being performed by the Douglas Aircraft Co.[2] The Douglas structural analysis technique is based on the Redundant Force Method. This method essentially idealizes the propellant grain and flexible case by means of a three-dimensional truss network. A linear elastic solution is obtained for this three-dimensional network by satisfying mixed boundary conditions together with compatibility conditions. The linear viscoelastic solution is obtained by replacing Young's modulus with the stress-relaxation modulus (Ref. 20). Responses are obtained employing the superposition principle.

It is noted that two basic methods of analysis prevail in the propellant community. These two are the force and displacement methods (Ref. 21). In the force method the redundant forces are taken as the unknowns of the problem, so that all the internal forces in the truss can be expressed in terms of the external and the redundant forces. Thus, by using the stress-strain relationship, the deformations of the elements can also be expressed in terms of the external and redundant forces.

Finally, by applying the compatibility criterion that the deformed elements must fit together, it is possible to formulate a set of linear simultaneous equations which yield the values of the redundant forces. The displacement method is similar, except that the equilibrium criterion at each point replaces the compatibility criterion.

A preliminary structural analysis was performed employing mechanical property data for TPH-1076 propellant, which was assumed to have a mechanical response similar to TPH-3062. This preliminary analysis was required to establish critical stress regions to guide the experimental phase of this program.

a. Thermal analysis. The thermal load analysis is predicated upon a preliminary heat-transfer analysis. The preliminary heat-transfer solution for rocket soakings to 20, 50 and 80°F from a cure temperature of 135°F is shown in Fig. 20. These heat-transfer solutions will eventually be evaluated by embedded thermocouples in the ¼-scale rocket.

[2]Figs. 20–27 were drafted from data supplied by Douglas Aircraft Co.

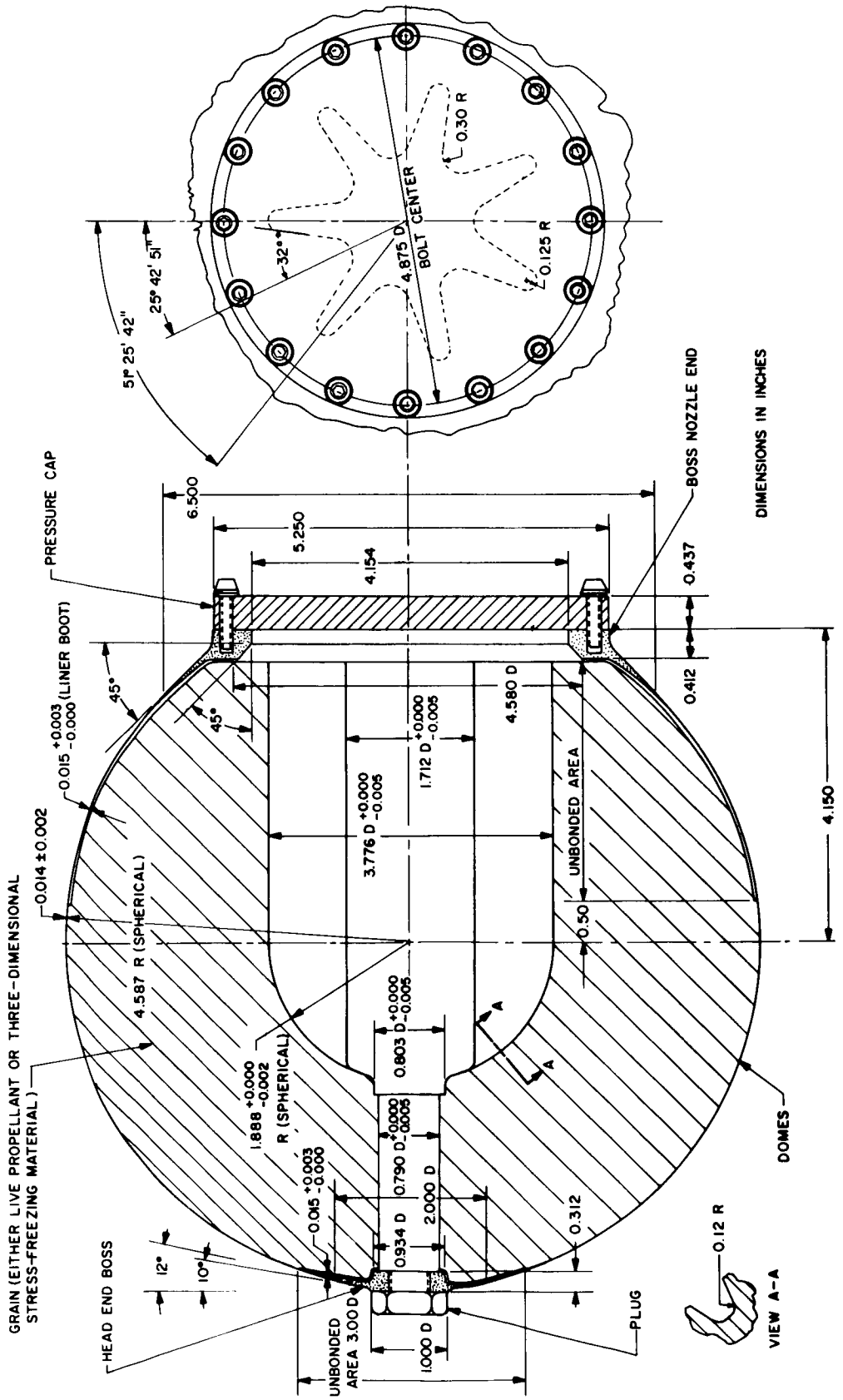


Fig. 19. Dimensions of 1/4-scale *Surveyor* retrorocket casing and mandrel

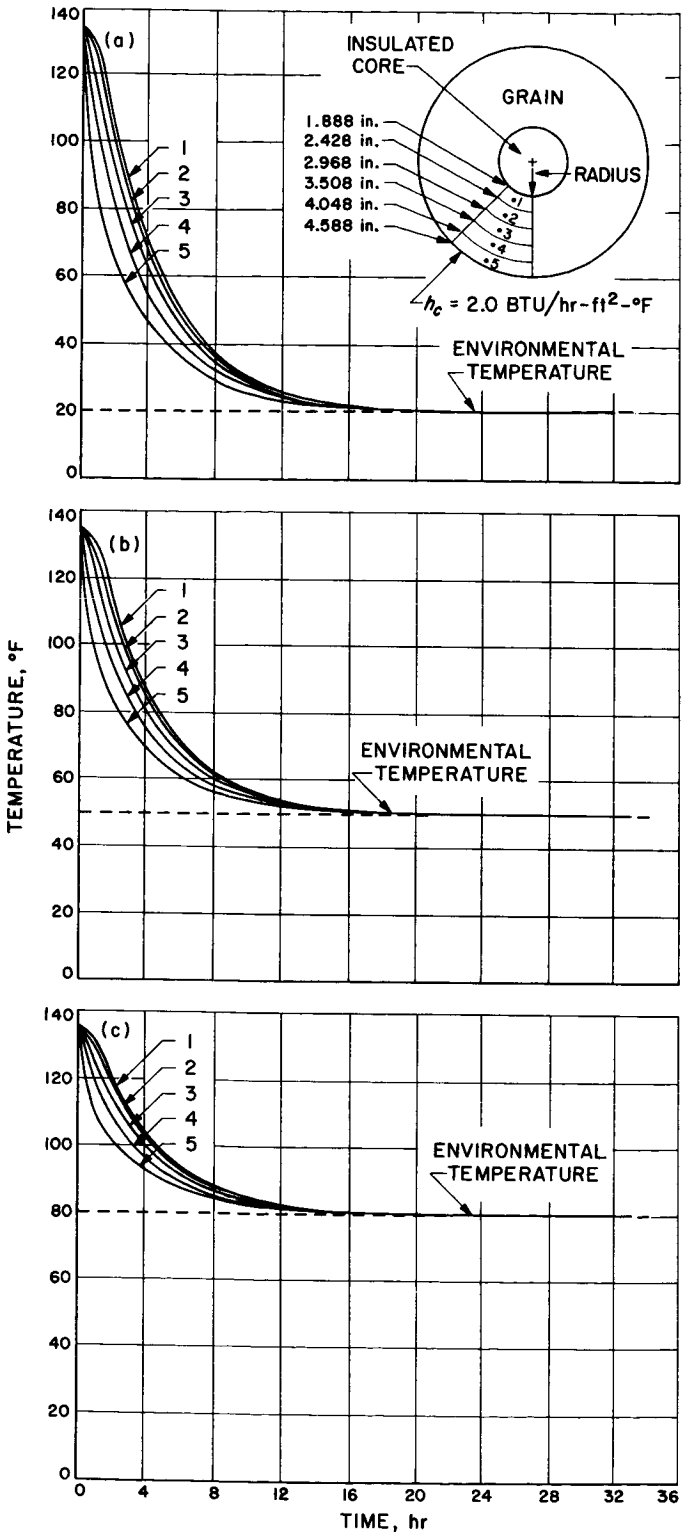


Fig. 20. Preliminary temperature-time distributions for 1/4-scale *Surveyor* retrorocket for environmental temperatures of (a) 20°F, (b) 50°F, and (c) 80°F

A finite-difference technique employing the method of successive overrelaxation is used to obtain the elastic stresses consequential to a 1% uniform shrinkage of the propellant with a Poisson's ratio equal to 0.45 (Fig. 21). The extrapolated stress values for Poisson's ratio equal to one-half are presented in Fig. 22. The bore-hoop strains obtained for 1% uniform shrinkage and Poisson's ratio equal to 0.45 are shown in Fig. 23. The corresponding extrapolated values for Poisson's ratio equal to one-half are presented in Fig. 24. Fig. 22 will be used to establish the locations where several miniature stress rosettes (Ref. 22) will be embedded in the ¼-scale rocket.

b. Pressurization analysis. The elastic stress distribution for an internal pressure of 600 psi is shown in Fig. 25. A viscoelastic analysis will eventually be performed where the ramp pressurization will be 50, 250, and 500 msec. These values were chosen because of convenience to the experimental phase of this study.

3. Task 3 Progress

The temperature-time mechanical properties of TPH-3062 are not available. Hence a formulation (JPL-EDI) and casting schedule believed to be the same as TPH-3062 is being examined. The following basic data have been obtained and will be reported in final form in the near future: (1) broad spectrum tensile properties, (2) stress relaxation modulus, (3) creep compliance, (4) bulk modulus, (5) specific heat, (6) thermal conductivity, (6) coefficient of thermal expansion, (7) glassy transition temperature, (8) density, and (9) cure shrinkage.

4. Task 4 Progress

The feasibility of carrying out the experimental evaluation of the stress and temperature state of the ¼-scale retrorocket has been demonstrated (Ref. 22). Fig. 26 illustrates exterior transducer mounting techniques which experimentally proved that the transducer positioning, using 3-mil mounting wire, can readily survive abusive handling and casting. The casing shown is that of a used *Syncom* motor.

An embedded stress rosette is shown in Fig. 27. Methods have been developed to remove mounting wires after curing is completed. These rosettes will be embedded in strategic positions within the grain, as suggested by Figs. 22 and 24.

Preliminary three-dimensional stress-freezing experiments have been performed to evaluate the transducer

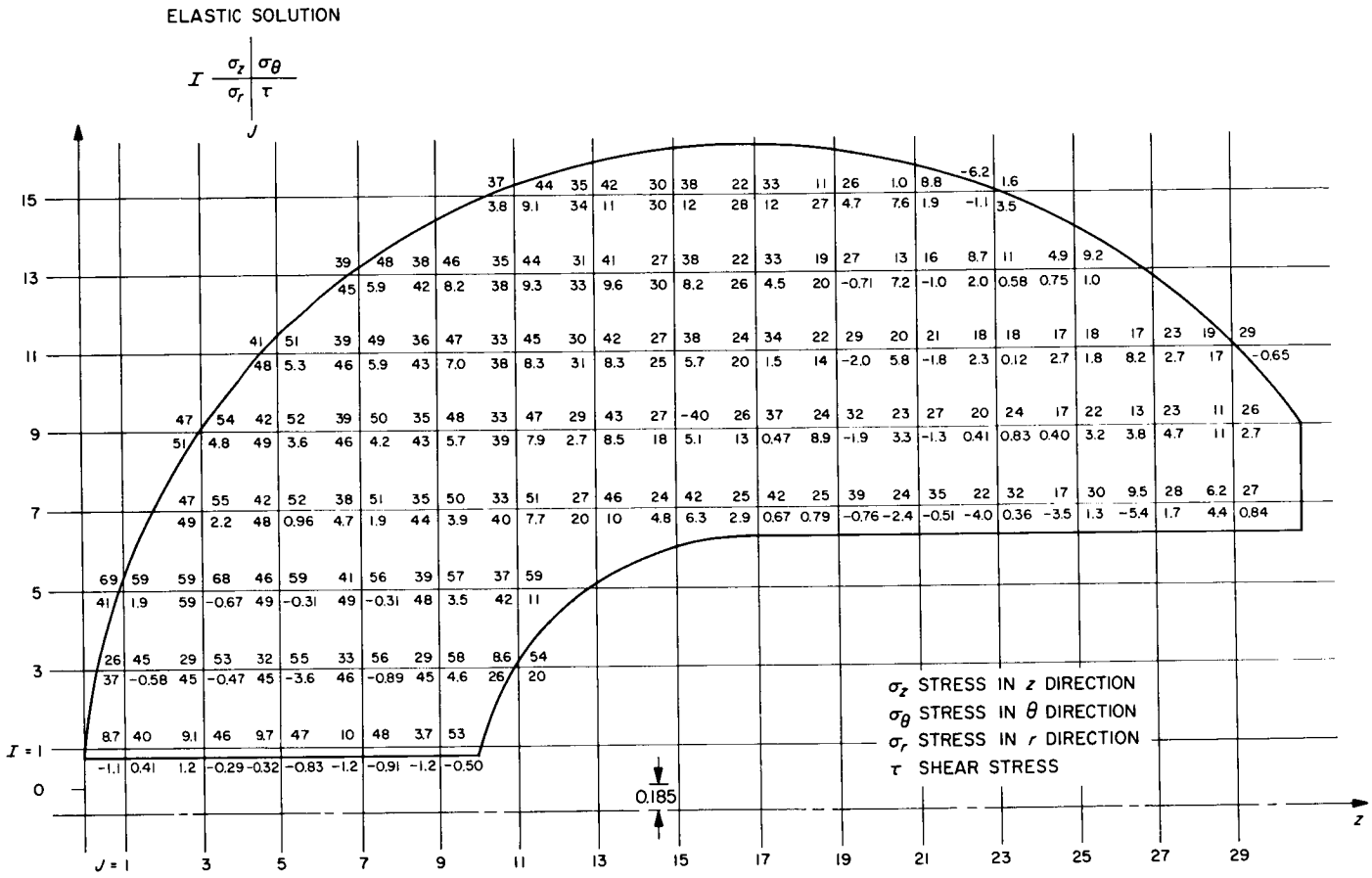


Fig. 21. Thermal loading due to 1 % uniform shrinkage and a Poisson's ratio of 0.45 (1/4 scale)

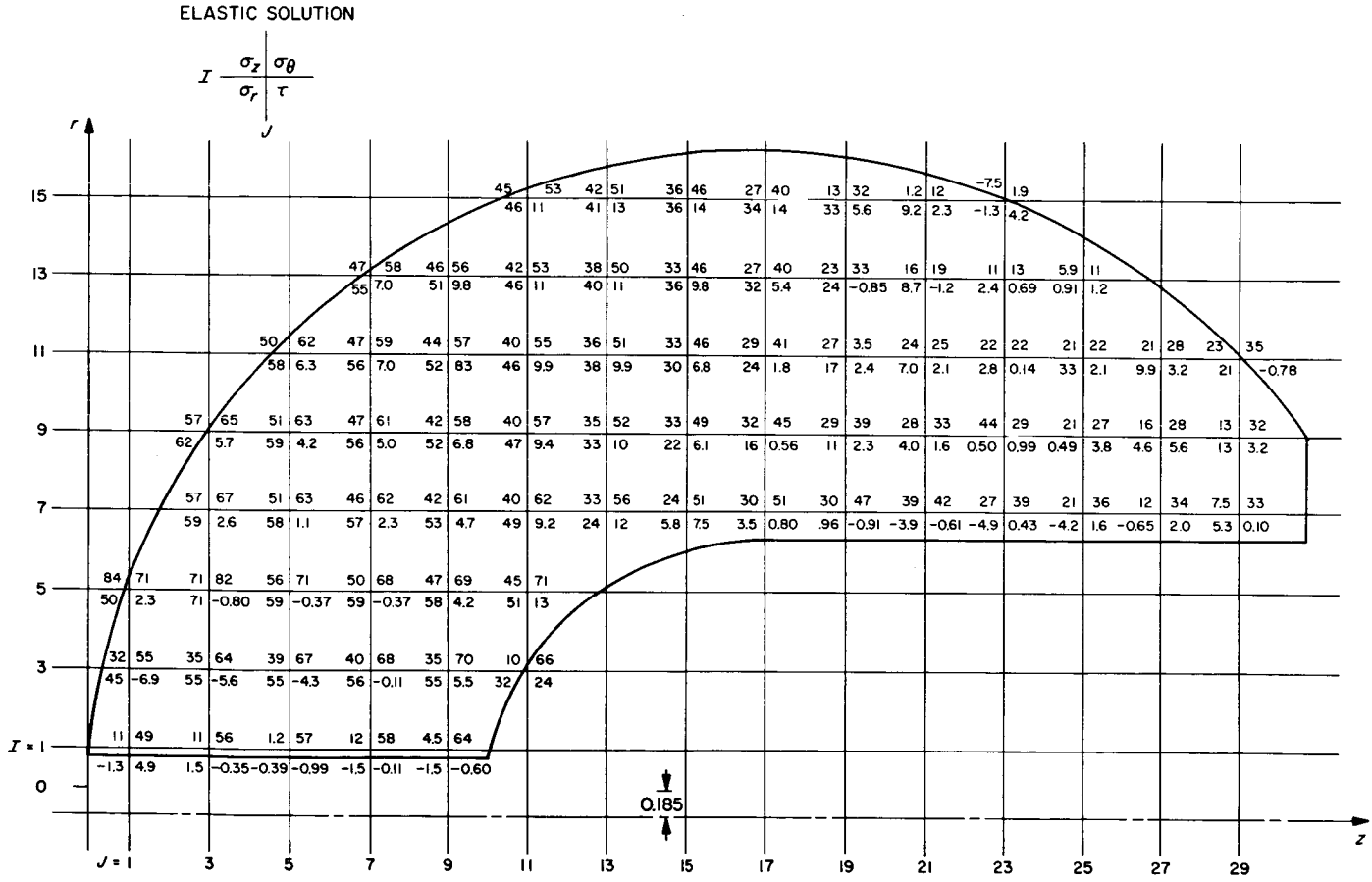


Fig. 22. Extrapolated thermal loading due to 1% uniform shrinkage and a Poisson's ratio of 0.50 (1/4 scale)

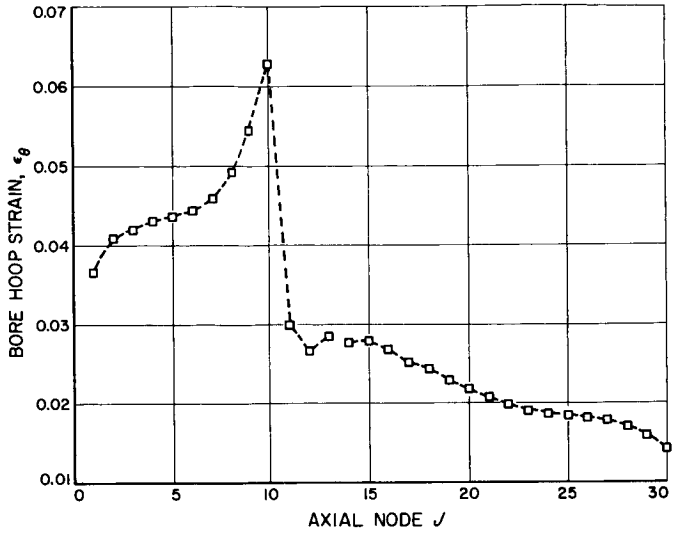


Fig. 23. Bore hoop strains obtained for 1 % uniform shrinkage and a Poisson's ratio of 0.45 (1/4 scale)

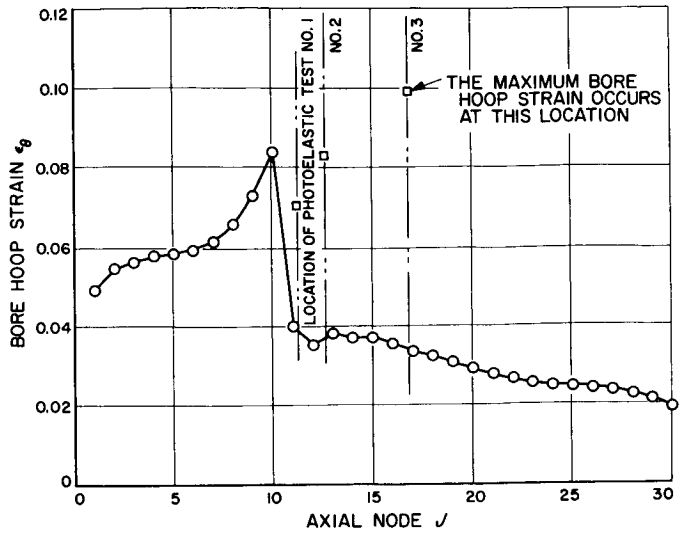


Fig. 24. Extrapolated bore hoop srtains obtained for 1 % uniform shrinkage and a Poisson's ratio of 0.50 (1/4 scale)

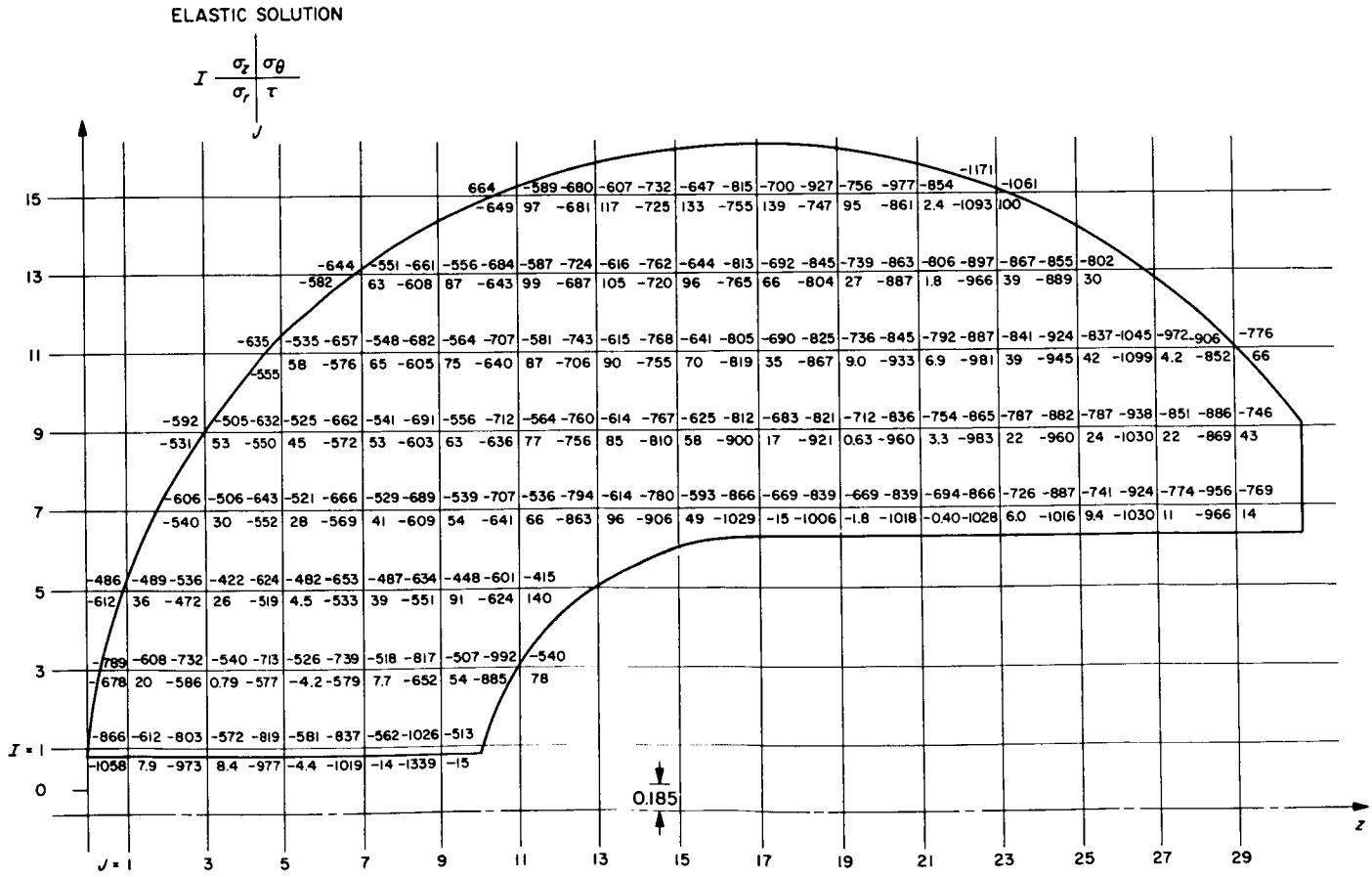


Fig. 25. Preliminary elastic stress distribution due to an internal pressure of 600 psi (1/4 scale)

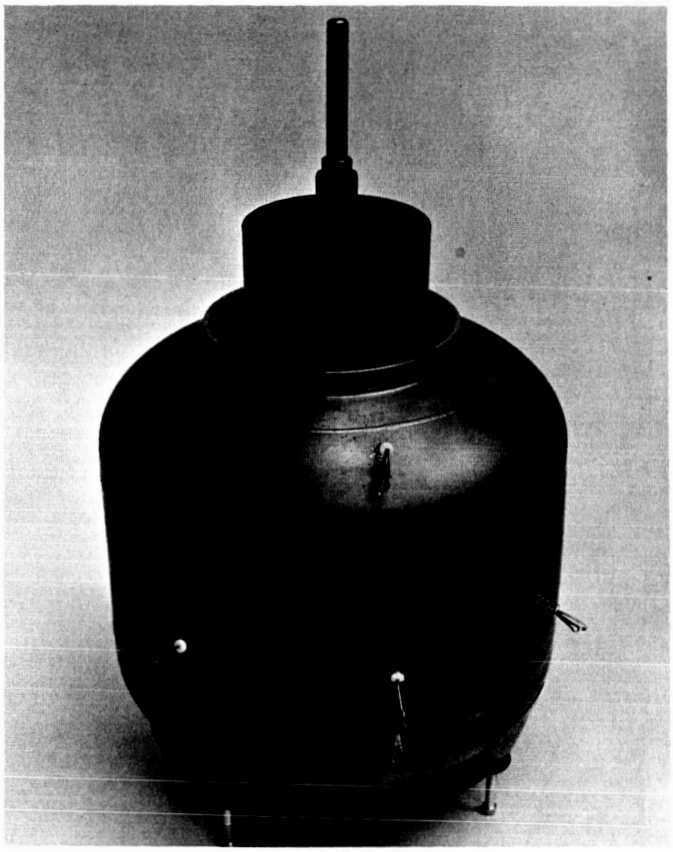


Fig. 26. Exterior transducer mounting techniques using a *Syncom* casing

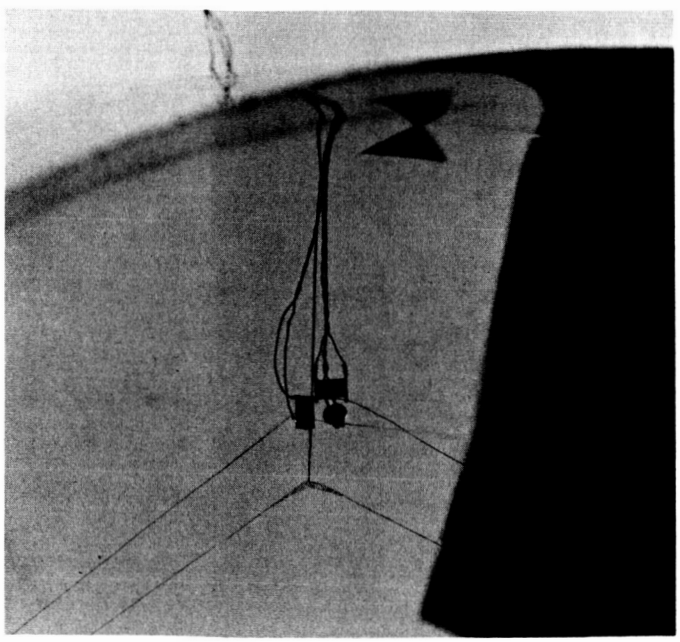


Fig. 27. An embedded stress rosette in an unfilled polyurethane resin

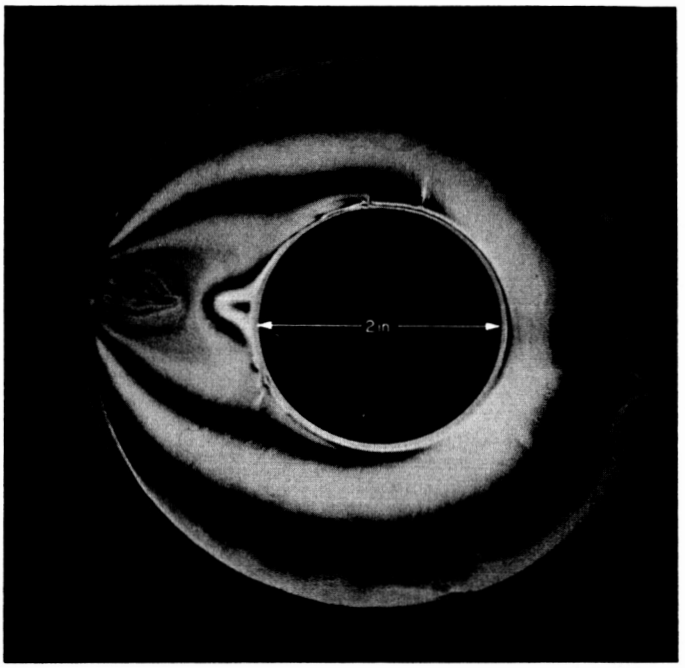


Fig. 28. Thermocouple inclusion effect

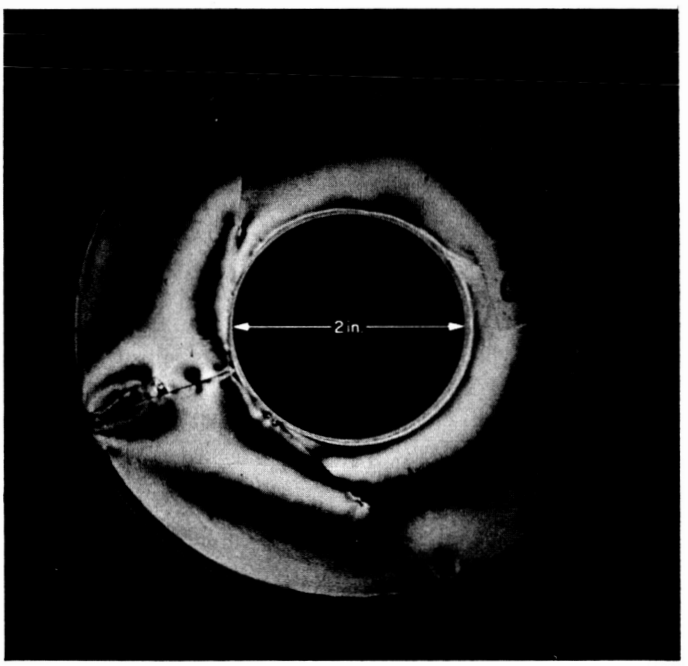


Fig. 29. Stress transducer inclusion effect

inclusion effects as well as to establish instrumentation requirements. Fig. 28 illustrates the inclusion effect of an embedded 1/16-in. thermocouple. Fig. 29 illustrates the inclusion effect of a miniature stress transducer subjected to the same environment as the thermocouple, i.e., 140°F and 600-psi internal cylinder pressure. Fig. 30 illustrates

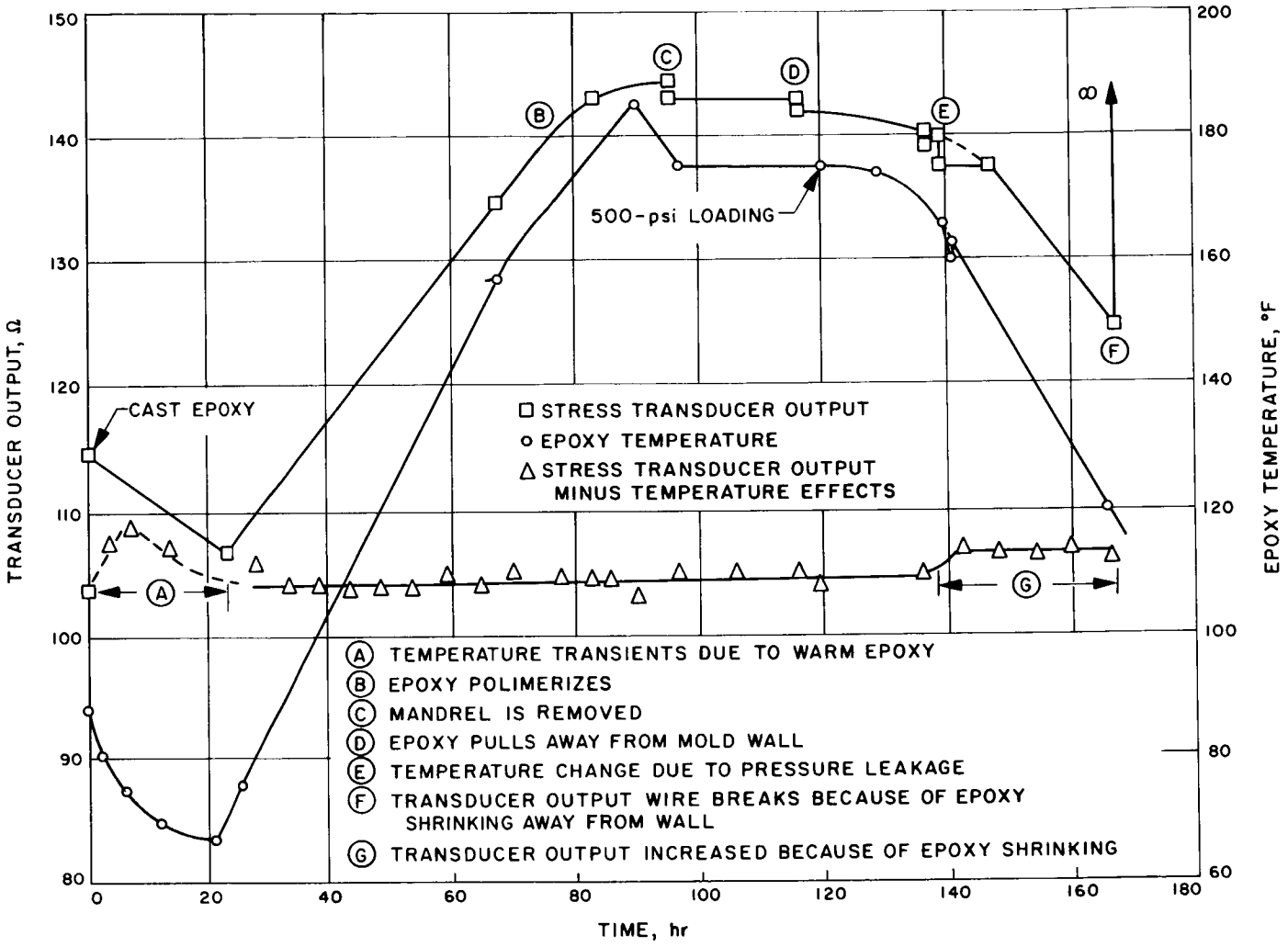


Fig. 30. Output response of embedded stress transducer and thermocouple during casting, curing, and pressurization

the output response of the miniature stress transducer and thermocouple embedded in a stress freezing epoxy. Miniature stress transducers are being constructed and calibrated for the environmental realm in which the ¼-scale retrorockets will be examined.

E. Electrical Properties of Poly-N-Vinylcarbazole-Iodine Charge Transfer Complex

A. Rembaum, D. S. Chapin[3], and L. E. Creasy

1. Introduction

Unsubstituted as well as substituted carbazole derivatives are known to form charge transfer complexes with such electron acceptors as 2,4,7-trinitro-9-fluorenone (Ref. 23) *ortho* and *para*-chloranil (Ref. 24), 2,7-dinitroanthraquinone (Ref. 25), picric acid (Ref. 26), etc. It was anticipated that similar complexes could be obtained with poly-N-vinyl-carbazole (PVCA). In fact, a variety of PVCA complexes were prepared, all of which exhibited electrical characteristics of organic semiconductors. In particular, the complex formed from PVCA and iodine was found to be of relatively low resistivity and to have the advantage of excellent moldability.

The purpose of this report is to present evidence for the formation of a PVCA-iodine charge transfer complex and to correlate some of the electrical properties of the complex with composition and temperature.

2. Experimental

a. Reagents. High molecular weight PVCA (intrinsic viscosity in benzene = 0.84 dl/g), from the Borden Chemical Company, was precipitated twice from a benzene solution with methanol and dried in a vacuum oven at 50°C. Resublimed iodine and spectroscopic grade dichloromethane (CH_2Cl_2) were used as received.

b. Complex preparation and encapsulation. A series of PVCA-iodine complexes of variable composition were prepared from their CH_2Cl_2 solutions by solvent removal at 0°C under vacuum in a rotating flask, fitted with a liquid nitrogen trap to collect traces of iodine lost during evaporation of solvent. PVCA-iodine discs (0.5 in. in diameter by ⅛ in.) were formed in a Perkin Elmer die at a pressure of 35 kbar. Encapsulation with epoxy cement was found effective in confining the iodine in the pellet and silver-filled conductive epoxy was used for electrical contacts. The encapsulation technique is shown in Fig. 31.

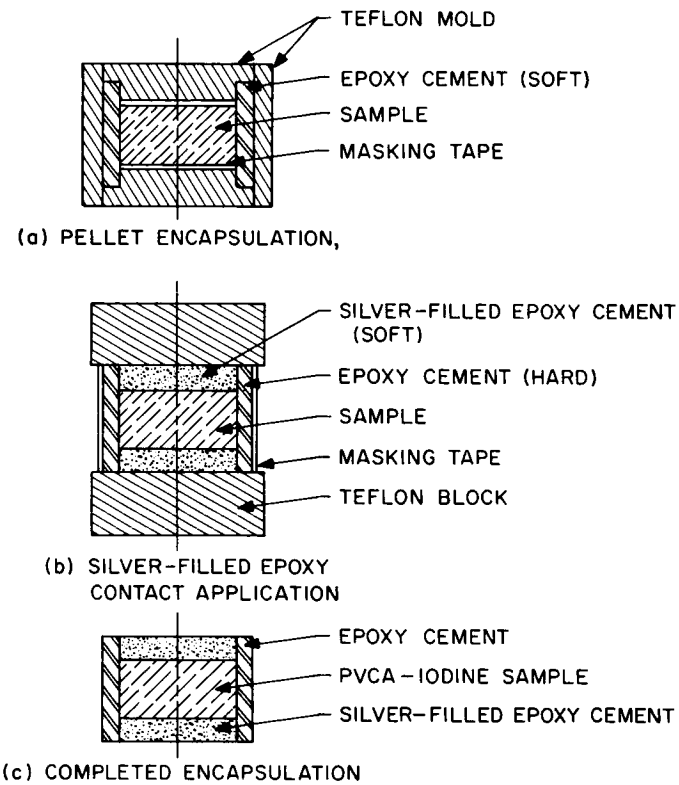


Fig. 31. Steps in the encapsulation of a PVCA-iodine pellet

c. Apparatus. Measurements of resistivity were carried out in electrostatically shielded enclosures using a Keithley 600A electrometer. The PVCA-iodine pellet in contact with a copper constantan thermocouple and held between brass electrodes was suspended in a Dewar flask, the temperature of which was adjusted by means of precooled gaseous nitrogen.

The determination of voltage current characteristics was carried out by means of the circuit of Fig. 32, using a low-frequency (10 cps) triangular wave generated by a 250 Exact Electronics Function Generator. Absorption spectra were obtained on a Cary Model 14 recording spectrophotometer against solvent in the reference cell.

[3] Associate Professor, Chemistry Department, University of Arizona, Polymer Research Section, Summer 1965.

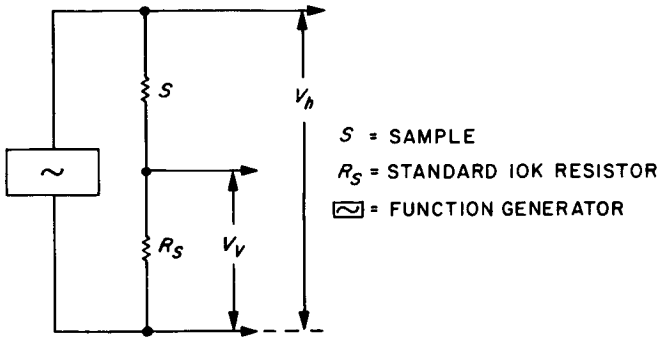


Fig. 32. Schematic diagram showing voltages applied to the vertical (V_v) and horizontal (V_h) oscilloscope inputs

3. Results

a. Spectrophotometric measurements. In Fig. 33 are recorded the spectra of iodine alone (Curve 1) and PVCA-iodine complexes (Curves 2, 3, 4, 5, and 6) all in CH_2Cl_2. The absorbancy changes occurring at 505 mμ with increased concentrations of PVCA are shown in Fig. 34.

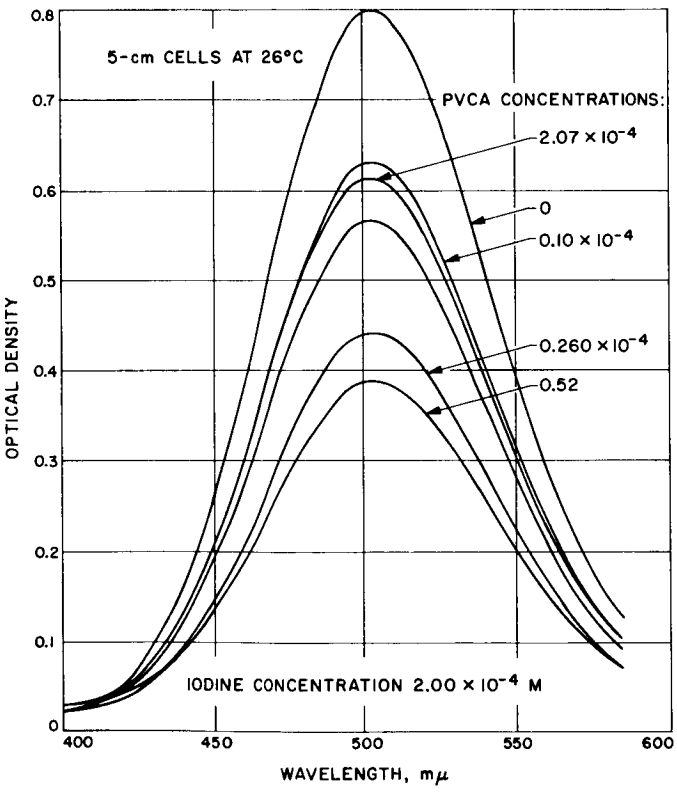


Fig. 33. Visible spectra showing the suppression and regeneration of the iodine peak for PVCA-iodine solutions in CH_2Cl_2

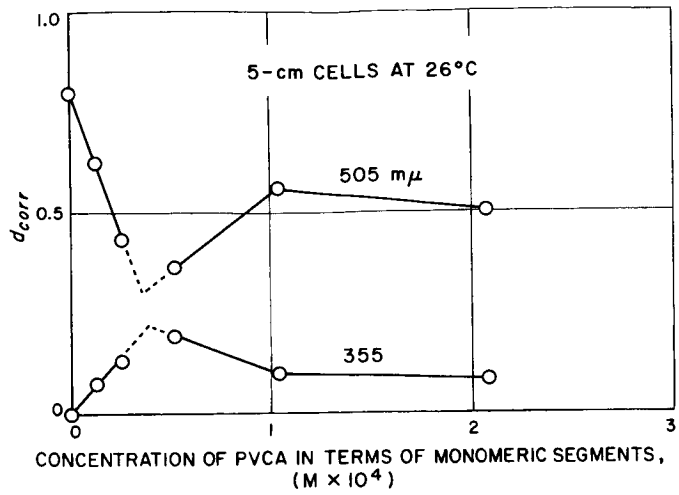


Fig. 34. Optical densities for solutions of PVCA at constant 2.00 × 10⁻⁴ M I_2 in CH_2Cl_2. Absorbancy due to PVCA alone was subtracted

The sharp decrease in the optical density of the iodine peak (505 mμ) is accompanied by an increase in the 355-mμ region (Fig. 34). A separate spectrophotometric study of N-ethyl carbazole serving as a model for the monomeric segment showed similar behavior for low concentrations (Fig. 35).

In contrast to solution measurements, a very definite charge transfer band was observed in the ultraviolet region of the spectrum when a KBr pellet of PVCA-I_2 was examined in the Cary spectrophotometer (Fig. 36).

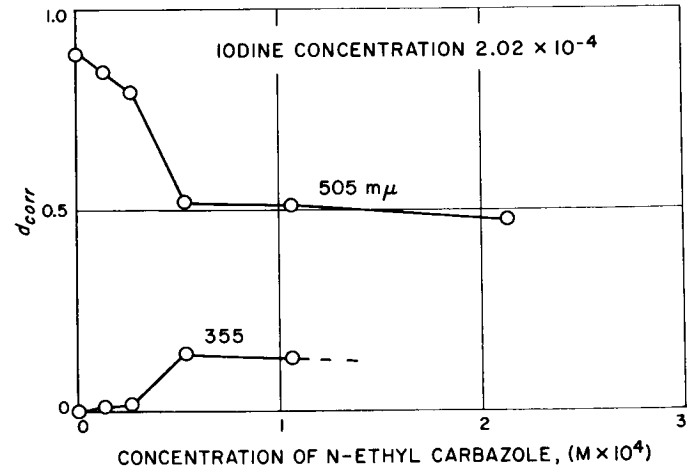


Fig. 35. Concentration of N-ethyl carbazole (M × 10⁴). Charge transfer band formation and iodine band decay in N-ethyl carbazole-iodine solutions in CH_2Cl_2

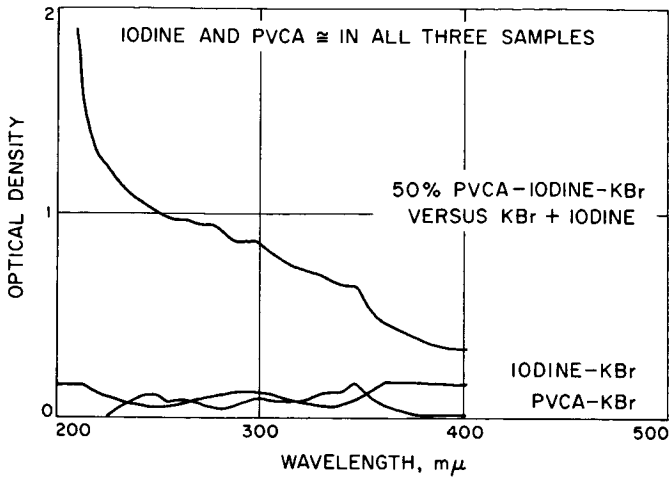


Fig. 36. Ultraviolet spectra of PVCA-iodine in KBr solid pellet

b. Resistivity measurements. The resistivity of PCVA-iodine as a function of composition is shown in Fig. 37. The resistivity dropped 6 orders of magnitude on increasing the amount of iodine to a mole fraction equal to 0.07, and 11 orders of magnitude on increasing the iodine content to the composition corresponding to one mole I_2 to one mole of the carbazole moiety (56.7% by weight, and hereafter referred to as the 1:1 composition).

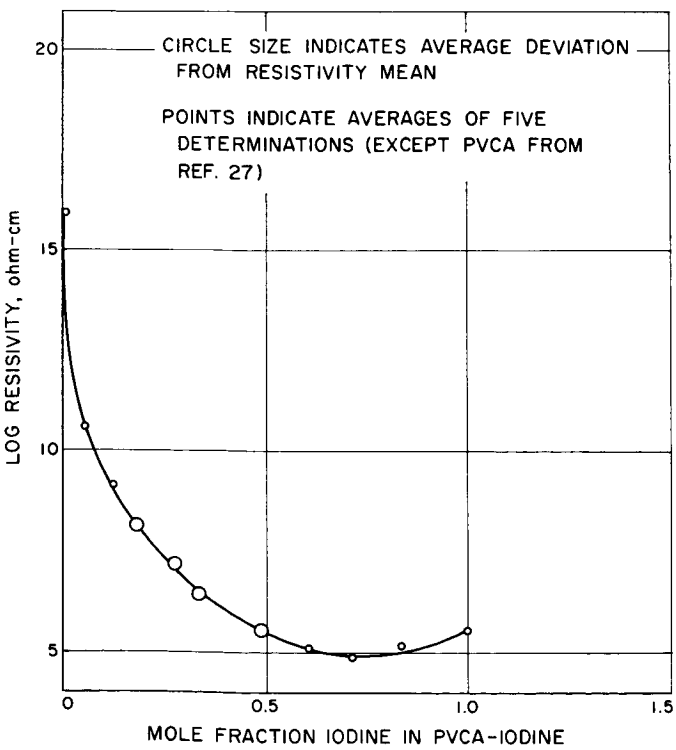


Fig. 37. Resistivity as a function of composition for PVCA-iodine of 26°C

The temperature variation of the resistivity for 3 PVCA-iodine compositions is shown in Fig. 38. If an Arrhenius type of relationship is assumed, then the activation energy E may be calculated from

$$\rho = \rho_0 \exp\left(\frac{E}{kT}\right) \tag{1}$$

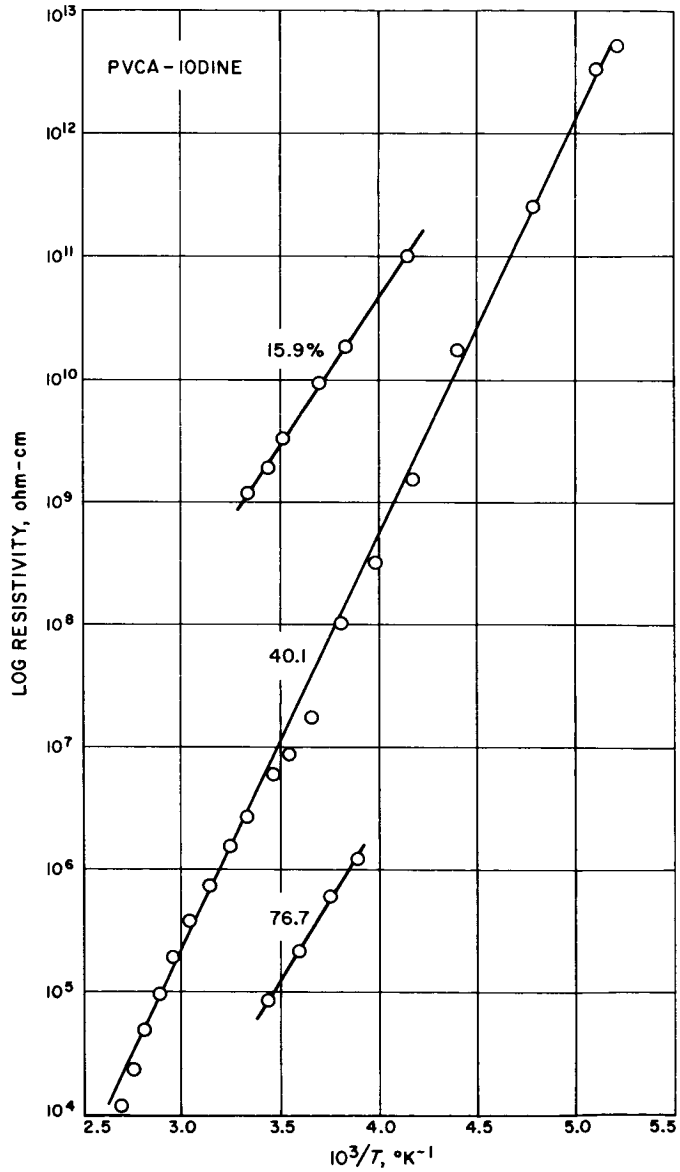


Fig. 38. Temperature variation of the resistivity of PVCA-iodine showing semiconducting properties for three iodine compositions

where ρ, ρ_0, k, and T are the resistivity, the pre-exponential factor, Boltzmann constant, and the absolute temperature, respectively. By means of Eq. (1) one calculates activation energies for the 15.9, 40.1, and 76.7% iodine in PVCA-I_2 pellets of 0.45, 0.70, and 0.55 ev, respectively. The obtained voltage current curves, proving ohmic behavior up to at least 400 v/cm are shown in Fig. 39.

The resistivities calculated from the slopes of the voltage current curves were in good agreement with the more accurate values obtained with the electrometer and are included in the averages recorded in Fig. 37.

4. Conclusions

(1) The decrease in the absorption intensity of the iodine band near 505 mμ and the increase in the region of 355 mμ (Figs. 34 and 35) could be attributed to formation of charge transfer complexes. This was confirmed by the study of absorption spectra of solid PVCA-iodine films[4] in which case the 505 mμ was eliminated almost completely. In very dilute solutions new charge transfer bands could be observed only in the vicinity of 355 mμ. Similar results were obtained by Ubbelohde, et al. (Ref. 28) during their study of condensed aromatic hydrocarbon-iodine complexes. Our own investigations of the pyrene-iodine complex failed to find new charge transfer absorption bands in dilute solutions. However, solid-state spectra show the existence of charge transfer bands quite conclusively (Fig. 37) and there is little doubt that the 200 to 400-mμ band does not appear in the pure components and is therefore due to the complex.[5] It should also be noted that the infrared solid-state spectrum of PVCA (from 2 to 15 μ) is identical to PVCA-I_2. Further evidence of charge transfer between PVCA and iodine is based on the very high free-spin concentrations determined by means of the electron spin resonance technique (Ref. 30), as well as on the minimum in the resistivity at a fixed composition (Fig. 37).

(2) The resistivity of pure polycrystalline iodine determined under identical conditions as used for the complexes was found to be lower than that reported in the literature (Ref. 31). In spite of this low value, there seems little doubt that when the latter forms a charge transfer complex with polymers containing aromatic hydrocarbons as side chains, its resistivity is lowered by approximately one order of magnitude, while the resistivity of the polymer is decreased quite considerably, namely, from 10^{16} to 10^5 ohm-cm at about 60% iodine content (Fig. 38).

[4] J. Moacanin and J. Farrar, unpublished results.

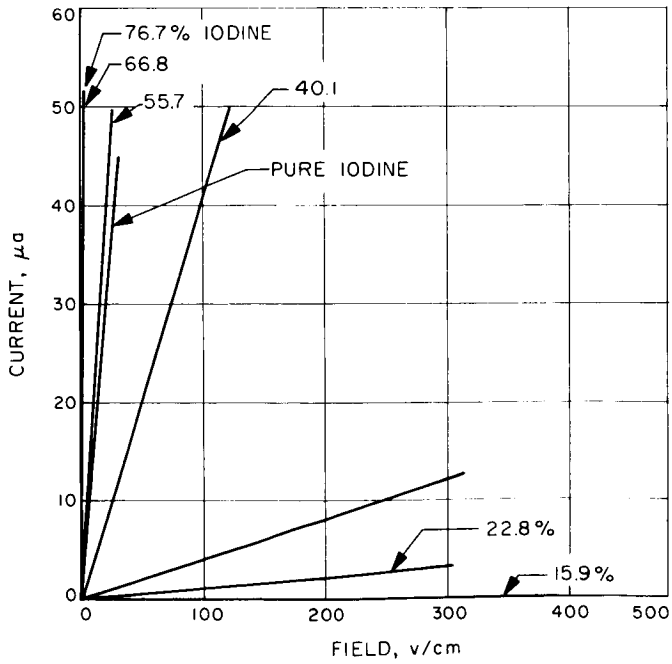


Fig. 39. Voltage-current curves for PVCA-iodine

F. The Dependence on Composition of the Optical Absorption of the Poly(N-Vinyl Carbazole)-Tetracyanoethylene (PVCA-TCNE) Charge Transfer Complex in the Solid State

J. W. Farrar and J. Moacanin

The PVCA-TCNE complex exhibits a very broad charge transfer absorption band in the visible region. The motivation for studying its optical absorption coefficient was suggested in previous studies on the dependence of

[5] The appearance of a charge transfer band at a lower wavelength agrees with the results of J. Ham for the benzene-iodine system (Ref. 29).

electrical resistivity on composition (*SPS 37-32*, Vol. IV) where the addition of TCNE in excess of 10 mole % did not result in additional changes of resistivity.

The principal results of this study are:

(1) Homogeneous solid films of the complex of certain compositions may readily be prepared.

(2) PVCA-TCNE does not form a 1:1 charge transfer complex as originally assumed, but more nearly an 8:1 complex.

(3) The optical absorption coefficient dependence on composition is consistent with the electrical resistivity dependence on composition.

(4) The optical absorption coefficient varies linearly with composition up to about 8 or 9 mole % TCNE.

(5) Increasing the temperature decreases the optical absorption slightly.

1. Experimental

The complex was prepared by mixing the desired proportions of the two components dissolved in dichloromethane to result in an approximately 1% solution. Solid films of a given thickness were cast by evaporating the solvent from an appropriate amount of the solution in a 3½-in. diameter flat bottom Petri dish floating on mercury. Fig. 40 shows the glassware used for casting the films. The rate of evaporation was controlled by a small hole in the lid of the jar. This technique enables homogeneous solid films of thickness from about 4 to 28 μ, measured with a micrometer, to be made out of the PVCA-TCNE complex of compositions up to about 14 mole % TCNE. Above 14 mole %, crystals of the TCNE appear distributed through the film in a nonhomogeneous way. A Cary 14 recording spectrophotometer was used to determine the absorbance. The spectra were taken in the wavelength region from 360 to 1300 mμ with the films between microscope slides (Fig. 41) and fastened to the sample cell holder.

The temperature dependence of the absorbance was carried out by holding the films between a microscope slide and the uncoated side of a NESA-coated glass plate (Pittsburgh Plate Glass Co.). This glass has a thin conducting-surface coating with resistivity of 80 Ω per

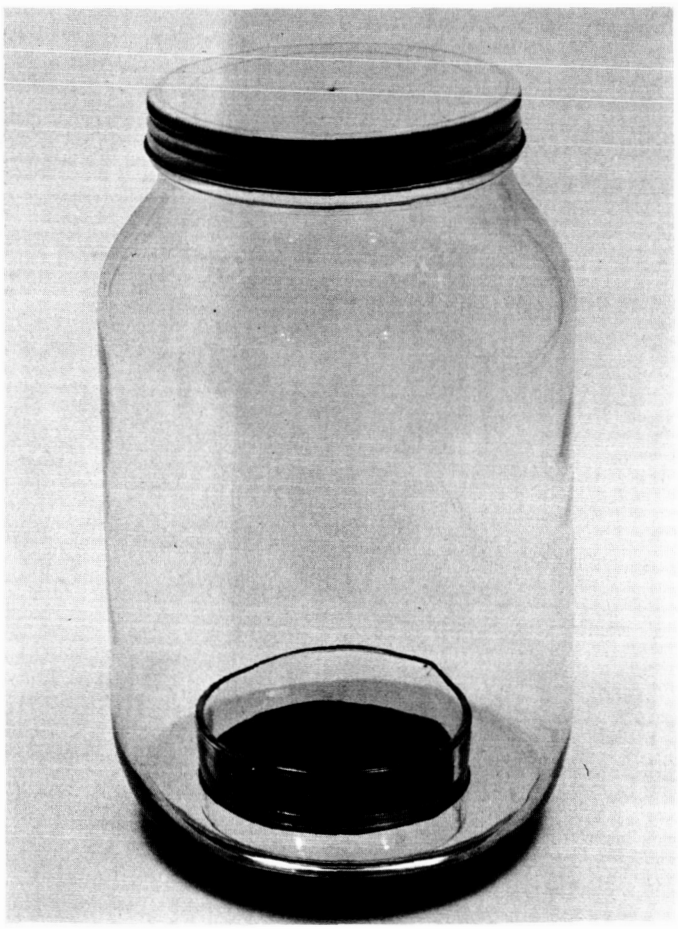


Fig. 40. Glassware for film casting

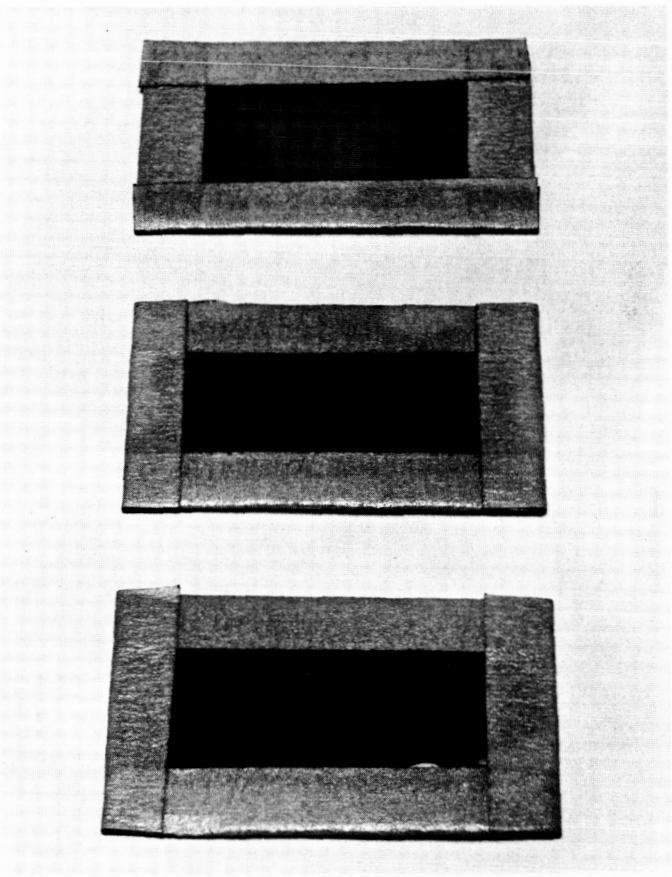


Fig. 41. Films held between microscope slides

square. A current of 0.1 to 0.3 amp was passed over the surface to heat the sample, and thermocouples were imbedded in the glass on the sample side to monitor the temperature. Fig. 42 shows the coated side of the heating device with the surface current electrodes and Fig. 43 shows the uncoated side with the thermocouples. Fig. 44 shows the device fastened on the sample cell holder.

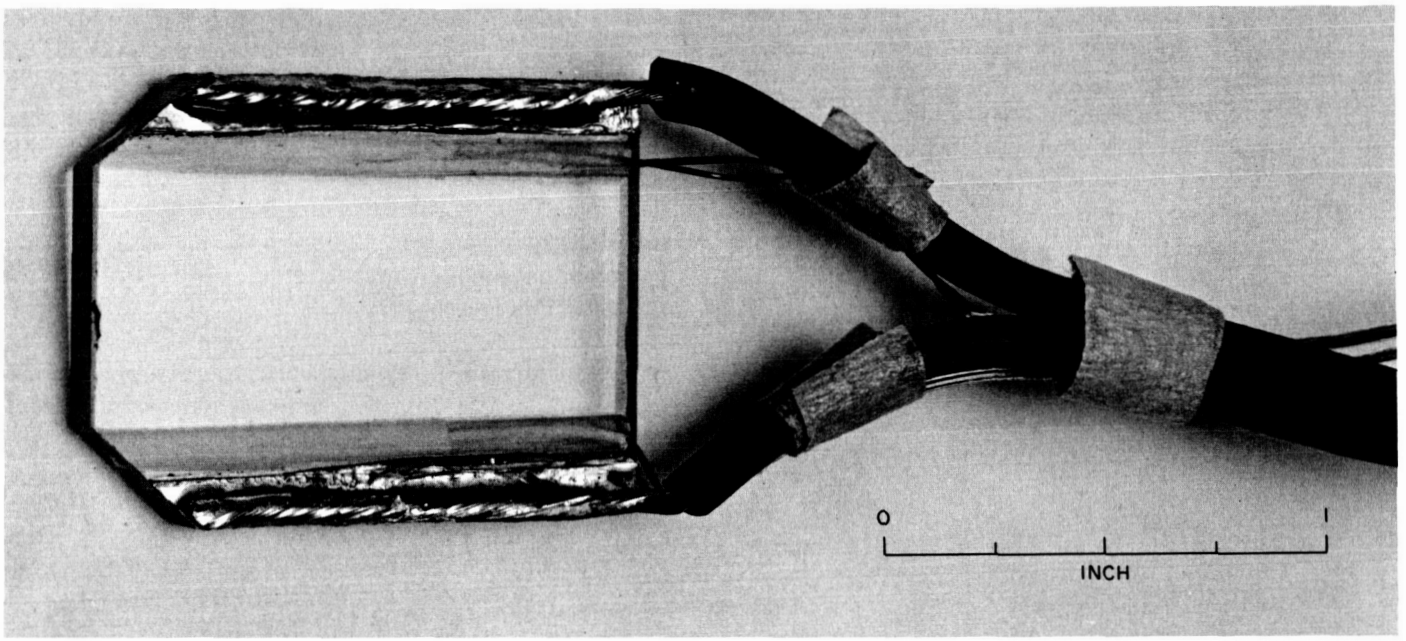


Fig. 42. Coated side of heating device showing the current electrodes

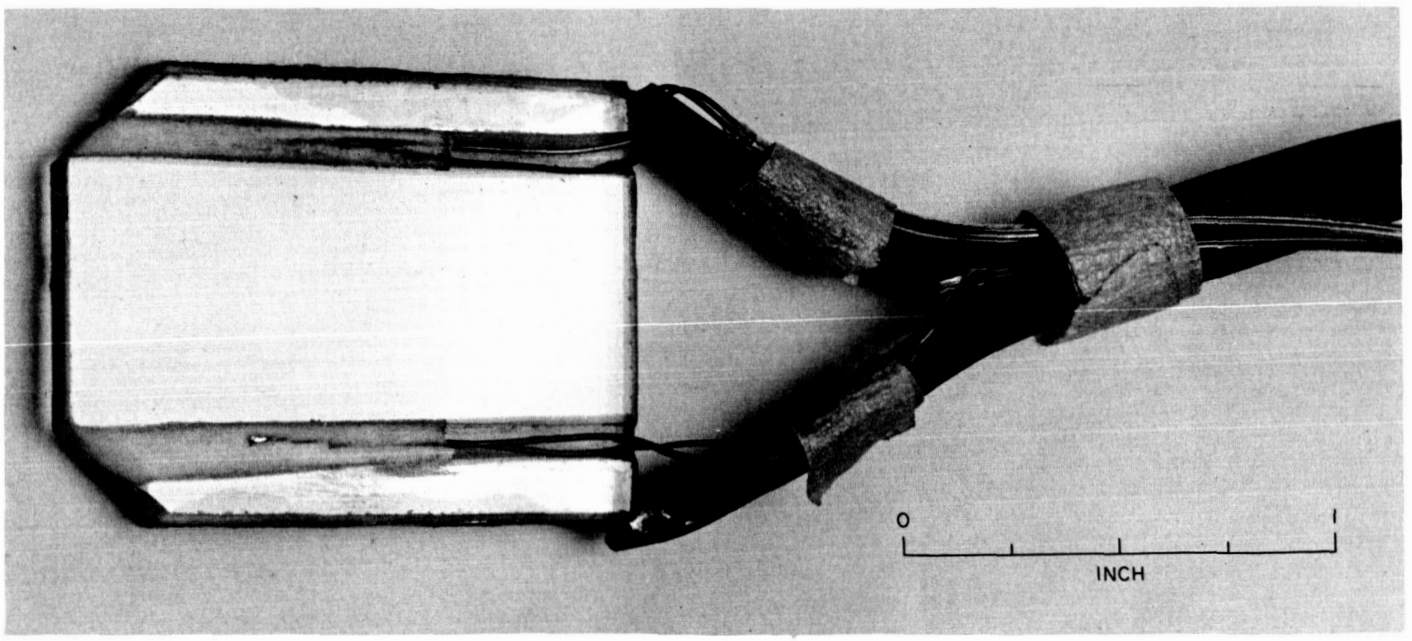


Fig. 43. Uncoated side of heating device showing the thermocouples

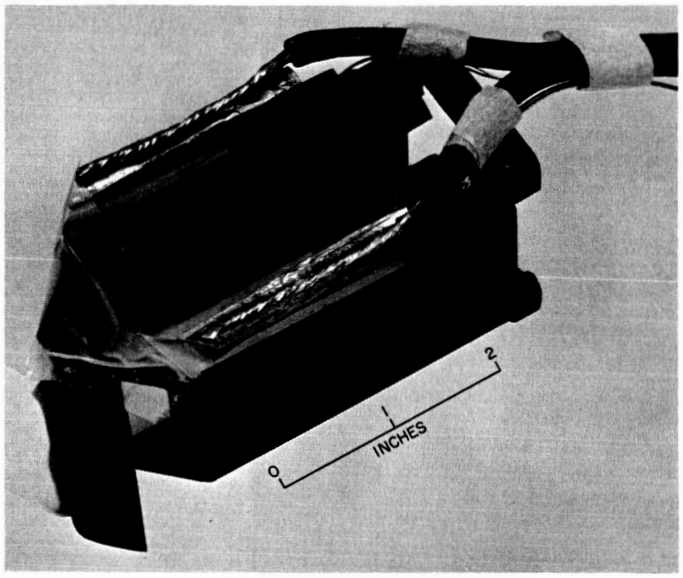


Fig. 44. Heating device fastened to cell holder

2. *Results and Discussion*

Absorption spectra between 360 and 1300 mμ were determined on nine PVCA films containing varying amounts of TCNE. Films containing up to 14 mole % TCNE were clear; but above this composition, TCNE crystals appeared, showing that its solubility limit has been exceeded. Fig. 45 shows the extinction coefficient k for the various compositions, where $k = (1/l) \log ((I_0/I)$. Fig. 46 shows that the extinction coefficient at maximum absorption (630 to 640 mμ) increases linearly with increasing TCNE content up to compositions of about 8 or 9 mole %. Above this composition range k levels off to a constant value of 0.12. This limiting k value, along with the observed solubility limit, suggests that a maximum of about one TCNE can complex with 8 carbazole units.

For increasing temperatures there is but a slight lowering of k, although the leveling off occurs about the same composition range. This change in k with temperature

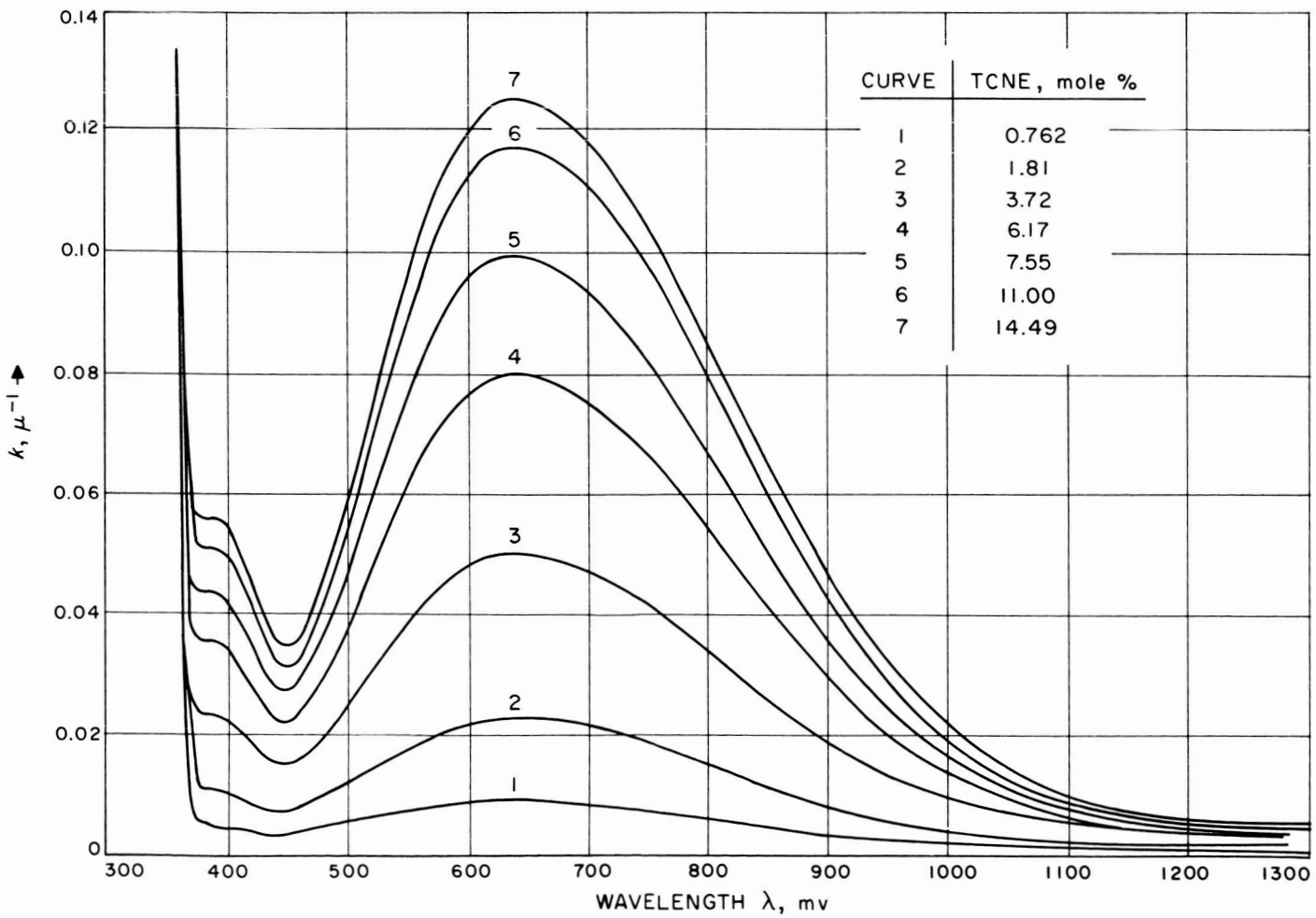


Fig. 45. Charge transfer absorption spectra

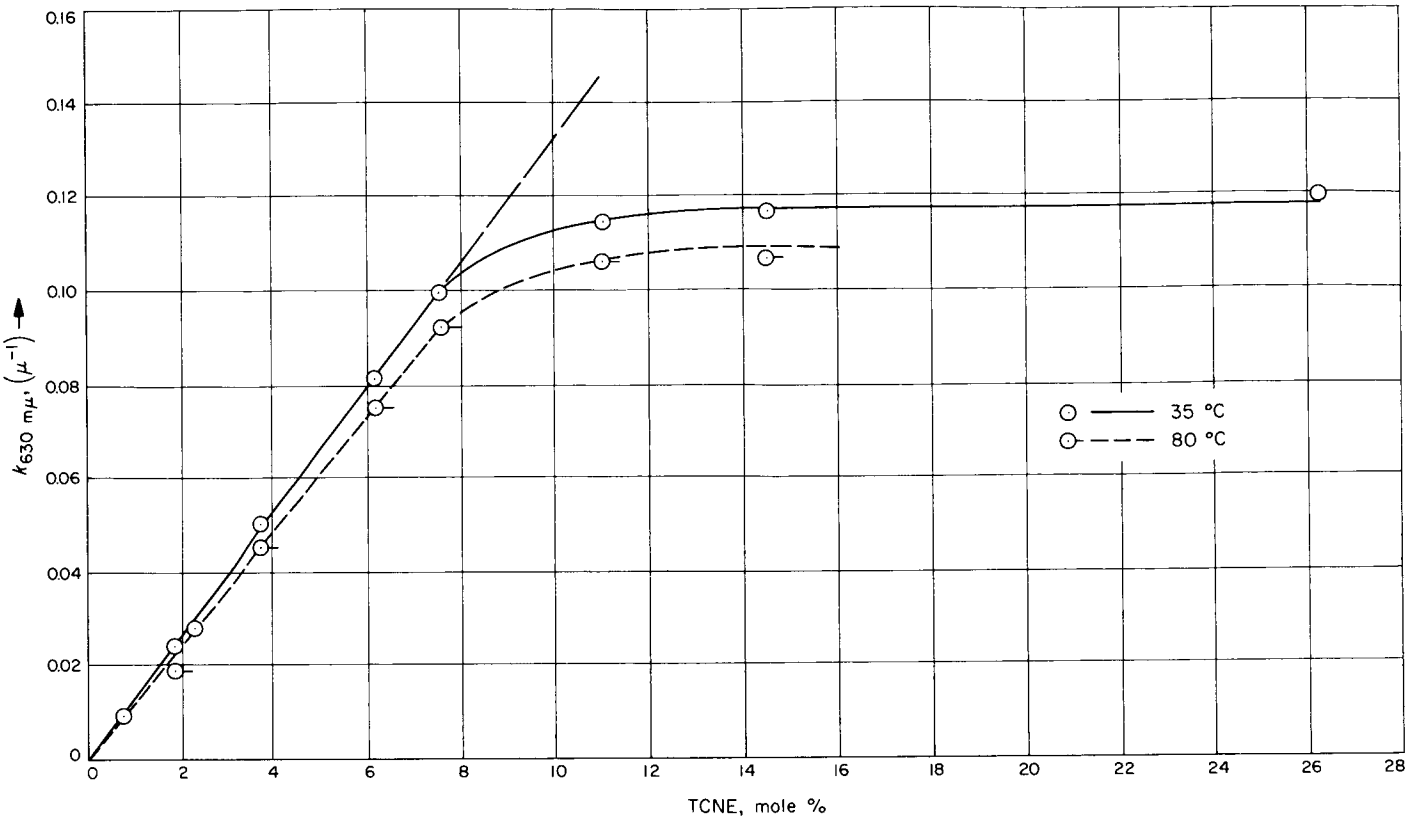


Fig. 46. Maximum of charge transfer band plotted versus mole % TCNE

should be related to the enthalpy of formation of the charge transfer complex. This point is now being investigated. Fig. 47 shows k versus T for the 6.17 mole % TCNE.

These results are to be compared with the previously reported dependence of resistivity on composition (*SPS 37-32*, Vol. IV). The resistivity decreases drastically with increasing TCNE (Fig. 48), and again levels off in the 9 to 10% range. This demonstrates the correlation between resistivity and complex concentration, and hence

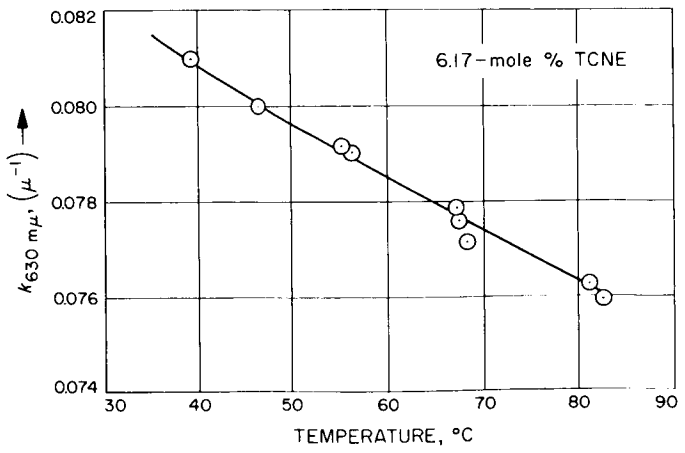


Fig. 47. Maximum of charge transfer band plotted versus temperature for the 6.17 mole % TCNE composition

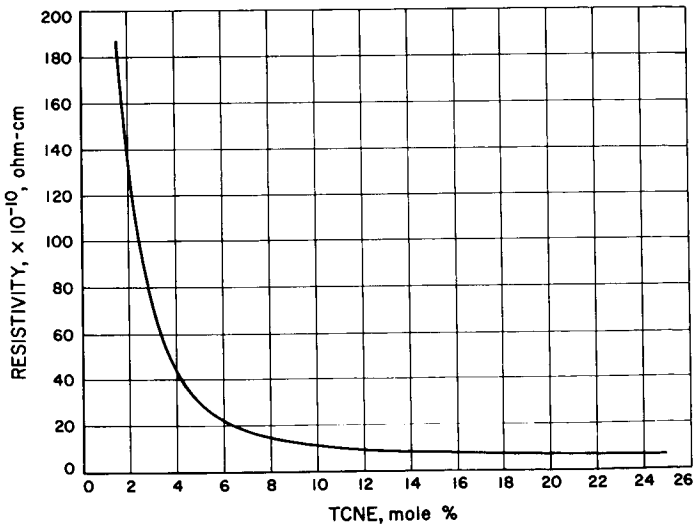


Fig. 48. Resistivity versus mole % TCNE data from SPS 37-32, Vol. IV

is in accord with the concept that the number of charge carriers is proportional to that of complexed carbazole moieties. Inasmuch as resistivity depends on both the number of carriers and mobility, the lack of simple proportionality between resistivity and composition (for less than 10% TCNE) implies also that the mobility varies with composition.

Now that solid films can be made, and some details of the absorption spectra and electrical properties are known, it becomes feasible to consider experiments in photoconductivity and the measurements of drift mobility by photo pulse techniques.

G. Carboxyl-Ending Poly(Alkylene Oxides) Purification and Curing

S. H. Kalfayan

1. Purification of Carboxyl-Ending Poly(Alkylene Oxides)

Previously, column chromatography and precipitation as "insoluble salt" methods (Ref. 32) were used to purify the crude carboxyl-terminated poly(alkylene oxides) (PAO-COOH), which contained by-products from the hydrolysis of the corresponding nitrile-terminated PAO. Although there was improvement in the purity of the products obtained, the pure dicarboxylated PAO could not be isolated.

Further attempts to purify the "insoluble salt" method were made by preparing the barium salts of the PAO-COOH, with barium hydroxide octahydrate. (Barium and other metallic salts were prepared previously from the sodium salts by metathesis). Products obtained from $Ba(OH)_2$ treatment showed the highest acidity yet, but still indicated the presence of hydroxyl groups. It is reasonable to conclude that monocarboxylic PAO, HOOC-R-OH, are largely responsible for the presence of hydroxyl groups and that these are not removable by the methods employed so far.

A typical purification with $Ba(OH)_2 \cdot 8H_2O$ is as follows:

Treat 100 g of carboxyl-terminated PPO-1000 with 200 ml 16% $BA(OH)_2 \cdot 8H_2O$. Extract the resulting solution several times with petroleum ether or hexane (total volume, 300 ml) and discard the extracts. Acidify the aqueous solution with hydrochloric acid after removal of the residual hydrocarbon *in vacuo*. Isolate the resulting oily layer, extract the aqueous layer with ether (100 ml) and join to the oily layer. Wash with distilled water until free of chloride ions. Dry and remove the ether. Yields are 60 to 70%.

2. Curing of PAO-COOH and Catalysis of the Carboxyl-Imine and Carboxyl-Epoxide Reactions

Imines and epoxides have been used to cure carboxyl-terminated poly(butadienes) (Refs. 33–35). Although cures with difunctional imines have been possible, only epoxides with functionality more than two have been capable of effecting satisfactory cures (Ref. 34).

Curing of PAO-COOH were tried with two triepoxides, namely, 2, 6(2, 3-epoxypropyl) phenyl glycidyl ether (Shell X 801) and N, N-diglycidyl-*p*-aminophenyl glycidyl ether (ERL-510, Union Carbide) and one triimine, tris 1-(methyl) aziridinyl phosphine oxide (MAPO). Two types of PAO-COOH were used: (1) carboxyl-terminated poly(propylene oxide) 1000 (PPO-COOH), and (2) carboxyl-terminated poly(ethylene oxide) 600 (PEO-COOH). Analysis of the specific preparations used in the experiments reported here showed 15% unconverted OH groups for PPO-COOH and 22% OH groups for PEO-COOH. The barium salt purification procedure was not applied to the latter. Cures were carried out at 100 and 145°C.

Some preliminary experiments were also carried out to study the catalysis of the carboxyl-epoxide and the carboxyl-imine reactions, using the PAO-COOH. Catalysts employed included dimethyl benzylamine (DMBA), stannous octoate (SO) and tetrabutyl tin (TBT). Acetic acid (AcOH) was used with MAPO in one run. The same percent by weight of catalyst was used in each sample so that a comparative evaluation of the activity of these catalysts could be made. Experiments were also carried out where the same mole percent catalyst was used. The experimental procedure consisted of mixing PAO-COOH and curing agent in stoichiometric proportions, adding to 5-g samples of the mixture the prescribed amount of catalyst and observing the gel time in air ovens set at 100 and 145°C.

Results of these experiments are given in Tables 4 and 5. Whenever hardness measurements were possible, they were made on samples that were post-cured for

Table 4. Cure and catalysis of PPO-COOH[a]

Curing agent	Catalyst or activator	Gel time, hr		Shore A hardness
		at 145° C	at 100° C	
ERL	None	>8, <22	30	15
	DMBA	1-1/2	4	14
	SO	3	—	Soft gel
	DMBA-SO	1-1/2	—	7
	TBT	>8, <22	—	Soft gel
Shell X 801	None			
	DMBA	>8, <22	No gelling after 1 wk	Soft gel, 145° C cure
	SO	>8, <22		Same as above
MAPO	None	remains fluid	3	Soft gel, 100° C cure
	DMBA	—	7	Same as above
	AcOH	—	6	Very soft gel, 100° C cure

[a] Mixing ratio: 100:18, PPO-COOH/curing agent. Catalyst concentration 1.8% by weight.

Table 5. Cure and catalysis of PEO-COOH[a]

Curing agent	Catalyst or activator	Gel time, hr		Shore A hardness
		at 145° C	at 100° C	
ERL 510	None	>8, <20	—	12
	DMBA	>4, <20	>5, <20	14
	SO	>4, <20	72	5
	TBT	>4, <20	—	Soft gel
Shell X 801	DMBA	>8, <20	—	Soft, tacky
	SO	>8, <20		
	TBT	No cure after 72 hr	—	Soft, crumbly —

[a] Mixing ratio: 100:30, PEO-COOH/curing agent. Catalyst concentration 1.8% by weight.

48 hr after gelation at the indicated temperature. In a number of cases gelation occurred overnight and exact gel times could not, therefore, be determined. Conclusions that can be withdrawn from results given in Tables 4 and 5 are as follows:

The order of reactivity of the curing agents in uncatalyzed reactions is:

$$\text{MAPO} > \text{ERL 510} > \text{Shell X 801}$$

Although MAPO is the fastest, it does not give satisfactory cures with the PAO-COOH samples used. Moreover, it reverts when cures at 145°C are attempted. Shell X 801 does not cure the PAO-COOH without catalysis.

The order of reactivity of the catalysts in the carboxyl-epoxy reaction on a weight percent basis is:

$$\text{DMBA} > \text{SO} \gg \text{TBT}$$

The same order is observed when equimolar quantities of catalysts (0.017 mole %) are used (not shown in tables).

DMBA and AcOH retard the MAPO-PPO-COOH cures. Cures with catalyzed Shell X 801 systems are invariably softer than catalyzed (or uncatalyzed) ERL 510 systems.

Softer products are obtained with SO than with DMBA. The uncatalyzed ERL 510 systems give harder products than the same systems catalyzed with SO.

3. Summary

Improvements have been made in the methods of obtaining pure dicarboxylated PAO. Purified materials still show the presence of OH groups (15%), thought to have their origin in monocarboxylic PAO. The effectiveness of two triepoxides and one imine as curing agents for PAO-COOH have been determined. The catalysis of the systems PAO-COOH-epoxide and PAO-COOH-imine have been studied in preliminary experiments.

4. Future Work

A comparative study of the properties of cured carboxyl-terminated poly(butadienes) and carboxyl-terminated poly(propylene oxides), using preparations with similar

carboxyl and nonfunctional material content, will be made. Other epoxides and imines will be investigated as curing agents. The effect of tertiary amine, organometallic and other types of catalysts, on cure rates and on physical properties at ambient and elevated temperatures will be studied.

H. Temperature Dependence of Outgassing Rates in Polymeric Foams

E. F. Cuddihy and J. Moacanin

In a previous study (*SPS 37-34*, Vol. IV), the dependence was presented of the outgassing rates in polymer foams on polymer type and density. Included in that report was the development of an equation relating the diffusion coefficient to the foam density. Also, for one-dimensional gas diffusion, the necessary equations were given to calculate D, the diffusion coefficient from data obtained by measuring the weight loss of gas from a foam as a function of time.

This study presents the results obtained on the temperature dependence of the outgassing rate for one polyurethane foam.

1. Experimental

The polyurethane foam used in this study was a foam-in-place material purchased from Emerson and Cummings. The material is designated Eccofoam FPH/12/6H and was prepared with a density of 4.5 lb/ft^3.

To obtain weight loss data for one-dimensional gas diffusion, foams of 3 in. diam and 1 in. deep were bonded into aluminum containers leaving only the top surface exposed. The bonding agent was Eccobond 55, a room temperature curing epoxy also distributed by the Emerson and Cummings Corp.

For elevated temperatures, weight loss was measured employing a Cahn RH Electrobalance which was situated inside a vacuum unit manufactured by the Consolidated Vacuum Corporation. The sample was heated by irradiation with a Variac-controlled 250-w bulb. A temperature differential of less than 1 deg was maintained across the sample by means of a reflecting aluminum shroud. Details of the experimental setup are shown in Figs. 49 and 50.

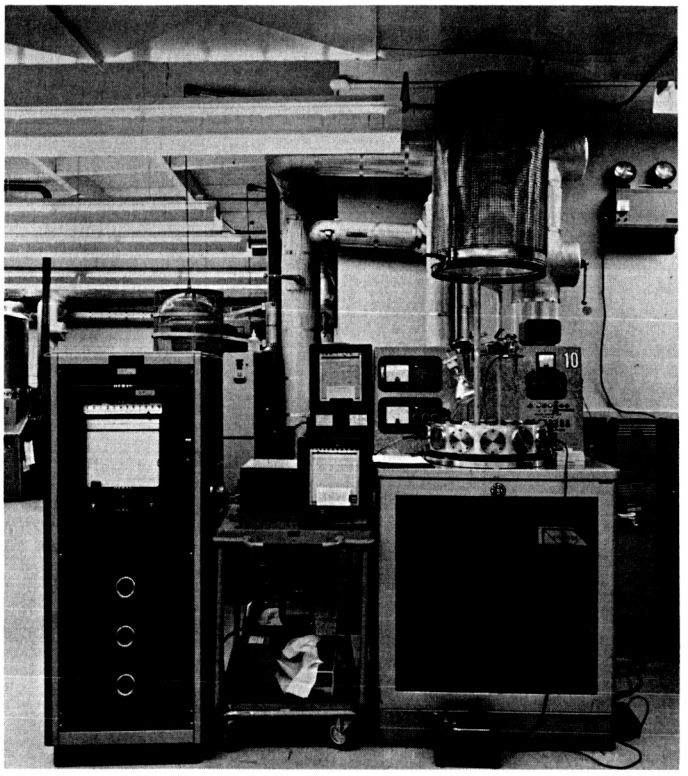


Fig. 49. Vacuum balance facility for elevated temperature operations

The initial weight of CO_2 gas Q_∞, in the foam was calculated using the perfect gas law.

2. Discussion

Outgassing rates were determined at 22, 41, 61, and 81°C. The tests were conducted by first heating the samples in air to the desired temperature and then subjecting them to a vacuum. It was found, however, that samples heated in air would lose an amount of CO_2 which increased with increasing temperature. It became a practice therefore to start recording the weight loss in air during heating. When the desired temperature was reached and the weight loss rate slowed significantly, vacuum was applied. The results are shown in Fig. 51, for weight loss both in air and in vacuum with zero time taken when the vacuum was initiated.

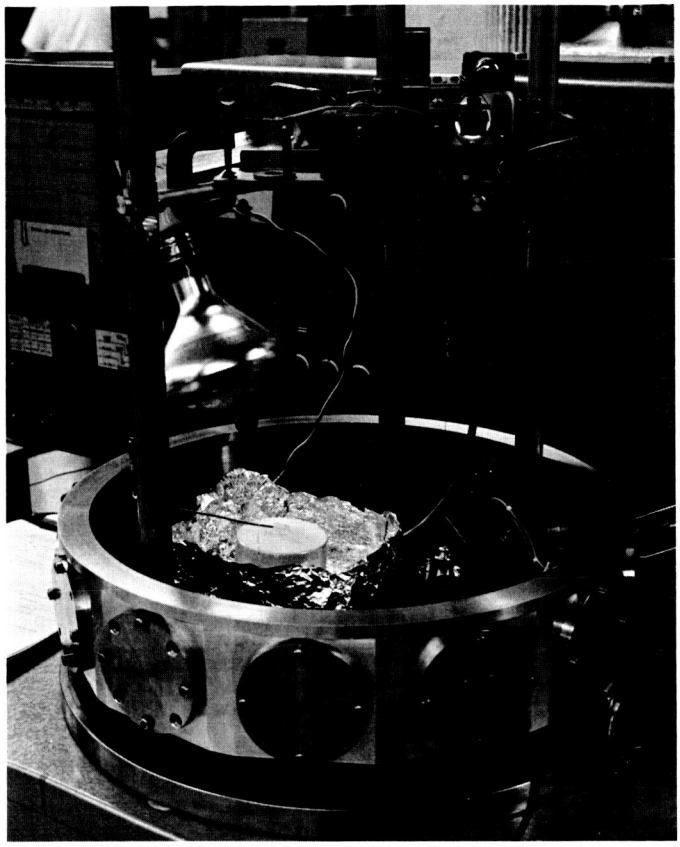


Fig. 50. Details of heating and balance assembly for elevated temperature operations

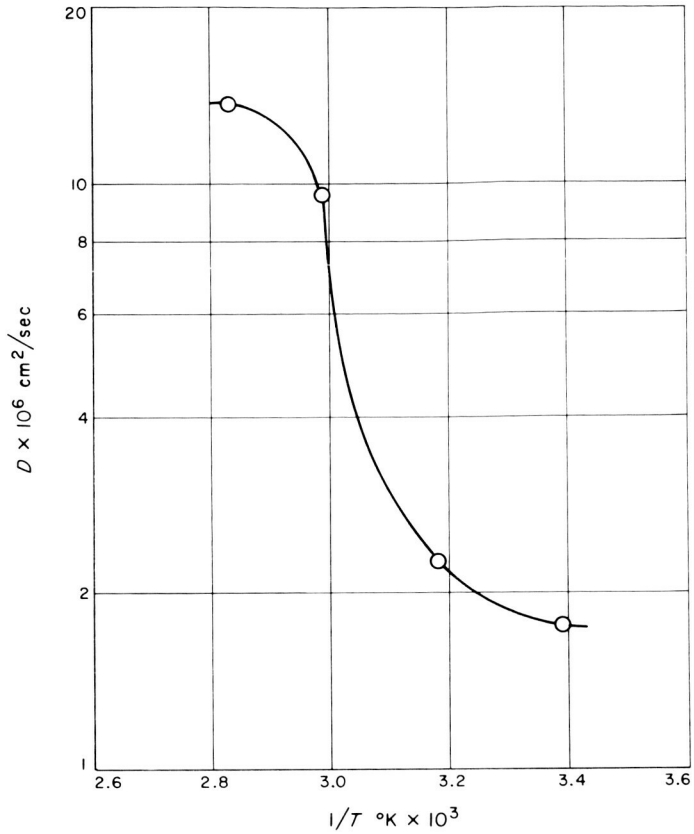


Fig. 52. *D* versus 1/T °K for an Eccofoam FPH/12/6H foam of density 4.5 lb/ft³

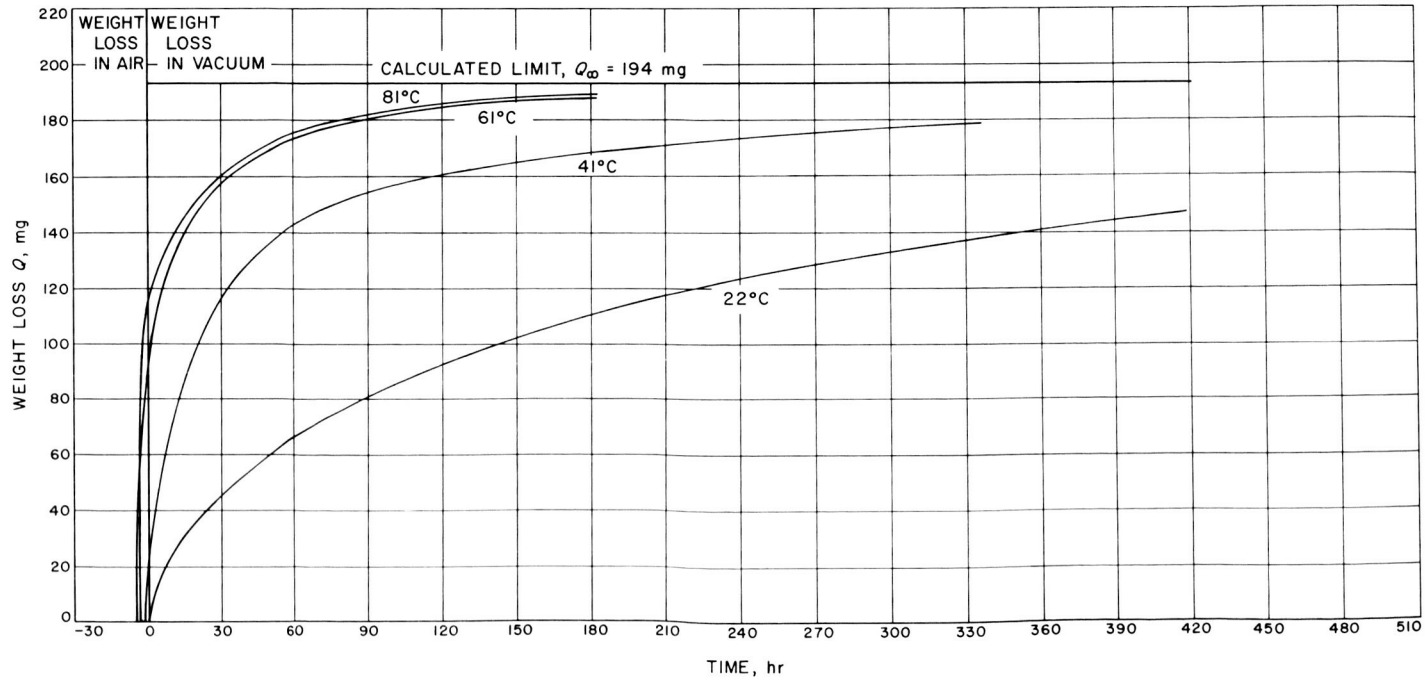


Fig. 51. Temperature dependence of the outgassing rate for an Eccofoam FPH/12/6H foam of density 4.5 lb/ft³

It can be seen that there is a dramatic increase in the rate of gas removal with increasing temperature, the largest change occurring between 22 and 61°C. Above 61°C, the changes are small. It is instructive to observe that at the higher temperatures, the weight loss asymptote appears to be the value of $Q_\infty = 194$ mg which corresponds with the value calculated from the perfect gas law. This strengthens the use of the perfect gas law in determining the values of Q_∞. Also using a value of $Q_\infty = 194$ mg, it is seen that this foam has lost approximately 97.5% of its gas content in 180 hr at 61°C and higher, whereas only 57% of the gas content was lost in 180 hr at 22°C.

The diffusion coefficients were obtained at the various temperatures and are shown plotted in Fig. 52 as log D versus $1/T$ °K. The general appearance of this curve is sigmoidal and indicates that the largest changes in the diffusion coefficient occurs between 22 and 81°C, above and below which there appears to be only a gradual dependence of log D on $1/T$.

References

1. Fedors, R. F., and Landel, R. F., "Mechanical Behavior of Styrene-Butadiene Rubber Filled with Glass Beads," *SPS 37-32*, Vol. IV., p. 120, Jet Propulsion Laboratory, Pasadena, California, 1965.
2. Smith, T. L., "Mechanisms of Reversible and Irreversible Loss of Mechanical Properties of Elastomeric Vulcanizates Which Occur at Elevated Temperatures," ASD-TDR-62-572, Wright-Patterson Air Force Base, Ohio, 1962.
3. Landel, R. F., and Fedors, R. F., "The Tensile Failure Envelope of Amorphous Elastomers: Effects of Statistical Variability and Crosslink Density," Proceedings of the Fourth International Congress on Rheology, Part 2, E. H. Lee, Editor, p. 543, John Wiley and Sons, Inc., Interscience Publishers Div., New York, N.Y., 1965.
4. "Effects of Filler Concentration on Mechanical Properties of Propellants," *Combined Bimonthly Summary 64*, p. 24, Jet Propulsion Laboratory, Pasadena, California, April 15, 1958 (Confidential).
5. Landel, R. F., and Smith, T. L., "Viscoelastic Properties of Rubber-like Composite Propellants and Filled Elastomers," *ARS Journal*, Vol. 31, p. 599, 1961.
6. Landel, R. F., "The Dynamic Mechanical Properties of a Model Filled System: Polyisobutylene-Glass Beads. II. Long-Time Studies," paper presented before the Division of High Polymer Physics, American Physical Society, Boston, March 1959.
7. Landel, R. F., and Moser, B. G., "Relative Sedimentation Volume of Ammonium Perchlorate in Polypropylene Glycol," *SPS 37-17*, Vol. IV, p. 189, Jet Propulsion Laboratory, Pasadena, California, October 30, 1962 (Confidential).
8. Guth, E., "Theory of Filler Reinforcement," *Journal of Applied Physics*, Vol. 16, p. 20, 1945.
9. Smallwood, H. M., "Limiting Law of the Reinforcement of Rubber," *Journal of Applied Physics*, Vol. 15, pp. 758–766, 1944.
10. Van der Poel, C., "On the Rheology of Concentrated Dispersions," *Rheologica Acta*, Vol. 1, p. 198, 1958.

References (Cont'd)

11. Hashin, Z., "Theory of the Mechanical Behavior of Heterogeneous Media," *Applied Mechanics Review,* Vol. 17, p. 1, 1964.

12. Weissberg, H. E., and Prager, S., "Viscosity of Concentrated Suspensions of Spherical Particles," Proceedings of the Fourth International Congress on Rheology, Providence, R. I., 1963, Part 2, p. 709 (abstract only), E. H. Lee, Editor, John Wiley and Sons, Inc., Interscience Publishers Div., New York, N.Y., 1965.

13. Landel, R. F., Moser, B. G., and Bauman, A. J., "Rheology of Concentrated Suspensions: Effect of a Surfactant," Proceedings of the 4th International Congress on Rheology, Brown University, August 27, 1963, Vol. 2, pp. 663–692, John Wiley and Sons, Inc., Interscience Div., New York, N.Y., 1965.

14. Moser, B. G., Landel, R. F., and Bauman, A. J., "Rheology of Suspensions II: Implications of Surfactants for High Energy Systems," *Bulletin of the 2nd Meeting of the ICRPG Working Group on Mechanical Behavior,* Hill Air Force Base, Utah, November 19, 1963, p. 273 (Confidential).

15. Kimizuka, H., "Binding of Calcium Ion to Lecithin Film," *Nature, Vol. 196,* p. 995, 1962.

16. Rosano, H., "Molecular Interaction Between Phospholipids and Salts at Liquid-Liquid Interfaces," *Journal of Physical Chemistry, Vol. 66,* p. 1928, 1962.

17. Anderson, P. J., and Pethica, B. A., "Biochemical Problems of Lipids," *Proceedings of the 2nd International Congress* (Ghent) pp. 24–29, G. Popjak, E. LeBreton, Editors, Butterworths, London, 1955.

18. Vold, R. D., and Sivaranokrishan, N. H., "The Origin of the Maximum in the Adsorption Isotherms of Associated Colloids," *Journal of Physical Chemistry,* Vol. 62 pp. 984–989, 1958.

19. Wiech, R. E., Jr., and Moser, B. G., "The Theory of Sedimentation Volume for Particle Systems of Uniform Spheres," to be published in the Preprints of the ICRPG Bulletin, meeting of November 17–19, 1965 in San Diego, California.

20. Bland, D. R., "The Theory of Linear Viscoelasticity," Pergamon Press, New York, N. Y., 1960, Chap. 4.

21. Pestel, C. E., and Leckie, F. A., "Matrix Methods in Elastomechanics," McGraw Hill Book Company, Inc., New York, N.Y., 1963, Chap. 9.

22. San Miguel, A., Silver, R. H., and Duran, E. N., "On the Practical Application of a Miniature Stress Rosette to Solid Propellant Grain Design," Bulletin of the Fourth Meeting of ICRPG Working Group on Mechanical Behavior, October 15, 1965, SPIA, The Johns Hopkins University, Applied Physics Laboratory, Silver Spring, Maryland.

23. Orchin, M., and Woolfolk, W., *Journal of the American Chemical Society,* Vol. 68, p. 1727, 1946.

24. Horner, L., and Spietschke, W., *Annalen der Chemie,* Vol. 591, pp. 1–20, 1955.

25. Bornstein, E., Schliewiensky, H., and Szcesny-Heyl, G. V., *Berichte der Deutschen Chemischen Gesellschaft,* Vol. 59B, p. 2812, 1926.

26. Midzuch, K. G., *Journal of General* Chemistry (USSR), Vol. 16, p. 1471, 1946.

References (Cont'd)

27. Busse, W. F., *Industrial and Engineering Chemistry,* Vol. 40, p. 2271, 1948.

28. Martin, G. C., and Ubbelohde, A. R., *Journal of the American Chemical Society,* p. 4948, 1961.

29. Ham, J., *Journal of the American Chemical Society,* Vol. 76, p. 3875, 1954.

30. Hermann, A. M., and Rembaum, A., "Hall Mobility and Electron Spin Resonance of Poly-N-Vinylcarbazole-Iodine Charge Transfer Complex," paper to be presented at the ACS Western Regional Meeting, Los Angeles, California,November 18, 1965.

31. Hermann, A. M., and Ham, J., "Apparatus for the Measurements of the Hall Effect in Semiconductors of Low Mobility and High Resistivity," *Review of Scientific Instruments,* in press. See also Mainthia, S. B., Kronick, P. L., Ur, H., Chapman, E. F., and Labes, M. M., "Electronic Complexes of Poly-p-phenylene," Polymer Preprints, American Chemical Society, Los Angeles meeting, 1963; and Many, A., Simhony, M., Weisz, S. Z., and Levinson, J., *Journal of Chemistry and Physics of Solids,* Vol. 22, p. 285, 1961.

32. Kalfayan, S. H., "Hydrolysis of Higher Molecular Weight Alkoxy-propionitriles," *SPS 37-29,* Vol. IV, October 31, 1964.

33. Reynolds, W. B., "Curing Polymers Having Terminal Reactive Acidic Groups," U.S. Patent 3,074,917 (issued to Phillips Petroleum Co.) January 22, 1963.

34. Gruver, J. T., "Butarez CTL–Terminally Reactive Polybutadiene," Phillips Petroleum Co., Research Division Report 3745-64R, 1964.

35. Thiokol Chemical Corp., "HC Polymer 434—Carboxyl Terminated Polybutadiene Polymer," Information Bulletin, 1964.

XII. Solid Propellant Engineering

A. Nozzle Thrust Misalignment

L. Strand

1. Introduction

A number of questions connected with the aerodynamics of nozzle flow of rocket motors remain largely unanswered. They include questions of: (1) the contributions of rutting of the nozzle entrance and expansion sections and variations of the nozzle throat cross section to misalignment of the rocket thrust vector, (2) the magnitude of thrust misalignment occurring during motor chuffing, and (3) the losses in motor performance due to submerging the forward portion of the rocket nozzle.

A decision was made to attempt to investigate Item (1), using both experimental and analytical techniques. For the initial test a three-dimensional cold flow test system was considered preferable to actual motor firings. This would greatly decrease the expense involved per test and would eliminate the problems of maintaining thermal equilibrium at high temperatures and maintaining the same nozzle throat cross-section throughout the test firing.

2. Auxiliary Flow Channel

The resulting test facility has been designated the "auxiliary flow channel." It consists of a gas line tapped off of the JPL Hypersonic 21-in. Wind Tunnel below the heated air supply section and prior to the test section inlet valve. The line in turn exhausts into the tunnel diffuser, thereby by-passing the wind tunnel test section. The by-pass line contains the test assembly shown in Fig. 1. The test assembly consists of a plenum chamber to straighten and stabilize the flow, a reducer section, the particular nozzle in use, an expansion joint to allow for nozzle removal, and a spacer section to accommodate test nozzles of increased length. Fig. 2 shows the 6-in. ID stainless steel line that delivers the compressed air to the plenum chamber. Fig. 3 is a closeup of the inlet valve located in the upper right-hand corner of Fig. 2. Fig. 4

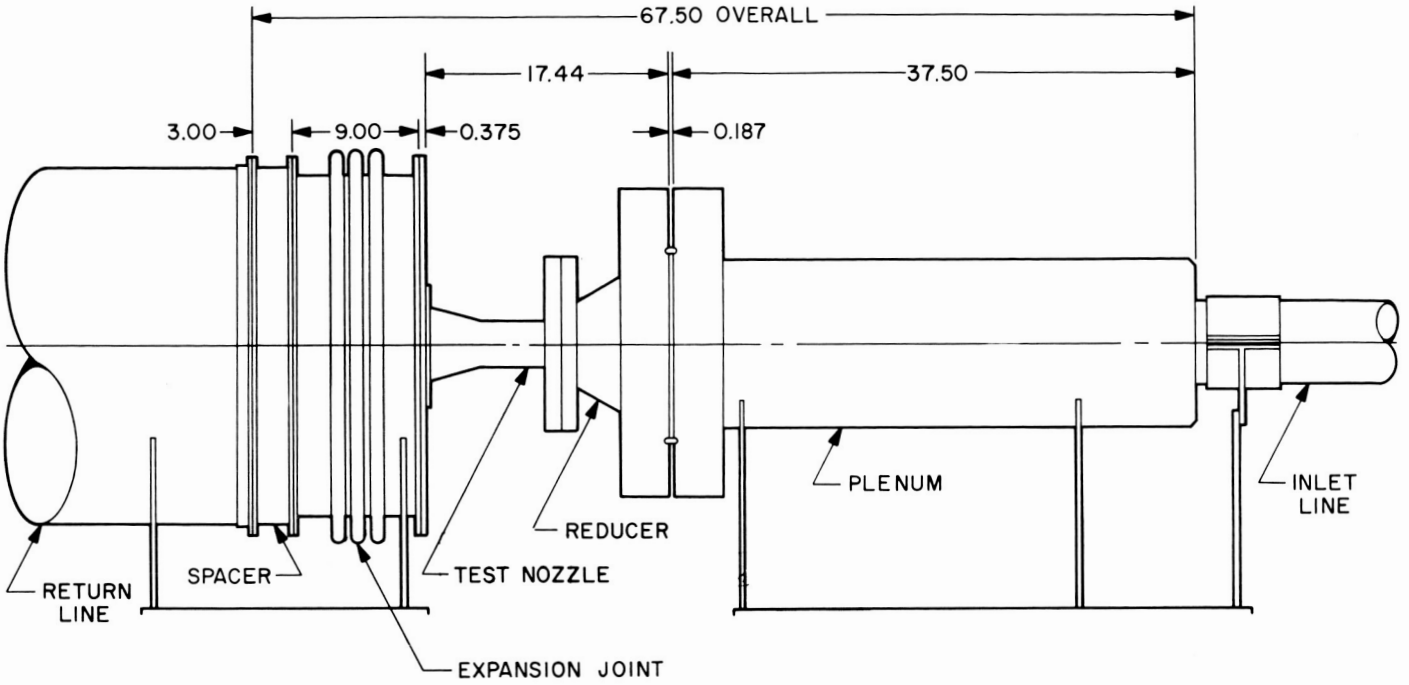


Fig. 1. Gas-flow nozzle test assembly

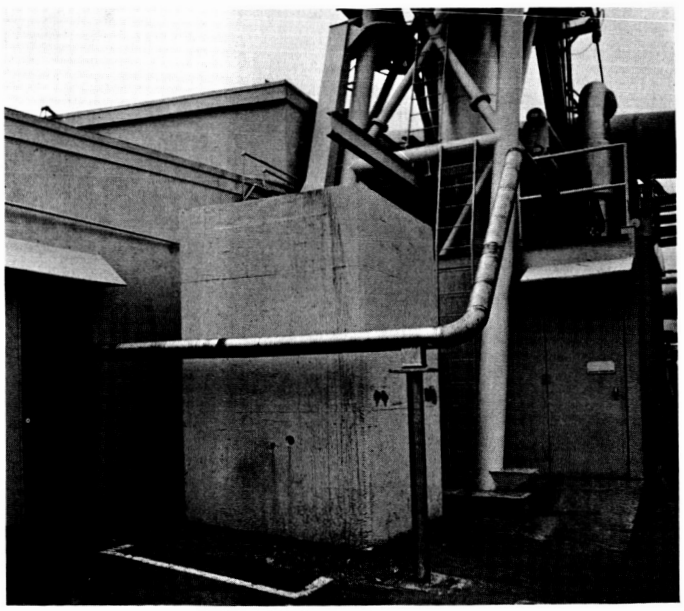


Fig. 2. Compressed-air inlet line to plenum chamber

shows the inlet line, test assembly, and 24-in. OD diffuser and return line. The diffuser is instrumented along its length with static pressure taps, read by the mercury manometer board shown. Figs. 5 and 6 show the return line crossing over and joining the wind tunnel diffuser. The outlet line valve is shown in Fig. 6.

The air supply capabilities of the auxiliary flow channel are the same as those of the hypersonic wind tunnel, neglecting losses in the inlet line: approximately 660-psia maximum supply pressure and 10-lb/sec maximum air flow rate. The wind tunnel heater was not used for these tests, so the inlet line was not insulated.

3. *Initial Tests*

In the initial tests the following basic questions are being attempted to be answered: Does a protrusion to the flow in the nozzle expansion region and an asymmetry in the nozzle throat cross section (an exaggeration of nonsymmetrical throat erosion) displace the thrust vector and, if so, to what magnitude? An attempt has been made to measure the thrust misalignment for two selected flow systems.

Two steel nozzles of equal throat area (1.0 in.^{2}), expansion ratio (18/1), and over-all length (9.56 in.) were fabricated as part of the test assembly (Fig. 7). The inner surfaces of the nozzles were nickel- and chrome-plated to resist corrosion. The nozzles were instrumented with static pressure taps mounted in the walls of both the nozzle throat sections and expansion cones, allowing maps of the wall pressure distribution to be determined. By integrating the pressure distribution over the nozzle

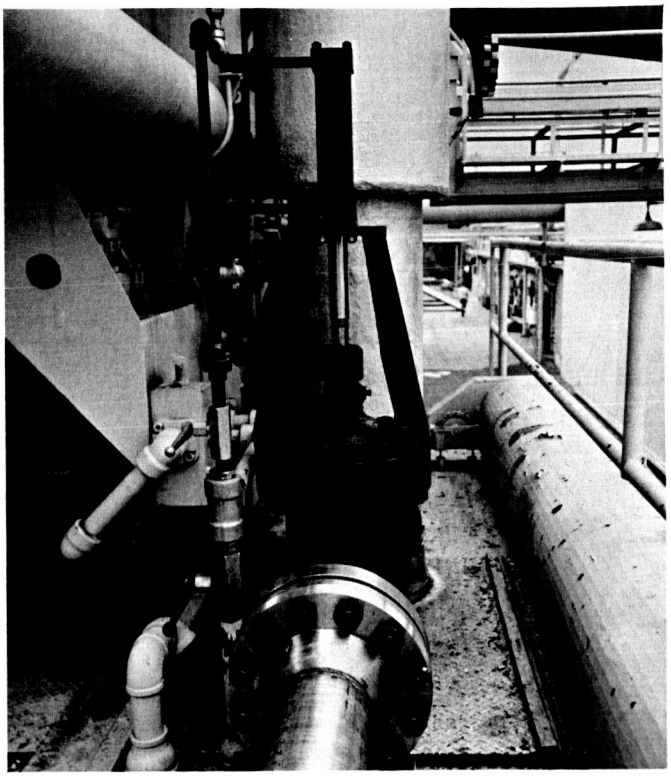


Fig. 3. Inlet valve in Fig. 2

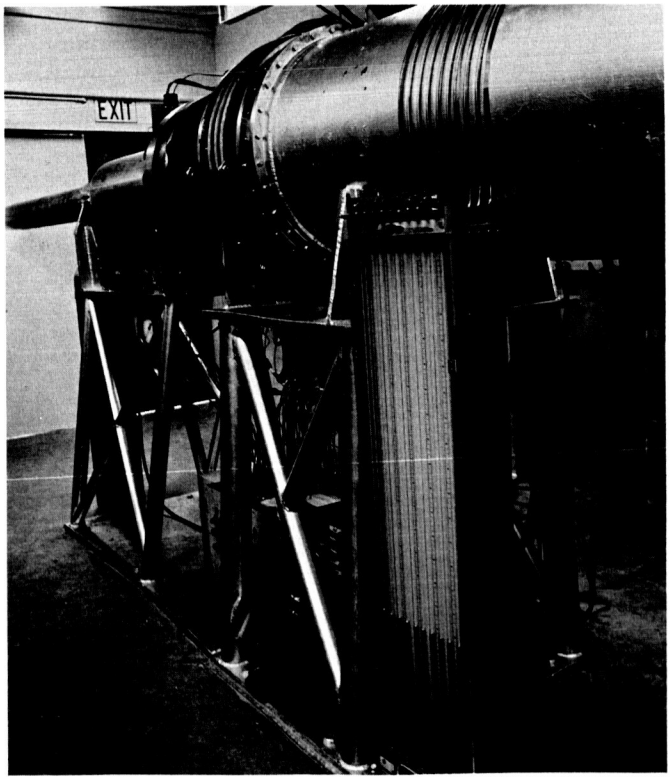


Fig. 4. Auxiliary flow channel test facility

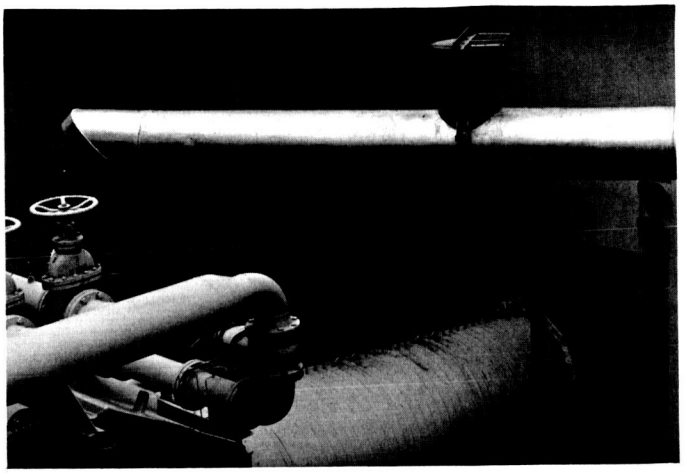


Fig. 5. Return line

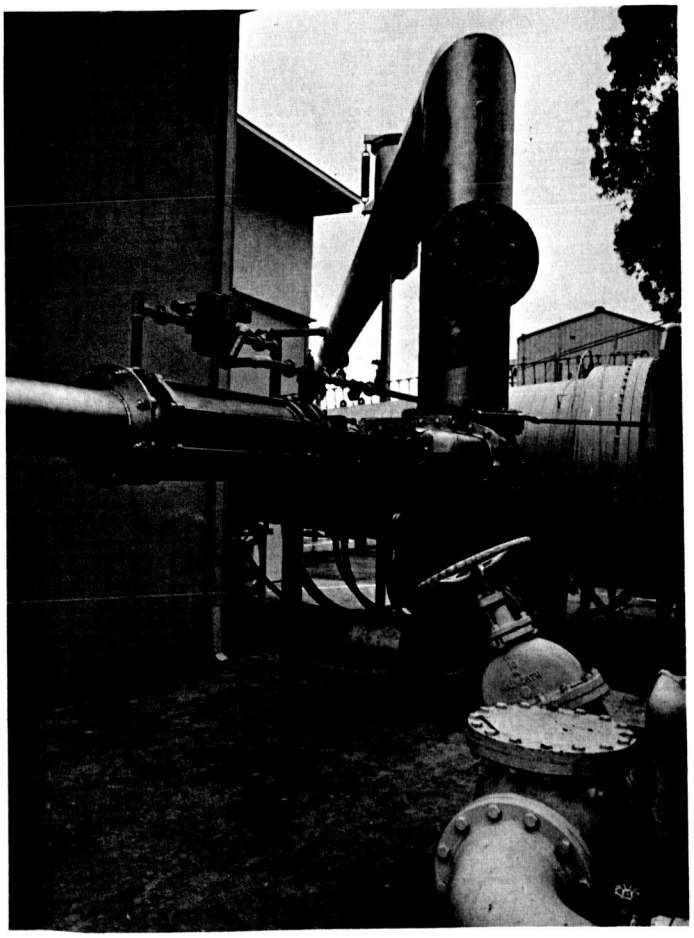


Fig. 6. Return line connection

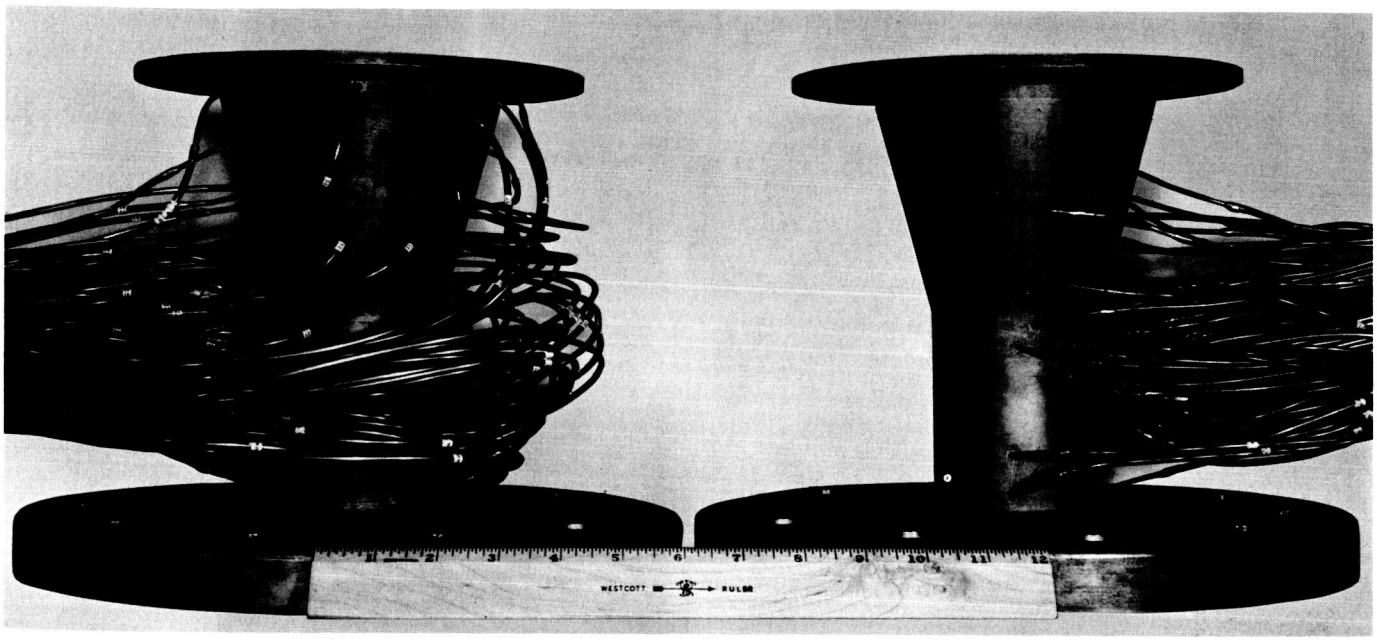


Fig. 7. Test nozzles

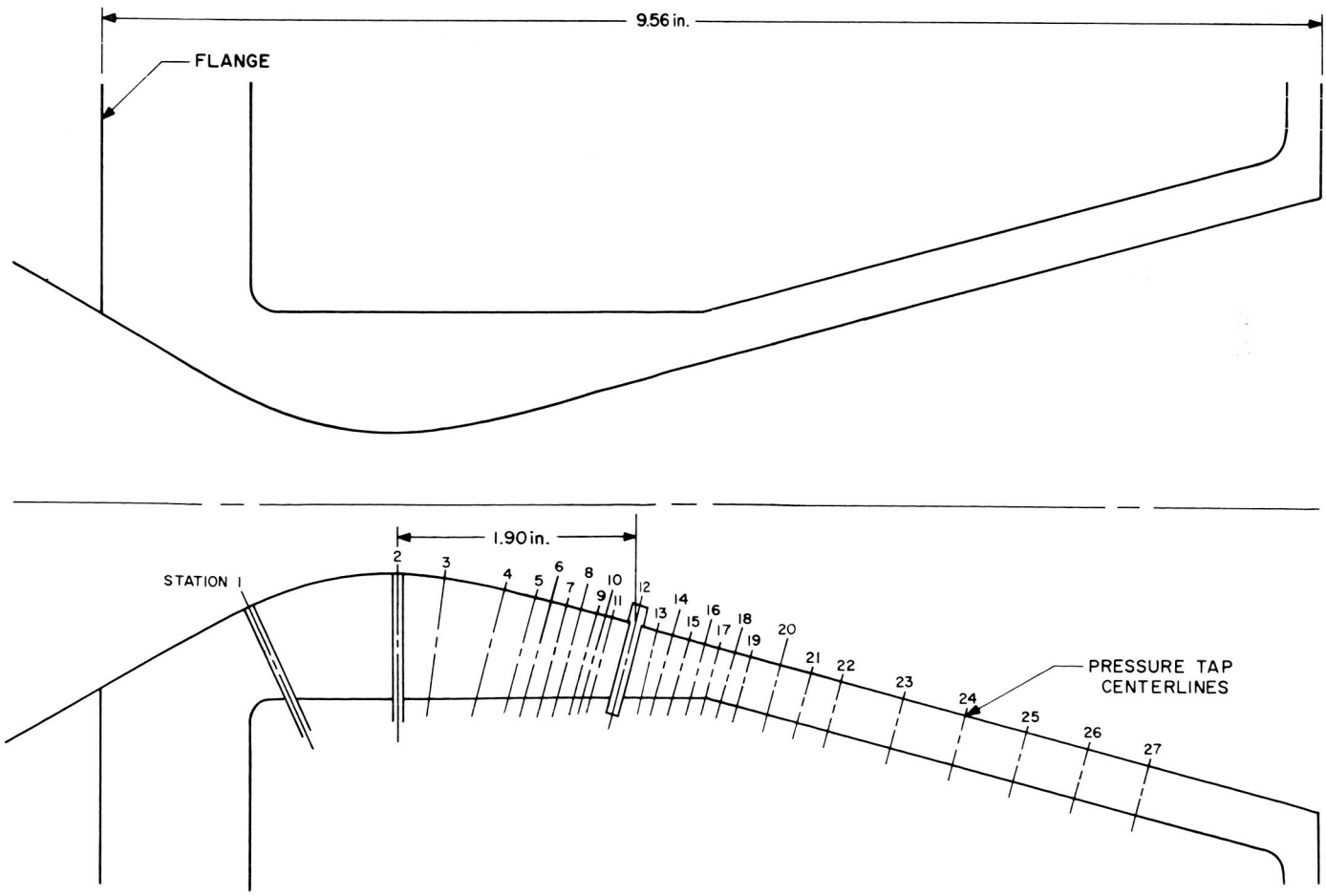


Fig. 8. Symmetric nozzle cross section

wall, the force unbalance normal to the nozzle axis could be obtained.

Both nozzles are conical with 30-deg entrance and 15-deg exit cones. One nozzle is axisymmetric. It has a throat radius curvature of 2.82 in. and a throat radius of 0.571 in., giving a curvature-to-throat radius ratio of 4.91. The exit diameter is 4.830 in. The nozzle has two leak-proof plugs, either of which mount in the nozzle wall 1.90 in. aft of the nozzle throat. The end of one plug lies flush with the nozzle surface. The second plug has a 0.125-in. square protrusion to the flow that extends 0.150 in. above the surface of the nozzle. The nozzle internal surface has 85 static pressure taps fabricated of 0.0625- and 0.093-in. OD stainless steel tubing. Fig. 8 is cross-sectional sketch along the pressure tap centerline, showing the axial spacing of the tap positions. Fig. 9 is an in-the-flat representation of the pressure tap spacing about the tap centerline. The protrusion, mounted in place, is shown in an aft end view of the nozzle (Fig. 10).

The second test nozzle was fabricated with a known throat region asymmetry or bias. One 180-deg section of the throat region is circular, with a radius at the design geometric throat of 0.526 in. The opposite 180-deg section is elliptical, with a fixed major axis of 0.705 in. (Fig. 11). The minor axis of the ellipse equals the radius of the opposite 180-deg section. The area centroid at the throat

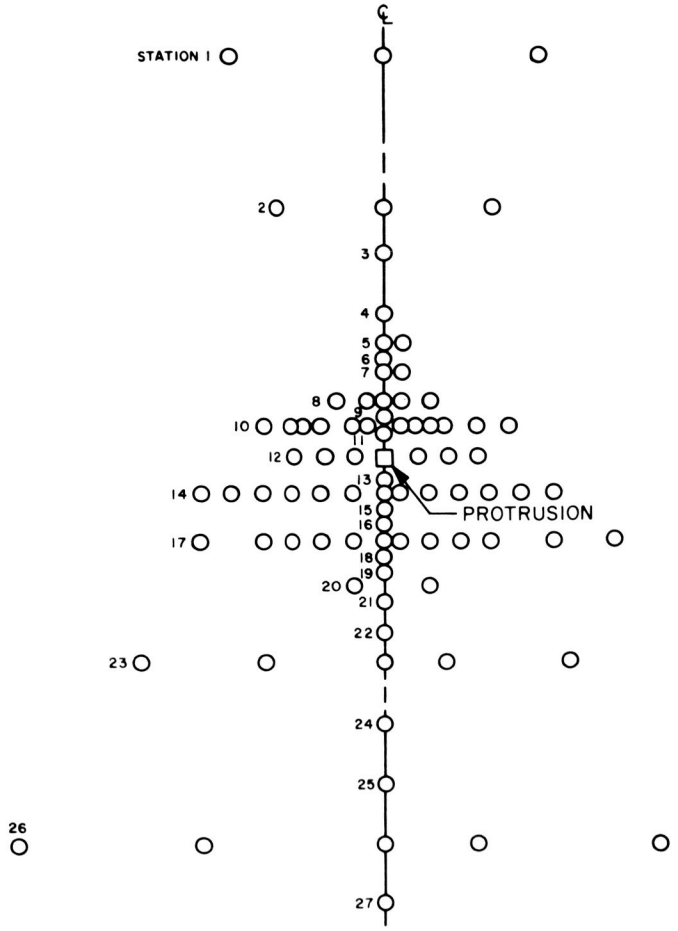


Fig. 9. Symmetric nozzle in-the-flat pressure tap spacing

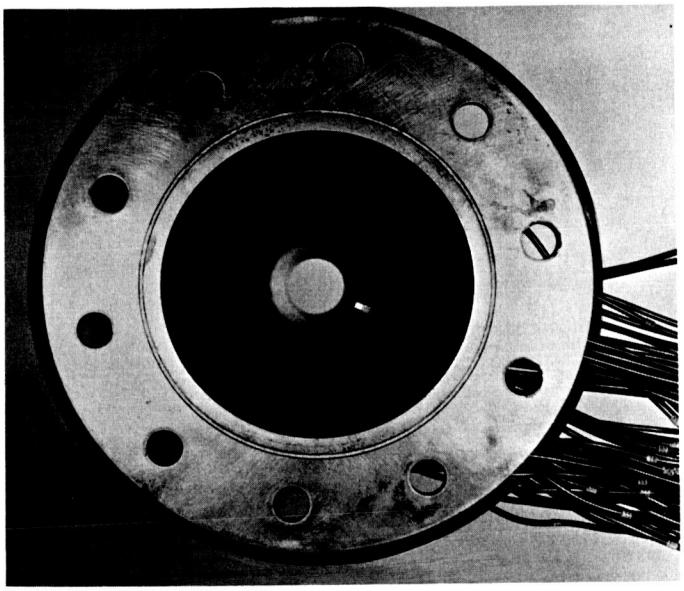


Fig. 10. Symmetric nozzle, aft view

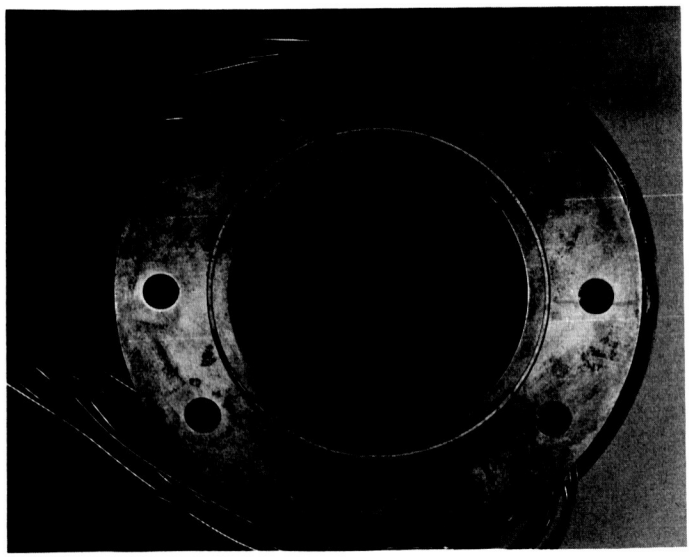


Fig. 11. Asymmetric nozzle, aft view

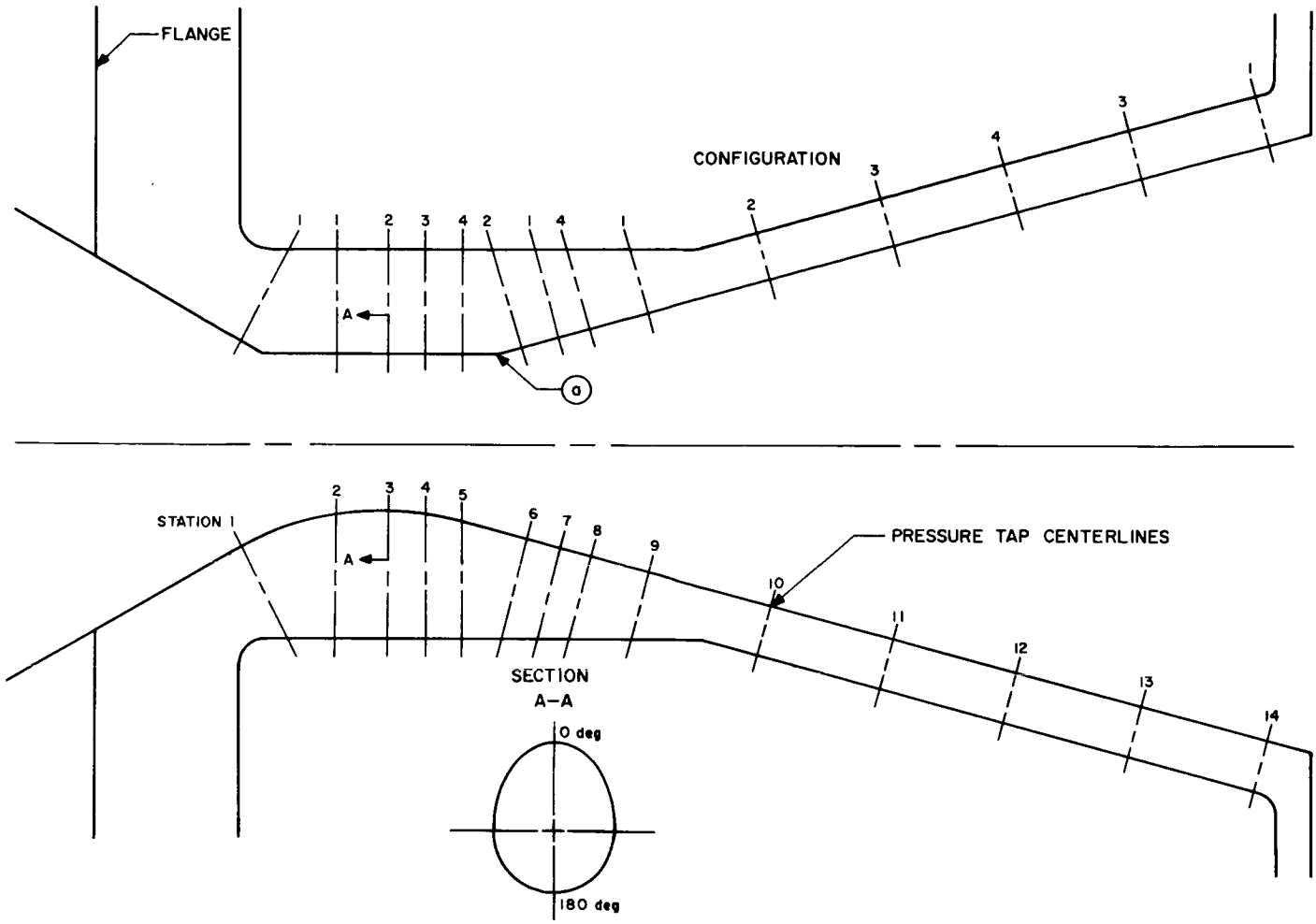


Fig. 12. Asymmetric nozzle cross section

is biased 0.073 in. from the nozzle centerline. At the nozzle axial positions of transition to symmetry on either side of the throat the major and minor axes of the ellipse become equal. The symmetrical throat region radiuses of curvature are 2.51 in. fore and 3.35 in. aft. This nozzle inner surface is instrumented with 100 static pressure taps, 0.093 in. OD tubing being used exclusively. Fig. 12 is a sketch of the nozzle cross section along the ellipse major axis, showing the axial spacing of the 14 pressure tap stations. One of the four cross-sectional tap configurations shown in Fig. 13 is used at each station.

4. Instrumentation

Static pressure measurements were made with the Section 373 multiple pressure measuring system (MPMS). The MPMS has been modified to withstand pressures greater than 600 psia. Each 12-port module of the system held 11 pressure tap lines and one reference pressure. Fig. 14 shows the symmetric nozzle mounted in the test assembly and hooked to the MPMS. Eight modules were used to measure the 85 static pressures plus the plenum supply pressure. The large range of pressures were measured by scanning the modules with 50, 100, 250, and 1000 psia pressure transducers. Reference pressures of 14.7, 100, and 200 psia were used. Likewise, nine modules were used for the 100 pressure port asymmetric nozzle. All transducer pressure counts data was converted by computer to millimeters of mercury pressure values and to several pressure ratios using the transducer calibration inputs. The plenum supply gas temperature was continuously monitored with a Chromel-Alumel thermocouple at 150°F reference temperature.

5. Test Procedure

The tests were grouped into three series. Series A tests were run with the symmetrical nozzle and no flow dis-

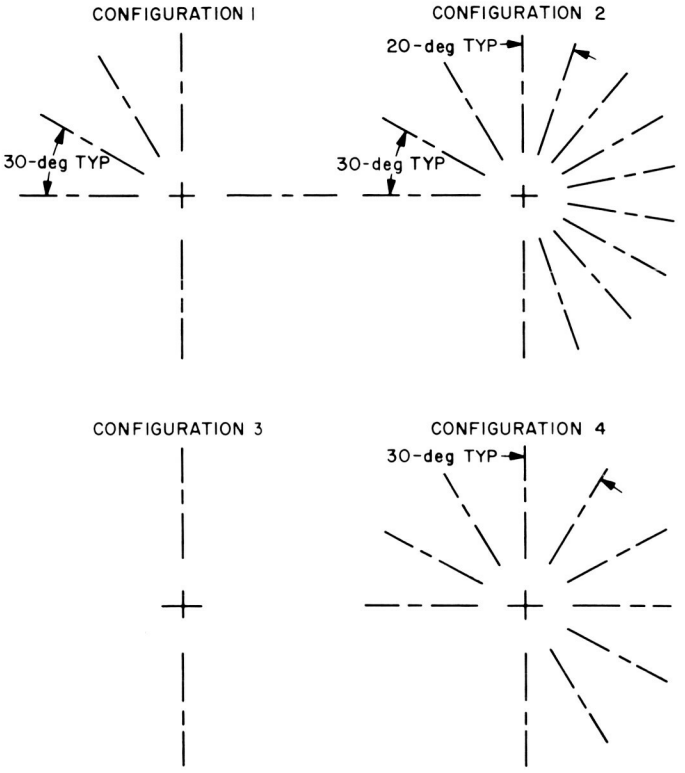


Fig. 13. Axial station pressure tap configurations asymmetric nozzle

turbance, Series B with the symmetrical nozzle with the flow disturbance, and Series C with the asymmetrical nozzle. The functions of the Series A tests were both to check the flow uniformity and instrumentation accuracy of the test system and to provide a base line for interpreting the results of Series B. Assurance of uniform flow with no vortexing action was obtained by rotating the nozzle in the test assembly approximately 90 deg from its normal position and repeating a few of the Series A runs. No significant changes in the wall static pressures were measured.

Each series of tests consisted of a test run at each of seven supply pressures: 100, 200, 300, 400, 450, 500, and 600 psia. The procedure for each run was to operate the tunnel compressor system at the required plenum pressure until the pressure stabilized, as indicated by the compressor operator's instrumentation. The MPMS system then made four scans of the nozzle static pressure taps, giving four sets of pressure data. The tunnel compressor output was then adjusted to each of the remaining required plenum pressures and the test procedure repeated. The supply air temperature was only raised to a high enough level (approximately 140°F) to ensure against air liquefaction in the nozzle expansion region.

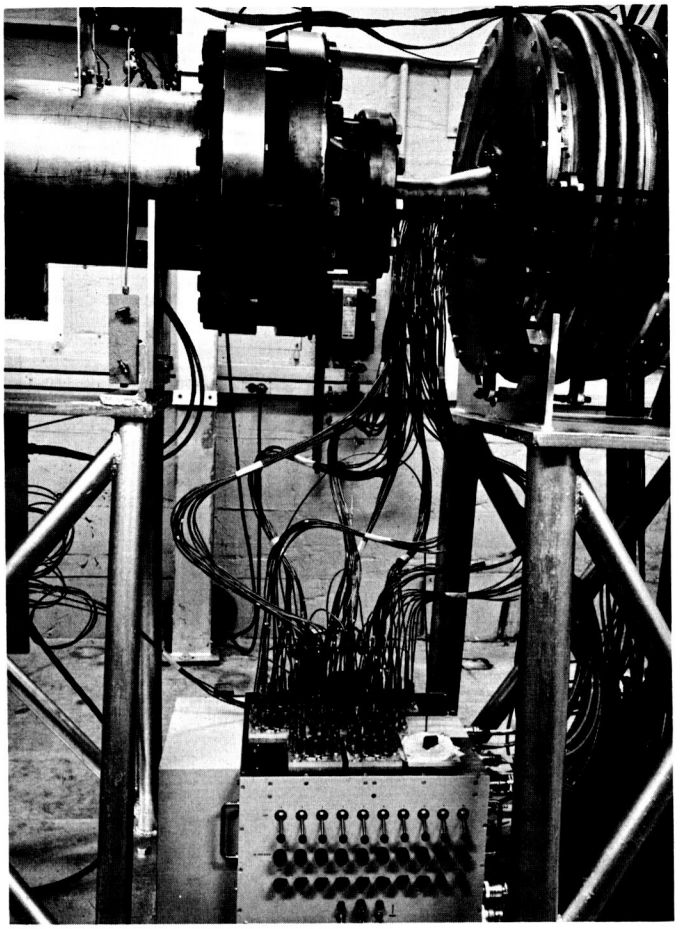


Fig. 14. Test assembly connected to MPMS

The air temperature was controlled by varying the compressor stage cooling.

During safety checkout of the entire auxiliary flow channel system the noise level in the wind tunnel operating area and outside the building in the region of the 24-in. return line became intolerably high. A temporary solution for these tests was obtained by insulating the expansion joint and 24-in. line inside the building (Fig. 15), barricading off the outside return line region, and locking all building doors exiting to that region.

6. *Test Results*

a. Series A. The Series A tests had to be run twice. Because of improper bonding of the chrome plating to the symmetrical nozzle surface, the plating became damaged when the protrusion for the Series B tests was installed in the nozzle. Subsequently, all of the plating was able to be peeled off, and the Series A tests were rerun.

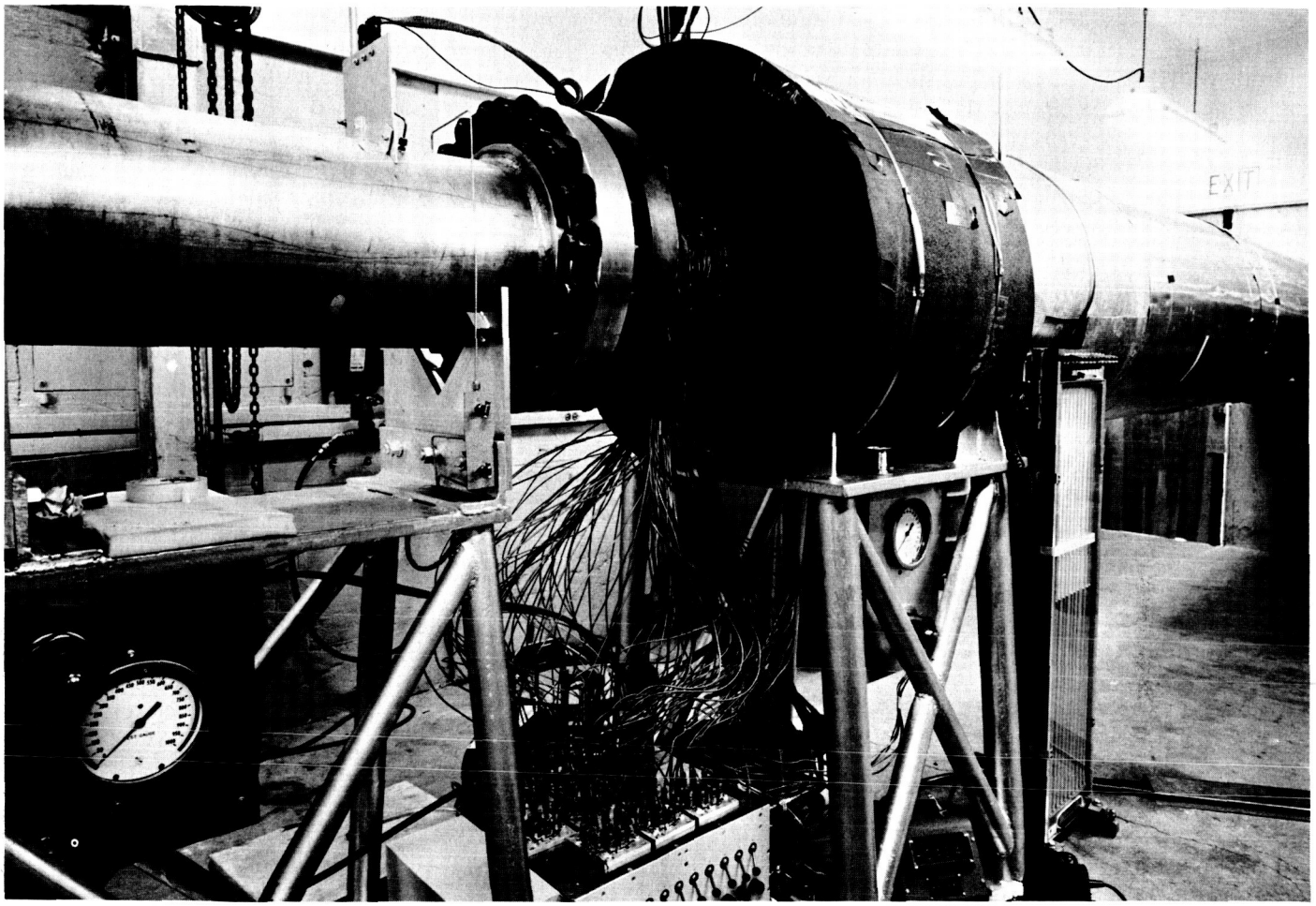


Fig. 15. Sound-insulated diffuser and return line

The performances of the auxiliary flow channel and the MPMS were very satisfactory. The compressors were able to maintain a diffuser back pressure from 1.2 psia at 600 psia down to 0.2 psia at 100 psia supply pressure. The nozzle flowed full at all supply pressures. The MPMS was randomly held at various pressure taps and the measurement compared with values obtained during scanning to ensure that the measurements were not being affected by the scan dynamics. Pressure measurement reproducibility over the four scans varied from within 0.5% at 600 psia supply pressure to 1% at 100 psia. With no air flowing through the system, atmospheric pressure measurement checks with the instrumentation system agreed with barometric values to 1% or better. These accuracies were considered adequate for these tests.

There were some slight variations in the pressures measured by the different taps at a given axial station. Such typical variation in pressure measurements at three axial stations is shown in Fig. 16. All of the pressures at

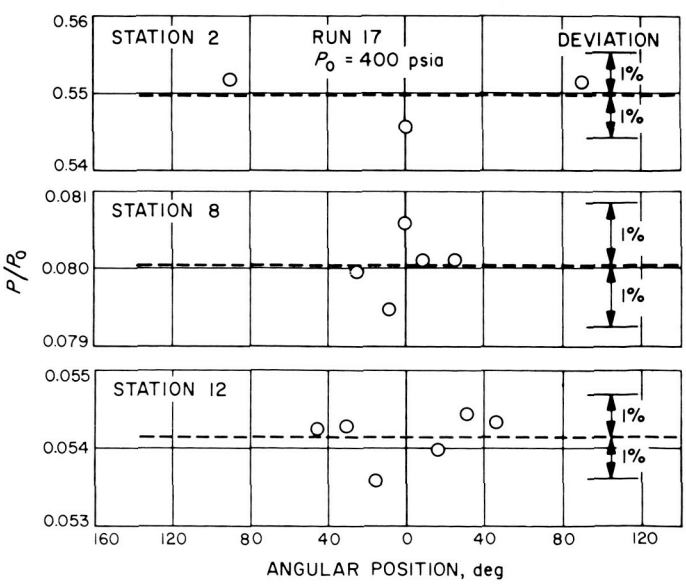


Fig. 16. Wall static pressure ratio versus nozzle angular position, symmetric nozzle

a given axial station fell within a ±1% band about the average pressure measured. This scatter could be attributed to slight variations in pressure tap axial location and surface geometry and to the pressure measurement errors previously mentioned.

In Fig. 17 the ratioed axial pressure data for a supply pressure P_0 of 400 psia is compared with a one-dimensional isentropic flow calculation for the symmetric nozzle. The measured pressure data is an average of the data from the four scans of each pressure tap. The measured data falls close to the one-dimensional theoretical curve, being slightly above the curve throughout most of the nozzle expansion as expected.

b. Series B. Fig. 18 shows the ratioed axial pressure data along the pressure tap centerline with and without the protrusion in place. The ratioed pressure data was essentially independent of supply pressure. The perturbed pressure ratio profile has characteristics similar to axial pressure profiles for nozzles with gaseous secondary injection. This similarity will be pursued. Fig. 19 shows the perturbed/unperturbed P_p/P_u pressure ratio axial profiles at angular cross sections 0, 10, 20, 30, and 40 deg from the pressure tap centerline. The curves were obtained by cross-plotting axial station pressure curves.

c. Series C. In Fig. 20 the ratioed axial pressure data for a supply pressure of 600 psia is shown for the angular positions designated in Fig. 12. The flow along the biased (0 deg) wall surface in the nozzle throat region expands more slowly than the flow along the symmetrical (180 deg) wall. At the aft axial position of transition to nozzle symmetry (Point ⓐ in Figs. 12 and 20) the flow along the biased wall surface rapidly expands. Downstream of Point ⓐ the two curves cross, and the flow from the biased throat region continues to expand more rapidly in the nozzle aft cone. The axial station pressure ratio profiles for the 400-psia run are shown in Fig. 21, the angular position designation shown in Fig. 12 again being used. As for the previous data, these pressure ratios also seem to be relatively independent of the supply pressure. Where data points were lacking, the curves were sketched in assuming asymmetry of flow in both 180-deg sectors. The nonsymmetry at Station 5 cannot be explained at this time.

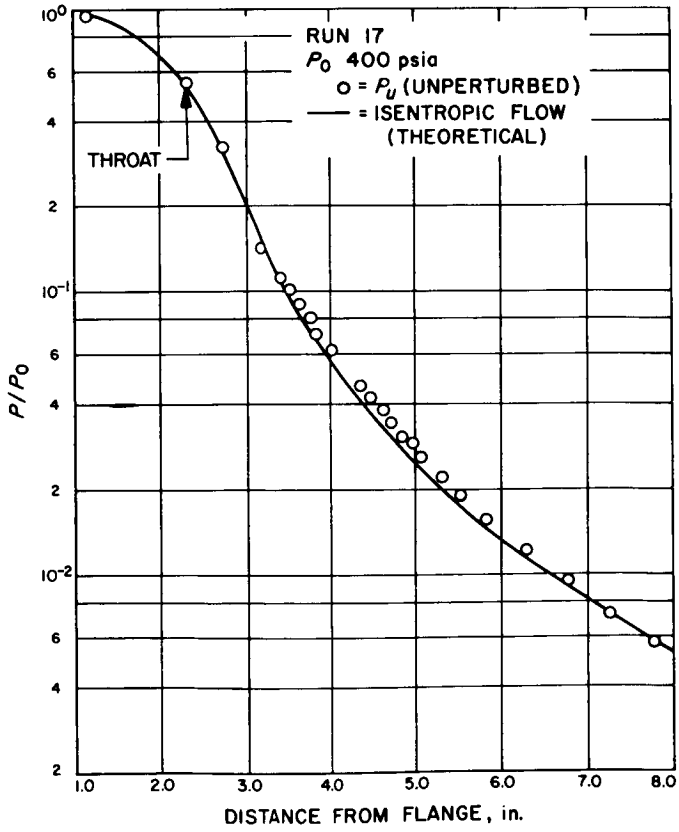


Fig. 17. Measured and theoretical wall static pressure ratios versus nozzle axial distance, symmetric nozzle

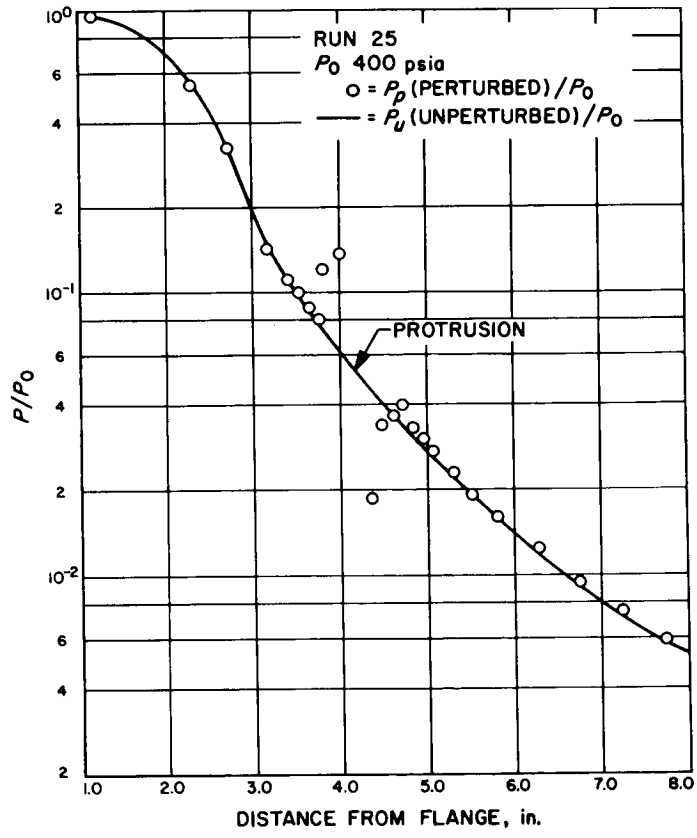


Fig. 18. Perturbed and unperturbed wall static pressure ratios versus nozzle axial distance, symmetric nozzle

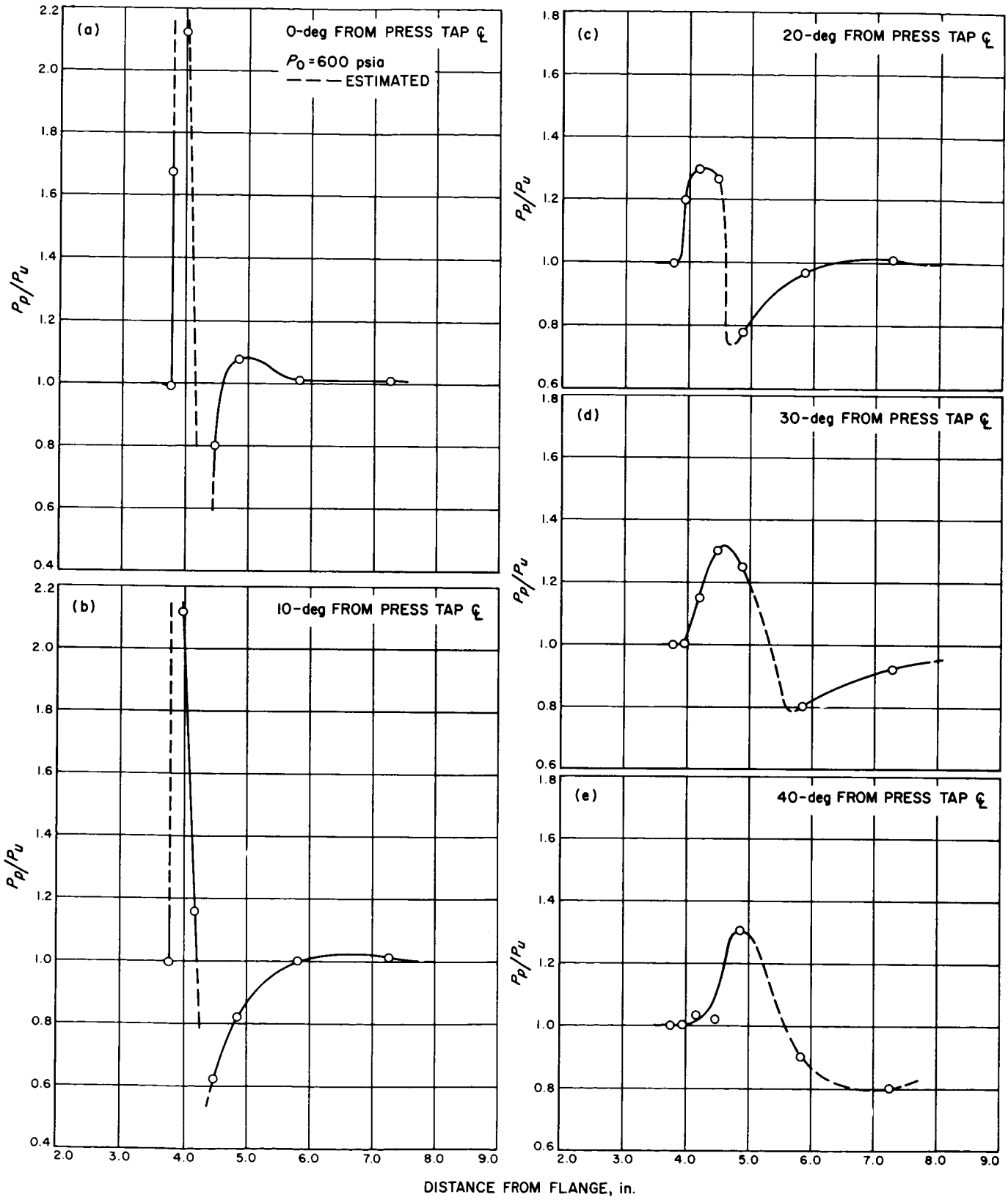


Fig. 19. Perturbed/unperturbed wall static pressure ratio versus nozzle axial distance, symmetric nozzle

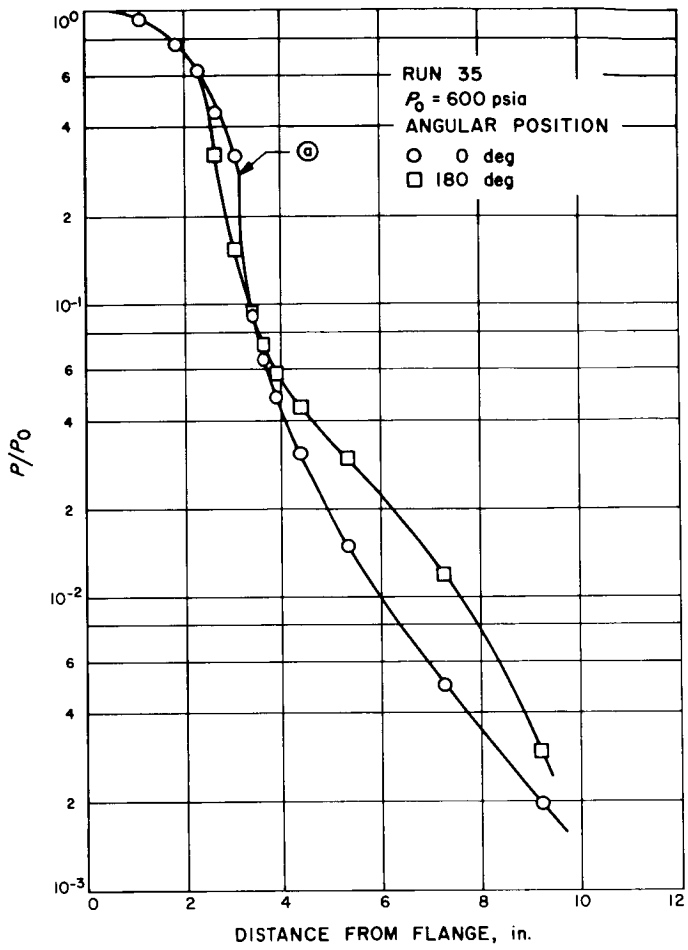


Fig. 20. Wall static pressure ratio versus nozzle axial distance, asymmetric nozzle

7. Discussion of Asymmetric Data

Because of the cancelling effect of the two wall pressure profiles shown in Fig. 20, it was not immediately obvious whether or not the asymmetric nozzle had any net side force. An analytical estimation of the side force resulted in a greater force unbalance in the throat region than in the conical expansion region. The net side force was slightly under 2% of the theoretical axial thrust.

The existence of the two opposed wall pressure regions might be due to the abrupt change in the wall contour at the transition point @ . A two-dimensional method of characteristics analysis of a similarly biased nozzle with smooth wall profiles yielded pressure ratio profiles similar to those in Fig. 20. Considering the side forces produced, the cancelling effect of the two opposed wall pressure regions is probably still present even for an asymmetric nozzle with smooth wall contours. This is one of the areas of analytical investigation being pursued.

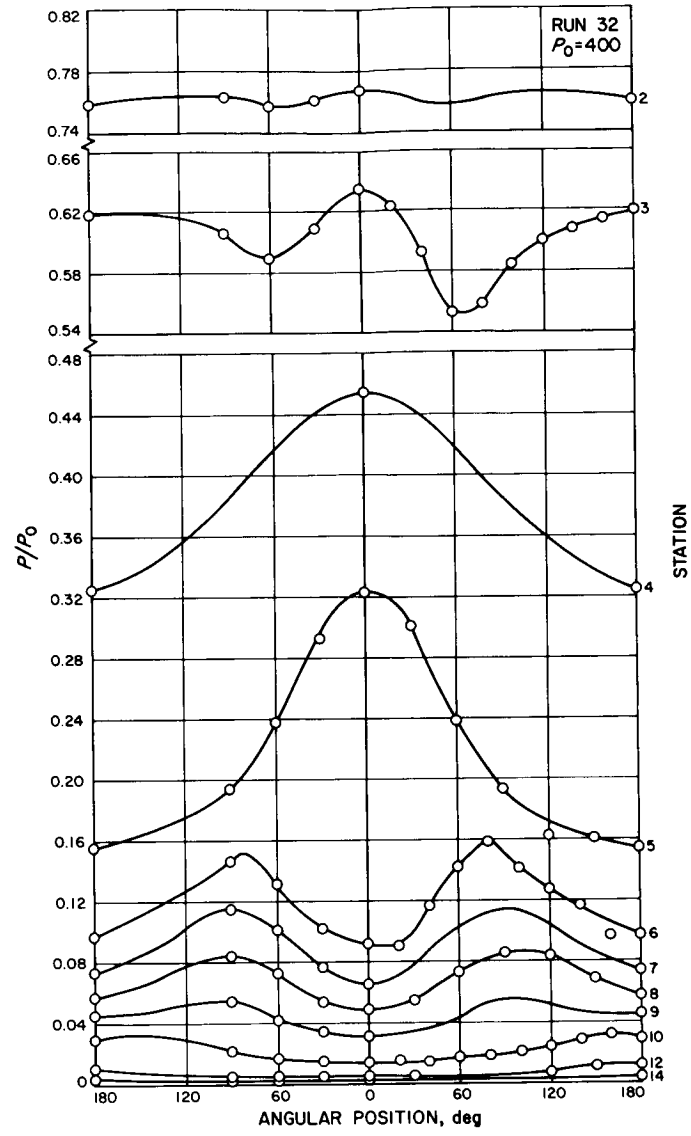


Fig. 21. Wall static pressure ratio versus angular position, asymmetric nozzle

B. Applications Technology Satellite (ATS) Motor Development

R. G. Anderson and D. R. Frank

1. Introduction

In January 1963 the Jet Propulsion Laboratory initiated a development program to provide a solid propellant apogee rocket motor for a second-generation *Syncom*

satellite. This program, under the management of the Goddard Space Flight Center, was designated *Advanced Syncom*. It was to result in a spin-stabilized, active repeater communications satellite weighing about 750 lb, operating at synchronous altitude (22,300 mi) which would handle voice communications, teletype, and monochrome and color television signals.

In January 1964 the *Advanced Syncom* communication program was redirected to include a number of experimental instruments in addition to the original communication instruments. This expanded program is the Applications Technology Satellite program and will result in a general-purpose satellite capable of operation at synchronous altitude with experimental instruments in the areas of meteorology, communications, radiation, navigation, gravity gradient stabilization and various engineering experiments. For those satellites to be placed in synchronous orbit, JPL will provide a solid propellant rocket motor to provide the final required velocity increment at the apogee of the elliptical transfer orbit. This rocket motor is designated the JPL SR-28-1 (steel chamber) or JPL SR-28-3 (titanium chamber) rocket motor. It is presently intended that only the JPL SR-28-3 unit will be delivered for flight use.

Previous reports of progress on the development of this motor have been published in *SPS 37-20* to *37-33*, Vol. V (Confidential) and *SPS 37-34*, Vol. IV.

2. Program Status Summary

The motor development program calls for static firing of four heavywall motors and 25 flightweight motors, including two with flight design titanium chambers, prior to conducting a nine-motor qualification program. To date, the four heavywall motors plus 18 flightweight motors have been static-fired, four of which were under simulated high-altitude conditions at Arnold Engineering Development Center (AEDC) Tullahoma, Tennessee. All of the flightweight motors tested to date have been with Type 410 chromium steel chambers.

During the period August to September 1965, two apogee motor units were static-fired at ETS and one inert loaded motor was delivered to Hughes Aircraft Company.

3. Inert Motor Delivery

Inert loaded motor T-2 (Code Dev. D-5T) was delivered to Hughes Aircraft Company for use as a dynamic test model. This unit represents the current ATS apogee motor using a titanium motor chamber. Combined spacecraft-apogee motor dynamic tests will be conducted at Hughes during the latter part of this year. Pertinent data for this motor are summarized in Table 1.

Table 1. ATS JPL SR-28-3 apogee motor data summary (Dev. D-5T)

Identification	Part name	Serial number
Chamber	J 390 1790B	T-2
Insulation	J 390 1797 N/C	T-2
Nozzle	J 390 1659 D	F-22
Igniter	—	—
Propellant	SYN D-9 (inert)	Batch SY 267

Weight data, lb	
Chamber (titanium)	25.00
Insulation	11.38
TDI (insulation rinse)	0.34
Nozzle	39.36
Igniter	not installed
Balance weights	no balance
Miscellaneous (nozzle bolts, washers, O-ring)	0.35
Total inerts (less igniter and nut)	76.43
Propellant	756.0
Total assembly (less igniter and nut)	832.4

Measurement data	
Thrust misalignment	0.0001 in./in.
Thrust offset at attachment	0.0022 in.
Attachment skirt	
Flatness, TIR	0.0004 in.
Out of round, TIR	0.0006 in.

	Empty assembly[a]	Loaded assembly[b]
Center of gravity (distance from edge of skirt)	21.06 in.	11.59 in.
Pitch moment of inertia (about CG)	26,500 lb-in.^{2}	101,400 lb-in.^{2}
Roll moment of inertia	7,910 lb-in.^{2}	82,400 lb-in.^{2}

[a]Includes igniter closure, nut and squib, but not igniter basket or charge.
[b]Includes igniter closure, nut, basket, charge and ignition squib.

4. Environmental Test Program

The environmental test program consists of nine motor assemblies which will be subjected to various combinations of environments prior to static testing for confirmation of design integrity. The first seven motors utilized steel chambers and these units have been cast and static-fired. The last two of the nine units will utilize titanium chambers. The final test in this series (G-9T) will be subjected to all environments and static-tested at + 10°F. A summary of environmental program and current status is given in Table 2.

5. Apogee Motor Ballistic Summary

Table 3 lists the apogee motor ballistic data obtained from all flightweight motor firings. This table includes 14 tests conducted at ETS and 4 units fired under simulated high altitude at AEDC. Tests at ETS are arranged by test temperatures, and the AEDC tests are summarized as separate units in a single block.

6. Chamber Insulation Evaluation

Development round C-7 temperature conditioned to 110°F, was tested August 10, 1965. This unit's prime purpose was evaluation of chamber insulation design for the titanium motor chamber. Twenty-four external chamber thermocouples monitored chamber temperature during and immediately after the run. This test employed a Type 410 chromium steel chamber. Fig. 22 shows the thermocouple placement. Figs. 23, 24, 25, and 26 show the external chamber temperatures.

The criteria for evaluating the titanium chamber insulation configuration is based on the margin of safety between the actual working stress in the chamber and the titanium yield strength at the operating temperature. Fig. 27 depicts a typical chamber temperature, chamber stress, and titanium yield strength curves for Test C-7. Thermocouple B-4, which recorded maximum temperature, was used for the above analysis. Fig. 27 shows that

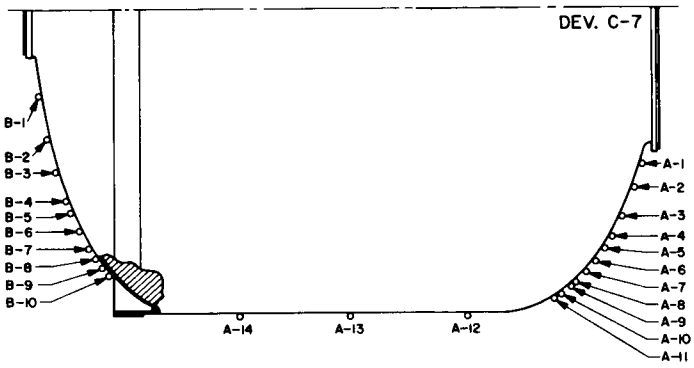


Fig. 22. Chamber thermocouple placement

Table 2. ATS apogee motor environmental test program summary

Test code	Motor number	Environmental testing					Firing temperature, °F	Current status July 30, 1965
		Temperature cycling	Shipping temperature exposure	Centrifuge	Vibration	Spinning		
G-1	P-11	√	√				110	Static firing successful
G-2	P-10	√	√				10	
G-3	P-14					√	60	
G-4	P-15				√	√	60	
G-5	P-16					√	60	
G-6	P-19	√		√	√	√	110	
G-7	P-20	√		√	√	√	10	
G-8T	Titanium chamber						110	Chamber fabricated, received at JPL
G-9T	Titanium chamber	√		√	√	√	10	Chamber fabricated, received at JPL

Code No.	Test Conditions				
	Type of test	Test location	Test date	Run No.	Grain condition temperature, °F
DEV. C-1	Atmospheric	ETS E60	2-26-64	E-127	60
DEV. C-2	Atmospheric	↓	3- 5-64	E-129	↓
DEV. C-3	Diffuser		6- 4-64	E-174	
DEV. C-3A	Diffuser		6-12-64	E-175	
DEV. C-6	Atmospheric		7- 9-64	E-416	
DEV. G-5	Spin		11-24-64	E-326	
DEV. C-4	Atmospheric	↓	7-15-64	E-218	10
DEV. H-1	Atmospheric		11- 4-64	E-314	↓
DEV. G-2	Atmospheric		12-16-64	E-343	
DEV. G-7	Spin		8- 6-65	E-434	
DEV. G-9T	Spin				
DEV. C-5	Atmospheric	↓	7-23-64	E-237	110
DEV. C-7	Atmospheric		8-10-65	E-435	↓
DEV. G-1	Atmospheric		12-16-64	E-342	
DEV. G-6	Spin		7-16-65	E-431	
DEV. G-8T	Atmospheric				
DEV. E-1	Diffuser	AEDC	7-16-64	1426-01	60
DEV. E-2	Diffuser	↓	10-13-64	1511-01	60
DEV. G-3	Diffuser and spin		5-28-65	1540-02	75
DEV. G-4	Diffuser and spin		5-26-65	1540-01	75

[a]Nozzle expelled approximately 17 sec following motor ignition, voiding all performance da
[b]Data unavailable because of electrical noise during ignition phase of motor.
[c]Thrust measurement was not made during spin tests.
[d]Vacuum correction does not apply because of incomplete expansion of exhaust gases.
[e]Data unavailable due to malfunction of oscillograph recorder.
[f]For definitions of motor ballistic parameters, see Table 4.

Table 3. ATS apogee motor static test

Physical hardware					Pressure			
Propellant weight, lbm	Chamber No.	Nozzle No.	Igniter	Squib	Characteristic velocity $\bar{W}^*$ ft/sec	Chamber pressure integral, psia-sec	Igniter basket peak pressure, psia (msec)	Chamber ignition peak pressure psia (msec)
760.8	P-2	F-13 $\varepsilon=8.5$	SYC-231	28-9	4981.3	8918.0	1997 (52)	255 (63)
761.3	P-7	F-7	SYC-233	10	4982.0	8921.1	1620 (50)	173 (60)
762.2	P-7F	F-6	SYC-234	11	[a]	[a]	1828 (59)	249 (71)
765.8	P-2F	F-5	SYC-235	12	4972.6	8952.2	2289 (48)	270 (63)
765.7	P-24	F-30	SYC-255	30	4957.8	8951.2	1950 (22)	269 (29)
768.7	P-16	F-14	SYC-245	28	4964.6	9002.8	1930 (33)	225 (42)
764.8	P-13	F-11	SYC-241	18	4960.3	8944.5	1853 (34)	228 (45)
640.4	P-5	F-10	SYC-244	27	4965.6	7514.3	2005 (20)	224 (40)
766.1	P-10	F-17 $\varepsilon=35$	SYC-252	58	4951.2	8945.5	1527 (22)	212 (32)
769.6	P-20	F-19 $\varepsilon=35$	SYC-263	60	4958.6	9019.7	1650 (14)	217 (23)
769.3	P-12	F-12 $\varepsilon=8.5$	SYC-243	20	4972.3	9012.2	3292 (21)	273 (29)
767.0	P-22	F-29 $\varepsilon=8.5$	SYC-265	20-9	4982.4	8999.2	3000 (12)	320 (16)
766.8	P-11	F-16 $\varepsilon=35$	SYC-248	38	4975.3	8997.2	[e]	[e]
767.2	P-19	F-21	SYC-262	59	4978.5	9021.9	3485 (16)	276 (23)
765.6	P-9	F-9	SYC-240	6	4982.9	8960.4	1524 (59)	239 (69)
764.4	P-18	F-15	SYC-242	26	4984.4	8937.8	1832 (38)	233 (49)
765.3	P-14	F-24	SYC-260	36	4965.2	8970.1	1626 (33)	241 (44)
768.6	P-15	F-20	SYC-261	31	4985.1	9042.7	1826 (32)	233 (45)

data[f] summary, development phase

		Time			Thrust			
Chamber starting pressure psia (sec)	Chamber run peak pressure, psia (msec)	Ignition delay t_{D_m}, msec	Action time t_R, sec	Run time t_R, sec	Peak vacuum thrust, lbf (sec)	Total measured impulse, lbf-sec	Vacuum total impulse, lbf-sec	Measured specific impulse, lbf-sec/lbm
103.9 (0.23)	265.2 (32.1)	34	42.31	42.42	5888 (32.9)	138,450	197,640	182.0
98.6 (0.27)	260.6 (33.0)	29	42.96	43.04	5746 (33.0)	137,560	196,580	180.7
115.2 (0.22)	[a]	40	[a]	[a]	[a]	[a]	[a]	[a]
99.6 (0.21)	264.4 (34.1)	35	42.76	42.89	5766 (34.2)	181,110	193,320	236.5
96.1 (0.25)	254.8 (33.7)	11	44.08	44.18	5686 (33.8)	138,710	199,620	181.1
98.6 (0.25)	258.9 (33.0)	[b]	[b]	[e]	[c]	[e]	[c]	[c]
98.0 (0.25)	252.3 (34.6)	17	44.45	44.52	5563 (33.3)	137,560	197,510	179.9
191.1 (0.36)	264.2 (20.3)	12	31.74	31.81	5774 (21.3)	117,080	163,930	182.8
92.4 (0.25)	242.7 (36.0)	2	45.48	45.74	[d]	127,400	[d]	166.3
99.3 (0.27)	245.7 (33.9)	[b]	[b]	45.09	[c]	[e]	[c]	[c]
98.8 (0.21)	274.6 (32.8)	9	41.68	41.76	6034 (32.8)	140,850	198,100	183.1
96.2 (0.20)	272.3 (33.4)	3	42.35	42.57	6018 (33.4)	140,690	200,080	183.4
99.4 (0.21)	277.4 (33.0)	[e]	[e]	[e]	[d]	131,360	[d]	171.3
100.4 (0.23)	268.5 (30.3)	[b]	[b]	41.74	[c]	[c]	[c]	[c]
98.1 (0.24)	263.4 (34.0)	26	42.93	43.42	6338 (34.2)	205,060	214,740	267.9
89.9 (0.20)	261.5 (34.6)	27	43.32	43.59	6319 (34.0)	210,730	214,360	275.7
99.5 (0.25)	261.7 (32.3)	16	42.52	43.09	6233 (33.0)	212,190	214,760	277.3
96.6 (0.25)	261.0 (33.1)	22	42.99	43.56	6226 (33.2)	213,240	215,740	277.4

2

3 ##

	Nozzle				
Vacuum specific impulse, lbf-sec/lbm	Nozzle throat diameter, in.			Nozzle exit initial diam., in.	Nozzle throat erosion, %
	Initial	Final	Average		
259.8	4.083	4.123	4.103	11.907	1.97
258.2	4.083	4.125	4.104	11.907	2.07
[a]	4.083	[a]	[a]	11.906	[a]
252.4	4.083	4.127	4.105	11.906	2.17
260.7	4.084	4.114	4.099	11.907	1.47
[c]	4.083	4.113	4.098	11.904	1.48
258.2	4.084	4.114	4.099	11.905	1.47
256.0	4.084	4.105	4.094	11.904	1.03
[d]	4.083	4.115	4.099	24.135	1.57
[c]	4.082	4.105	4.094	24.126	1.13
257.5	4.083	4.118	4.100	11.907	1.72
260.9	4.083	4.120	4.101	11.908	1.82
[d]	4.083	4.114	4.098	24.137	1.52
[c]	4.082	4.108	4.095	24.127	1.28
280.2	4.085	4.124	4.105	24.134	1.92
280.2	4.090	4.124	4.107	24.135	1.67
280.6	4.083	4.110	4.096	24.135	1.33
280.7	4.082	4.112	4.097	24.106	1.48

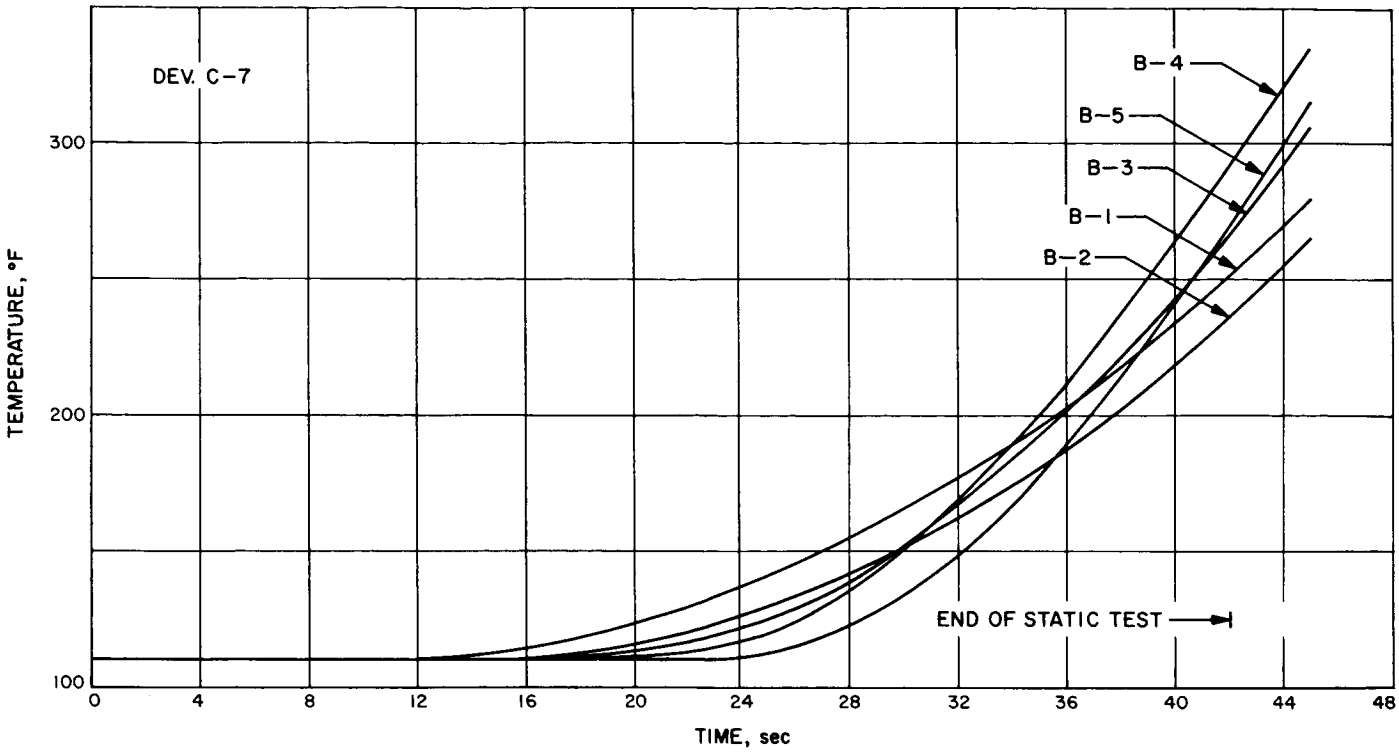


Fig. 23. Chamber temperature versus time, Thermocouples B-1 through B-5

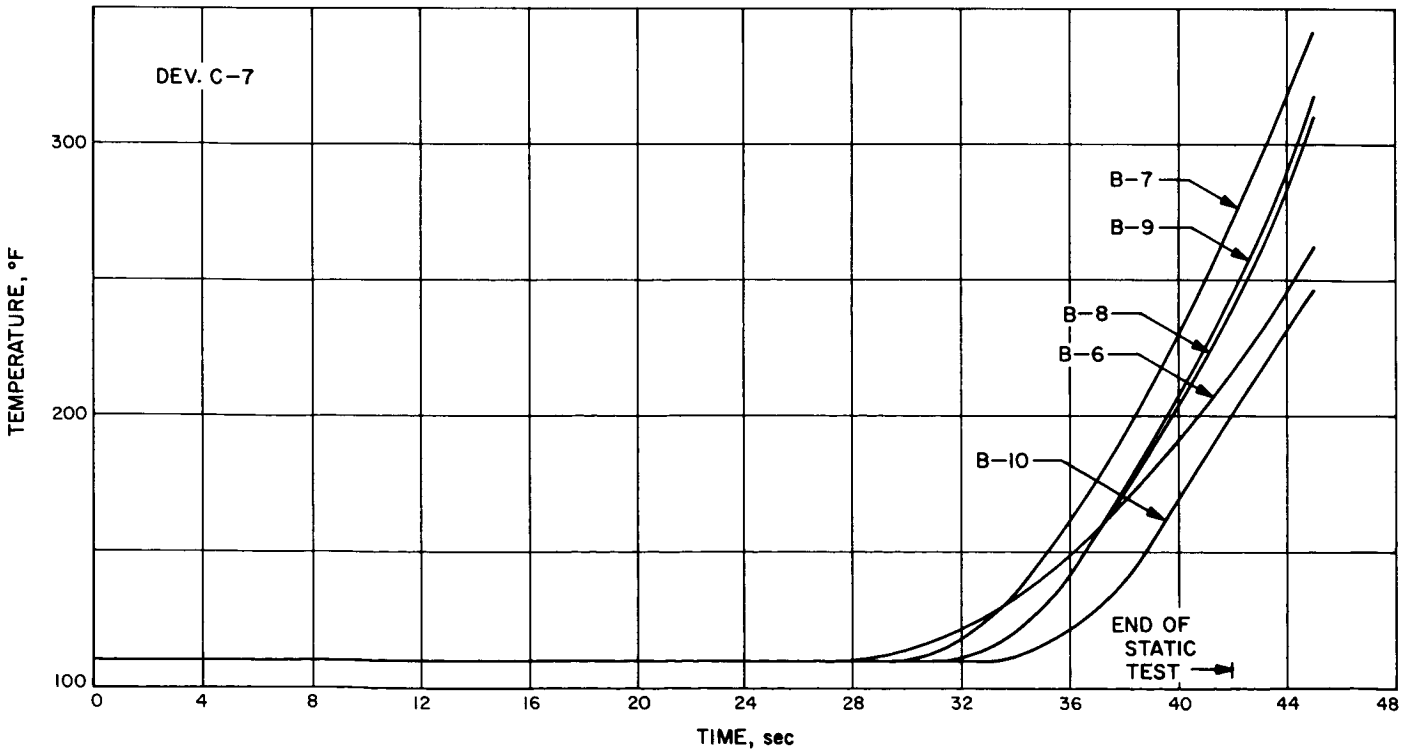


Fig. 24. Chamber temperature versus time, Thermocouples B-6 through B-10

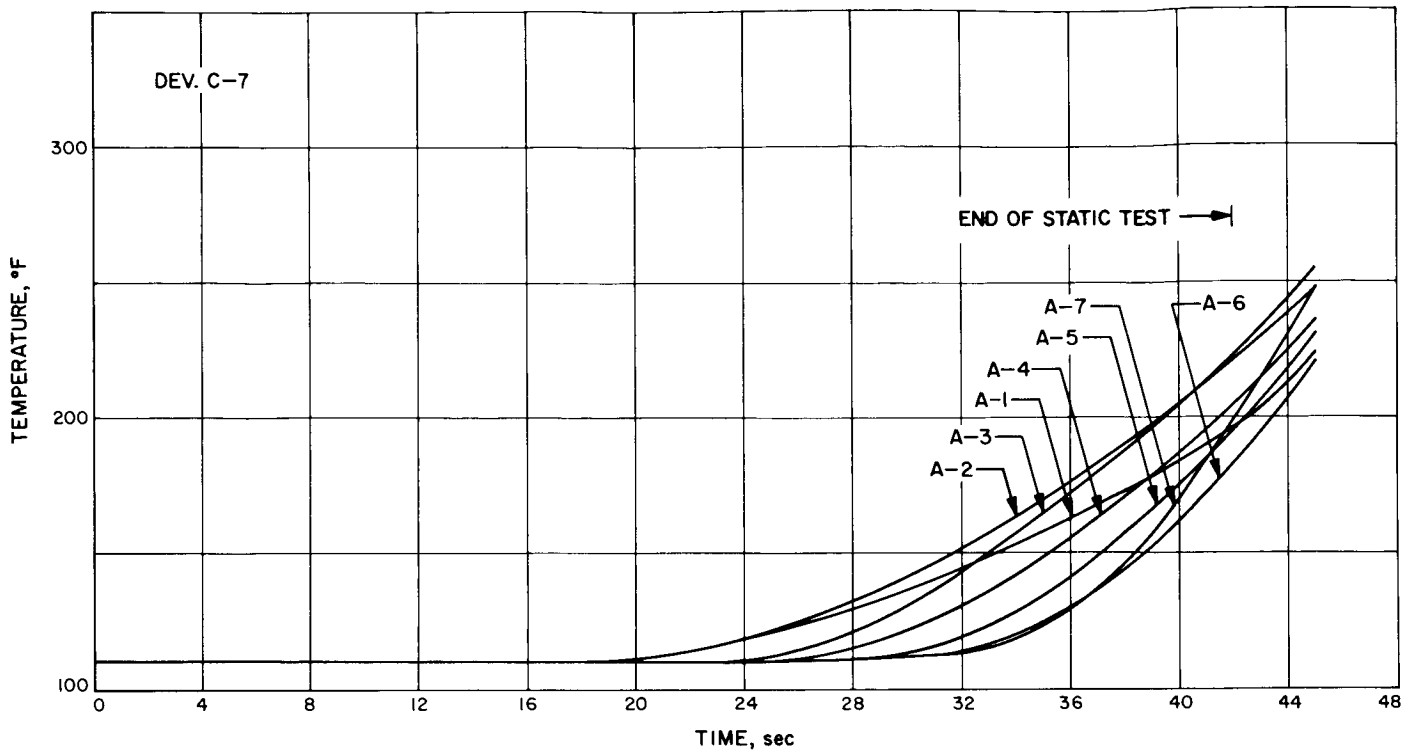


Fig. 25. Chamber temperature versus time, Thermocouples A-1 through A-7

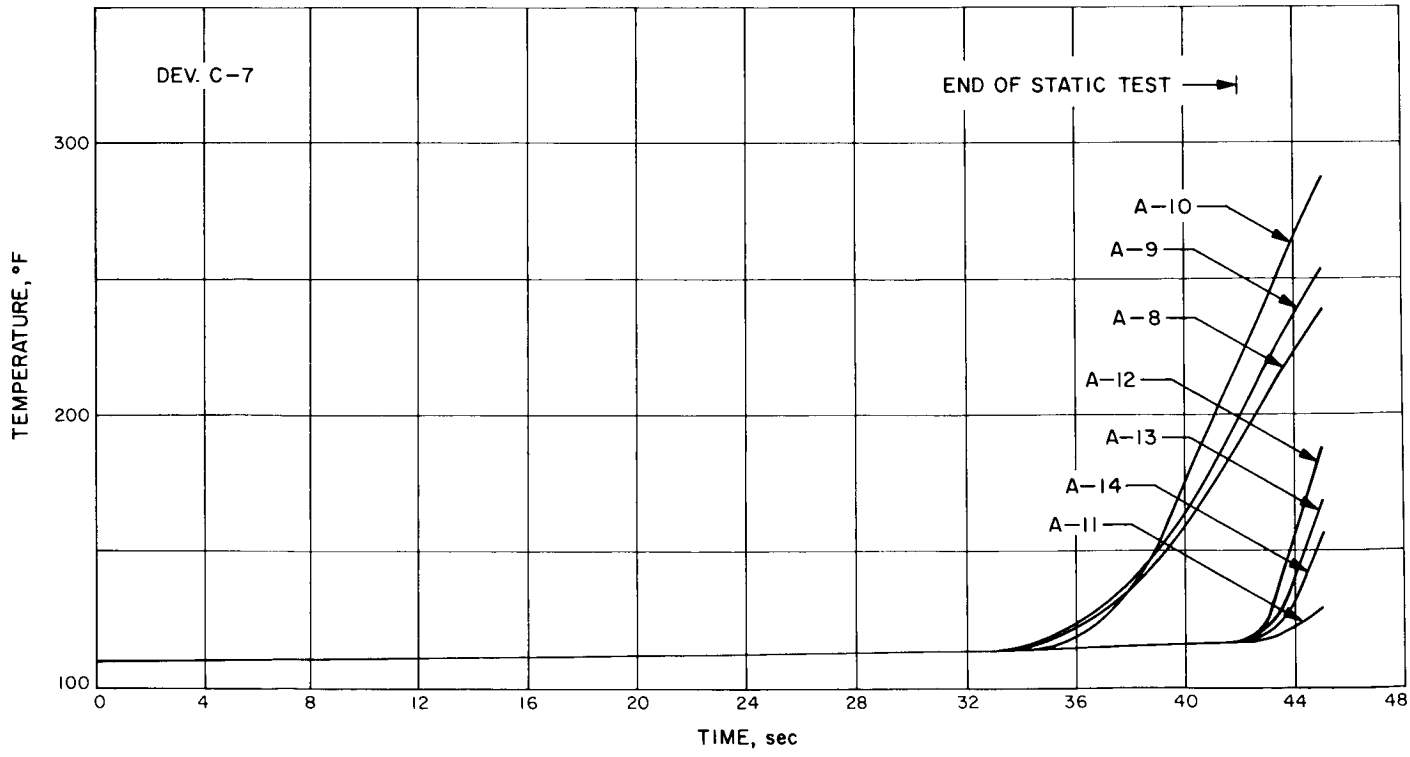


Fig. 26. Chamber temperature versus time, Thermocouples A-8 through A-14

Table 4. Parameter definitions used in ATS static test data summary

(1) $\bar{W}^*$, ft/sec, $\dfrac{g_0 A_t \int_{t_0}^{t_r} P_c dt}{W_p}$

where

g_0 = gravitational constant (32.14 lbm-ft/lbf-sec^2)

A_t = average nozzle-throat area (based on the average of the prefire and postfire measurements)

W_p = loaded propellant weight

t_r = run time (Item 10)

t_0 = zero time (Item 7)

(2) Chamber pressure integral, psia-sec, $\int_{t_0}^{t_r} P_c \, dt$

(3) Igniter basket peak pressure, psia, P_{I_i} (Fig. 28)

(4) Chamber ignition peak pressure, psia, P_{I_m} (Fig. 28)

(5) Chamber starting pressure, psia, P_c (start) (Fig. 28)

(6) Chamber peak pressure, psia, P_c (peak) (Fig. 29)

(7) Zero time, sec, t_0, t_0 is the instant electrical current is supplied to the squib (Fig. 28)

(8) Ignition delay, msec, t_{D_m}, t_{D_m} is the time period between the application of electrical current to the squib until the first indication of motor-chamber pressure (Fig. 28)

(9) Action time, sec, t_a, t_a is the time period between the time the motor-chamber pressure has reached 10% of its run peak value at ignition until it has decreased to 10% of the run peak value during motor tailoff (Fig. 29)

(10) Run time, sec, t_r, t_r is the time period between the application of electrical current to the squib until the motor chamber pressure has returned to ambient conditions following motor tailoff (Fig. 29)

(11) Peak vacuum thrust, lbf, F_{vac} (peak), measured at the same time as the chamber peak pressure value (Fig. 29)

(12) Total measured impulse, lbf-sec, $I_{(mea)} = \int_{t_0}^{t_r} F_{(mea)} \, dt$

(13) Vacuum total impulse, lbf-sec, $I_{(vac)} = \int_{t_0}^{t_r} F_{(vac)} \, dt$

(14) Measured specific impulse, lbf-sec/lbm, $I_{s(mea)} = \dfrac{\int_{t_0}^{t_r} F_{(mea)} \, dt}{W_p}$

(15) Vacuum specific impulse, lbf-sec/lbm, $I_{s(vac)} = \dfrac{\int_{t_0}^{t_r} F_{(vac)} \, dt}{W_p}$

(16) Nozzle throat erosion, %, $\dfrac{D^2(\text{final}) - D^2(\text{initial})}{D^2 \text{ (initial)}} \times 100$

where D = the average nozzle throat diameter

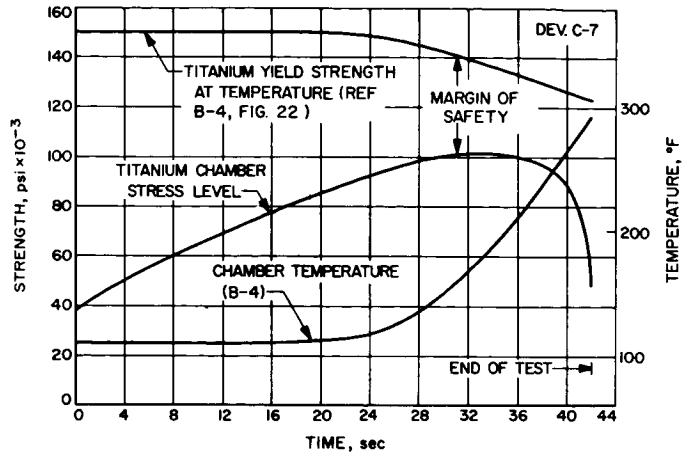


Fig. 27. Chamber temperature, chamber stress, and titanium yield strength

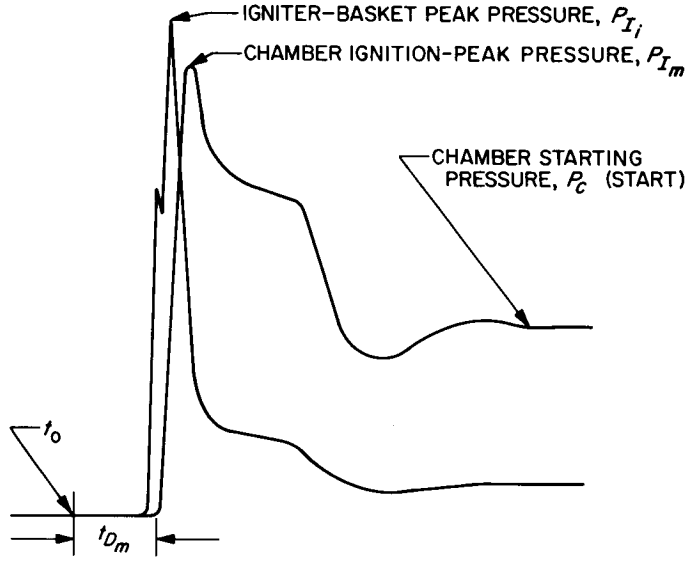


Fig. 28. Measurements of ignition phase of motor operation

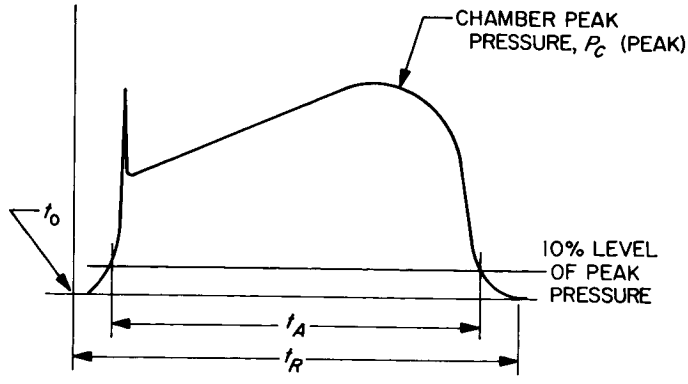


Fig. 29. Measurements of entire motor test

the insulation configuration employed in this test provides adequate chamber temperature control during the motor burning phase. Actual stress levels were kept well below the minimum yield strength for titanium.

Development unit G-8T will be the initial titanium motor firing. This unit will employ chamber thermocouples as a recheck of the insulation design. This test is scheduled for November of this year.

XIII. Liquid Propulsion

A. Advanced Liquid Propulsion Systems

R. N. Porter, D. D. Evans, H. B. Stanford, and W. F. MacGlashan

1. Introduction, R. N. Porter

The Advanced Liquid Propulsion Systems (ALPS) program is investigating selected problems generated by spacecraft operational requirements for propulsion systems capable of high inherent reliability, long-term storage in space, multiple start in free-fall (zero-gravity), and engine throttling. The solutions proposed to satisfy these requirements have been incorporated into the ALPS system.

Periodic reports in the *SPS*'s (starting with No. 37-8) describe the progress of work on the various parts of the ALPS system. Recent accomplishments are outlined below. These include three tests of the ALPS pressurization subsystem, optimization of a 25-lb-thrust impinging sheet injector element, temperature cycling of a tank containing hydrazine to simulate heat sterilization procedures, and further development work on a simple fill valve.

2. System Development, R. N. Porter

The first three in a continuing series of tests of a breadboard version of the ALPS pressurization subsystem have been made. The purposes of these tests are to learn as much as possible, within the program limitations, about the stability, response, and general behavior of the circuit, to detect troublesome interactions between the component parts, and to gain test experience with the components performing their specific functions within a working subsystem which simulates part of the ALPS system. A previous series of tests was accomplished in 1961 with a system composed of interim components. Those tests adequately confirmed the general feasibility of the ALPS pressurization scheme but told little beyond that, since none of the ALPS components were available for testing then. It is hoped that the current tests will be a more definitive demonstration of what the actual ALPS system performance would be, using the presently existing ALPS components.

In the ALPS system the fuel and oxidizer are stored within two positive expulsion devices (bladders or diaphragms) located in a single spherical tank. Gas pressure in the tank collapses the expulsion devices to expel the

propellants and to force them into the engine. Valves in the engine injector vary the propellant flow rate, metering it as necessary to achieve the desired thrust, as well as acting as leak-tight shutoff devices between firing periods.

The gases which pressurize the propellant tank are generated by the catalytic decomposition of hydrazine in a small reactor called a "gas generator." Because the gaseous mixture of nitrogen, hydrogen, and ammonia flowing from the gas generator is too hot (approximately 1500°F) to be pumped directly into the propellant tank, it is first cooled by two heat exchangers (one cooled by the fuel and the other by the oxidizer flow to the engine) so that the temperature on entering the tank is no greater than 150°F.

A "generant controller" meters the flow of liquid generant (hydrazine) into the gas generator so that the rate of gas generation is just sufficient to maintain a constant predetermined pressure in the propellant tank during all firing periods, regardless of whatever variations may occur in the rates at which the propellants flow out of the tank or in the hydrazine source pressure. This controller is essentially a remote-sensing, single-stage, spring-loaded regulator. Its operation is strictly mechanical and automatic. An electrically operated *on-off* valve located between the controller and the gas generator is used for leak-tight shutoff of the hydrazine during nonfiring periods.

The hydrazine flow originates from the "generant tank and cell assembly" in which the hydrazine is stored under pressure. Expulsion of the hydrazine under zero-*g* conditions can be accomplished with little risk of entrained bubbles of pressurizing gas in the liquid because the cell, which is a flexible rubber hemisphere, inside the tank acts as a barrier between the gas and the liquid. A "blowdown" mode of operation is used to expel the generant. This means that during the loading operation the tank is only partially filled with generant and the remaining volume, on the other side of the diaphragm, is filled with inert pressurizing gas (nitrogen). The gas expands as the liquid is withdrawn; no gas is added during the withdrawal processes, therefore, the tank pressure decays (from an initial maximum of 1500 psig down to a final minimum of 500 psig).

In *SPS 37-32*, Vol. IV, the above-described system is depicted schematically and detailed descriptions of the generant controller and generant tank and cell assembly are given.

The basic circuit under test in the current series is comprised of two basic parts: (1) a breadboard gas generation subsystem, composed mostly of ALPS components, and (2) a water flow subsystem which simulates the volumetric flow characteristics of the ALPS propellant feed subsystem. In the first-mentioned, an ALPS generant tank and cell assembly feeds hydrazine through a turbine-type flowmeter, the ALPS generant controller, and solenoid shutoff valve to a gas generator. The gas generator discharges into the ALPS heat exchangers which cool the gases before they are delivered into the water flow subsystem tank. This large tank contains a sufficient amount of water to volumetrically simulate the full-thrust propellant flow rate (20 lb/sec) from a propellant tank for slightly longer than 200 sec. From the tank the water flow is split into two parallel circuits, one representing the oxidizer flow and the other the fuel. After passing through turbine-type flowmeters, each flow path continues through an ALPS heat exchanger, a rotary ball shutoff valve, and finally the ALPS 2000-lb thrust dual-pintle throttleable injector, which discharges both streams of water to the atmosphere. In addition to the basic components a variety of fill, pressurization, bleed, venting and relief valves and flow-balancing orifices are provided in the test setup. Numerous pressure transducers, thermocouples, and valve position indicators complete the instrumentation complement.

While the simulation of the ALPS system is fairly accurate as far as inclusion of actual ALPS components is concerned, the line lengths and volumes, unfortunately, could not be duplicated because of the need to incorporate the extra control and instrumentation devices; this lack results in some of the response times and pressure losses, particularly in the water part of the circuit, being somewhat greater than in the ALPS system. Also, to avoid any possibility of water entering the gas generator, the gas line from the heat exchangers was routed externally from the bottom to the top of the tank so that a shutoff valve could be incorporated just at the tank inlet. This long line and the valve have considerably greater volume and pressure loss than the equivalent line on the ALPS system. If sufficient confidence is gained during the tests, some of the extra components may be eliminated later to obtain a more accurate simulation.

For the first test a conventional gas generator of all-welded construction with a commercial poppet-type hollow-cone spray nozzle injector was filled with Shell 405 catalyst and mounted adjacent to the heat exchangers. The generant tank and cell were loaded with approximately 44 lb of hydrazine and then pressurized to

1480 psig with nitrogen gas. A moderately small ullage remained in the water tank after filling and this space was pressurized to 215 psig with nitrogen gas. The water flow from the tank and the hydrazine flow to the gas generator were started simultaneously. Water tank pressure dropped to 187 psig within 1 sec but reached and stabilized at 219 psig 1½ sec later. This recovery resulted from a surge in gas flow from the gas generator which is the expected reaction of the system when the generant controller senses a pressure below its set point. Once stabilized, water tank pressure and all flows continued steadily for 12 sec. At that time, 14½ sec after start, the first of several leaks from the hot gas joints apparently started with the result that tank pressure and the water flow rates began to decline. These leaks were not detected on the monitoring television system for some time, but when the cause of the pressure decay was finally confirmed by visual evidence, the test was terminated. Shutdown was commanded 131 sec after start. Several muffled pops were heard shortly thereafter.

Post-test examination of the records and hardware tended to indicate that a high-amplitude pressure spike had occurred in the gas generator about ¾ sec after start.

This event was not unexpected, as similar spikes were almost always obtained when the Shell 405 catalyst was used in gas generators of nonoptimum design (such as the one used for this test). It was not expected, however, that the pressure spike would cause damage in the downstream hardware; in this case, the fuel heat exchanger was rather severely distorted. Additional damage to the heat exchanger appeared to have been caused by overheating. Fortunately, the hot gas leaks did no harm since protective shields had been installed to deflect the gas streams in case of leakage from the joints. It was later determined by hardware examination that one leak was at a joint where a nonstandard gasket had been substituted for lack of the proper seal. Another gas leak occurred when a weld, which fastened a small flange to the outlet tube of a heat exchanger, cracked where there were sizeable voids in the bead. Several holes caused by the severe distortion were located in the bellows of the fuel heat exchanger. Excessive temperatures were evident in the record of the temperature of the gas where it left the fuel heat exchanger to enter the oxidizer heat exchanger, but this problem was attributed to the damaged condition of the fuel heat exchanger. After shutdown, the poppet in the injector of the gas generator was blown into the catalyst bed when the residual hydrazine trapped upstream of the poppet exploded; the external water sprays which had been used apparently were ineffective in preventing heat soak back from the catalyst bed to the injector after shutdown.

Replacement hardware was installed with some changes prior to the second test. The gas generator was refitted with a new injector of the same design, but a conventional HA-7 catalyst, with a layer of iodine pentoxide-coated catalyst pellets for hypergolic ignition. This catalyst was installed in place of the Shell catalyst in order to avoid the pressure spike. (Elimination of the spike through improved bed and injector geometry is one objective of a current study by a NASA contractor, consequently little could be done toward this goal at JPL without duplicating parts of this existing effort.) Repetition of the injector explosion after test termination was guarded against by installing the gas generator in a special section in the fuel-side water line; in this section, the stream of water in the fuel-simulating leg of the circuit cools the injector and upper end of the gas generator. The outlet tube from the gas generator was welded to the fuel heat exchanger inlet to eliminate one hot gas joint. Reinforcing rings were fastened around the bellows on a spare fuel heat exchanger to prevent recurrence of the distortion; the internal parts were also slightly modified to reduce the likelihood of changes in geometry during operation.

The second test was shut down shortly after start when the generant controller failed to regulate properly. Disassembly revealed corrosion of internal parts to which the failure was attributed. Excessive temperatures were developed in the oxidizer heat exchanger, which resulted in slight distortion; these high temperatures were certainly caused by insufficient cooling capacity in the fuel heat exchanger, which was essentially unaffected by this test.

A newly overhauled and calibrated generant controller and a spare oxidizer heat exchanger were installed for the third test. Once again the generant tank was filled with hydrazine and pressurized. The small ullage over the water was pressurized to 209 psig. Two successful operations were accomplished. The first consisted of 47½ sec of full flow, a fast transient and 20 sec at approximately 20% of full flow, a fast transient and 20 sec at about 90% flow, a fast transient and 20 sec at nearly 75% flow, a fast transient and 26 sec at full flow, and shutdown. The second operation began 17 sec later by starting at a little over 40% of full flow. After 12 sec, the flow was gradually increased over a period of 10 sec to full flow. Full flow was allowed to continue for 44 sec and then the entire system was shut down. No

external water spray or internal gas purge was used on the gas generator, but neither shutdown caused the trapped hydrazine to be heated to detonation temperature.

The transients from one water flow rate to another were accomplished by using the only externally controlled throttling element in the circuit, the dual pintle valve at the ends of the water lines, which was rapidly repositioned by means of a hydraulic servosystem. The rate of hydrazine flow, and therefore the rate of gas generation, was controlled automatically by the generant controller. (At start and shutdown, the solenoid valve between the generant controller and gas generator was used to turn the hydrazine flow on and off.) The generant controller performance is judged by the constancy of the sensing pressure, which is essentially the same as the water tank ullage pressure. When operating at constant water flow rate, the sensing pressure was held within ± 4 psi; most of this variation was a gradual downward drift, probably a result of the decay in hydrazine feed pressure and the friction in the controller. This drift was somewhat surprising, since a controller of the same design had been used repeatedly in similar service (i.e., same range of inlet pressures) and drift had been undetectable (*SPS 37-35*, Vol. II). Overall, throughout the entire test period of 214 sec, the sensing pressure ranged from a maximum of 226 psig down to a minimum of 210 psig (excluding the short start transient when the pressure dropped to 204 psig at ¼ sec after start). Most of the variance was a level shift which seemed to occur in response to the changes in water flow rate, with the higher pressure levels being associated with the lower flow rates. This degree of constancy, ± 8 psi while throttling over a 5-to-1 range of flows, is not considered to be adequate for the ALPS requirements. A newer generant controller design with reduced friction is scheduled to be fabricated and tested later this year.

In this third test, the fuel heat exchanger performance was poor again. At full flow rate the temperature of the gases leaving it are not supposed to exceed 250°F but values as high as 950°F were recorded. Such temperatures are far beyond the reasonable maximums for gases entering the current oxidizer heat exchanger which is constructed of aluminum. On the other hand, the oxidizer heat exchanger proved to be more efficient than was expected; the maximum temperature of the gas leaving it was only 97°F. Both fuel and oxidizer heat exchangers were being cooled by the water from the main tank, which was at 67°F. Maximum temperature of the water leaving the fuel heat exchanger was 100°F, and that leaving the oxidizer heat exchanger was 88°F. These two exchangers (*SPS 37-27*, Vol. IV) were not subjected to calibrations of their heat-transfer coefficients prior to these three pressurization subsystem tests, consequently their performance was unknown. The chief design problems are: the assurance of adequate performance over 20-to-1 range of flow rates, the maintaining of the pressure drops to as low as possible (less than 5 psi is desired), and the prevention of any points of coolant flow stagnation in the fuel heat exchanger.

The comments above comprise but a portion of the conclusions derived from the tests to date. Test number three yielded a great quantity of data on the operating characteristics of the components and the pressurization subsystem. These data are still being analyzed to determine what further changes should be made in the test setup before the next tests are performed.

3. *Injector Development*, D. D. Evans

In recent months the ALPS injector development program has been concentrating upon the optimization of impinging-sheet doublet elements at the 25-lb-thrust level with the nitrogen tetroxide–hydrazine propellant combination. This element was chosen over impinging jets in order to avoid problems of jet misimpingement due to manufacturing tolerances and to provide a degree of film cooling to the injector face. Initial test results were discussed in *SPS 37-29*, Vol. IV, and data relating impingement angle and developed sheet length to combustion efficiency were presented in *SPS 37-34*, Vol. IV. Further analysis of velocity effects and radii of curvature effects has now been completed and is presented in this report.

A sketch of a typical impinging-sheet element is shown in Fig. 1. The important parameters that control each sheet thickness and that determine the degree of mixing of the two propellants are shown on this figure. In excess of 100 short-duration firing tests have been conducted with a variety of deflector geometries. In each case propellants were nitrogen tetroxide-hydrazine, chamber pressure was approximately 130 psia, and mixture ratio was nominally 1.2. These single-element firings were conducted at the 25-lb-thrust level in an uncooled thrust chamber with a contraction area ratio of 7.11 and a characteristic length (L^*) of 19 in. Fig. 2 shows a cross-sectional view of this engine. Note that the injector is designed for two discrete deflector spacings, and that relatively long orifices were used ($L/D > 60$) in order to ensure that reproducible injector hydraulic characteristics would be obtained.

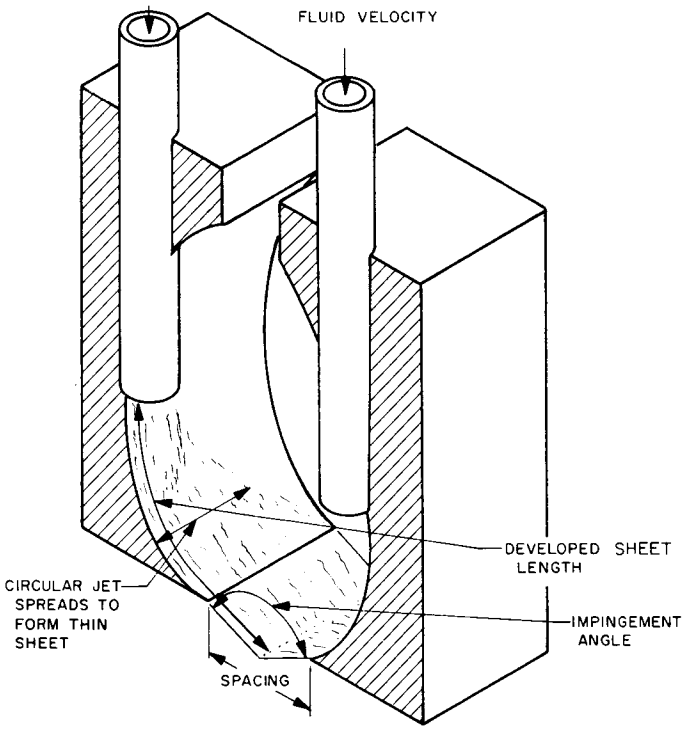


Fig. 1. Key parameters involved in impinging-sheet doublet element optimization

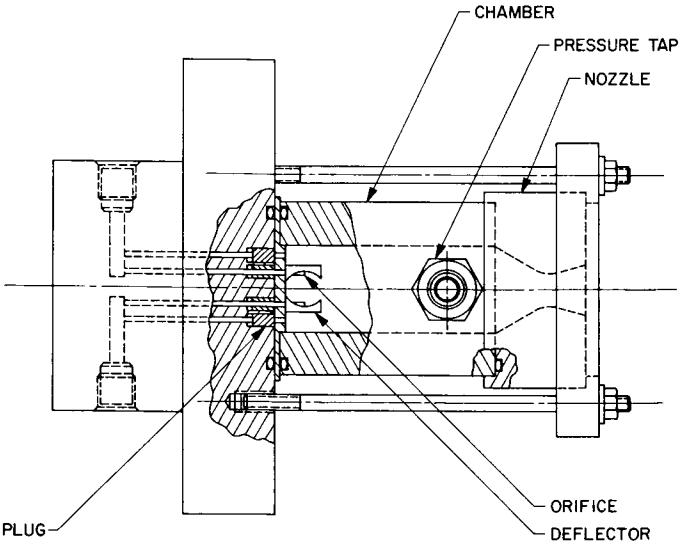


Fig. 2. Rocket engine assembly used to test 25-lb-thrust impinging-sheet injector elements

In order to investigate the effects of incoming fluid velocity, four orifice sizes were selected and mated to deflectors which had been designed to accommodate each particular size. Orifice inside diameters, ID, of 0.033, 0.035, 0.040, and 0.043 in. were chosen to give fuel velocities (V_{fuel}) ranging from 60 to 105 ft/sec and oxidizer velocities (V_{ox}) from 50 to 90 ft/sec. The corresponding pressure drops, ΔP_{inj}, varied from 50 to 150 psi. Figs. 3 and 4 illustrate the variation in experimental combustion efficiency (η_{c^*}) as a function of the fuel and oxidizer velocities. Performance appears to remain relatively constant over the velocity ranges tested, and considerable scatter is evident, especially with far-spaced deflectors. More important, perhaps, is the increase in combustion roughness observed at the lower fluid velocities. The data envelope shown on Fig. 5 indicates that

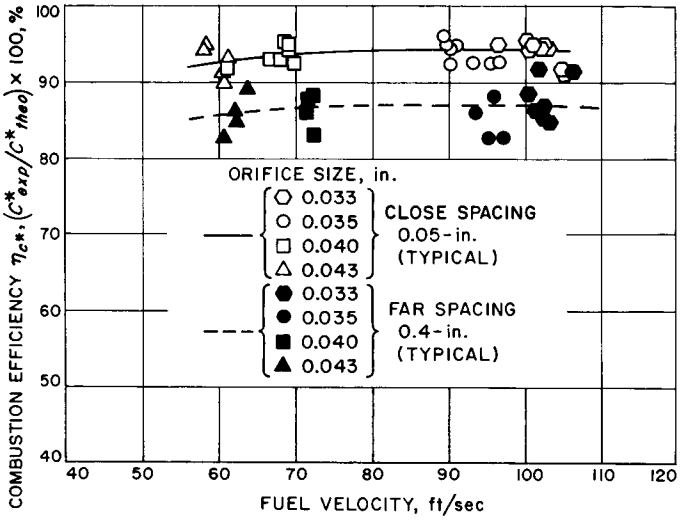


Fig. 3. Effect of fuel velocity on combustion efficiency

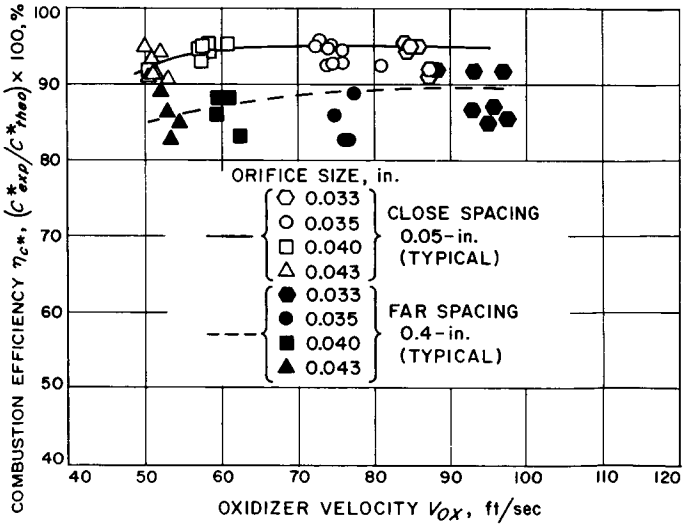


Fig. 4. Effect of oxidizer velocity on combustion efficiency

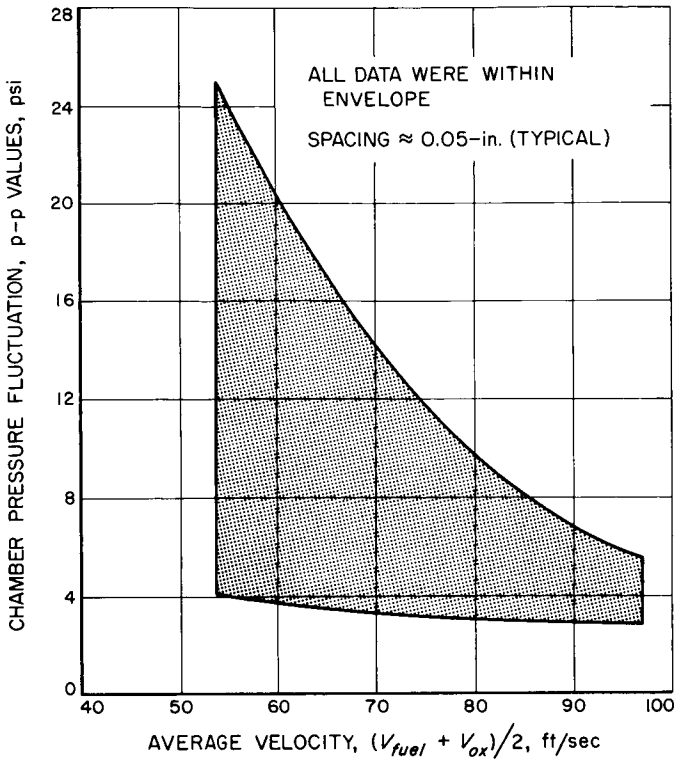


Fig. 5. Effect of average fluid velocity on combustion roughness

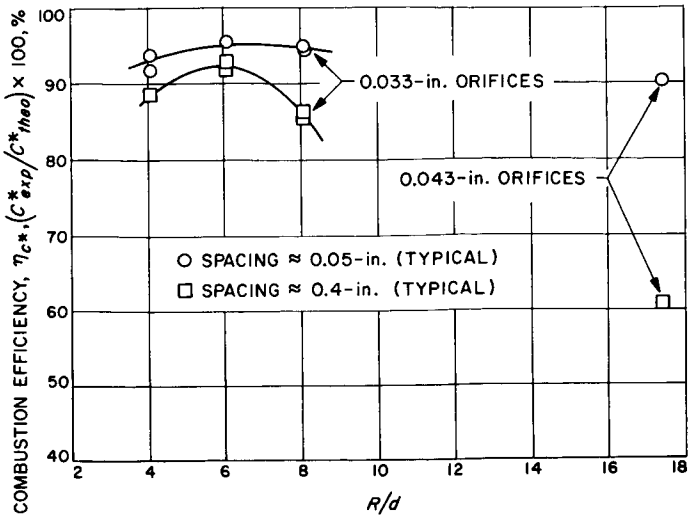


Fig. 6. Variation of combustion efficiency with deflector radius-to-orifice diameter ratio R/d

the combustion roughness, as measured by the peak-to-peak (p-p) excursions in the chamber pressure, is considerably more pronounced at the lower average fluid velocities. Values of 25 psi peak-to-peak (20% of chamber pressure, P_c) were observed at average velocities of 55 ft/sec, and values of 8 psi peak-to-peak (6% of P_c) were observed at velocities of 85 ft/sec. More scatter in the data was evident in the lower velocity regimes also. These combustion roughness values should be treated as relative, however, since no attempt was made to obtain precise pressure transient measurements with high-frequency response instrumentation. The test setup employed a conventional Taber transducer connected to the combustion chamber with a 6-in. length of ¼-in. tubing.

As discussed in *SPS 37-34*, Vol. IV, performance optimized at relatively short developed lengths, on the order of 0.25 in. It was observed that the developed sheet width was a function of the radius of curvature of the deflectors as well as the developed length. Since several orifice sizes were used, it was of interest to treat the data in terms of the ratio of radius of curvature to orifice diameter, *R/d*. Performance data for several deflector geometries are presented in Fig. 6. Note that for close deflector spacings the combustion efficiency η_{c^*} is relatively insensitive to the deflector radius, and that the optimum occurs at an *R/d* of about six when 0.033-in. orifices are used. However, if deflector spacings on the order of 0.4-in. are used, more sensitivity is apparent, and the maximum combustion efficiencies are lower. The sensitivity of the injector to pressure fluctuations is also apparently a function of the *R/d* parameter. An envelope of peak-to-peak values at various *R/d* ratios is shown in Fig. 7. In this case, lower peak-to-peak values are obtained as *R/d* increases, and less data scatter occurs.

The data and conclusions presented above and in previous reports may be summarized as follows:

(1) Impinging-sheet injector elements are simple to fabricate and are capable of delivering in excess of 95% combustion efficiency at the 25-lbf-thrust scale (no heat correction has been applied to the data).

(2) Impingement angles of 90 to 120 deg are acceptable.

(3) Close spacings, on the order of 0.050 in., are desirable.

(4) Developed sheet lengths on the order of 0.25 in. are optimum.

(5) Orifice fluid velocity does not appear to have a strong effect on performance; velocities of 90 to 100 ft/sec produce smoother combustion.

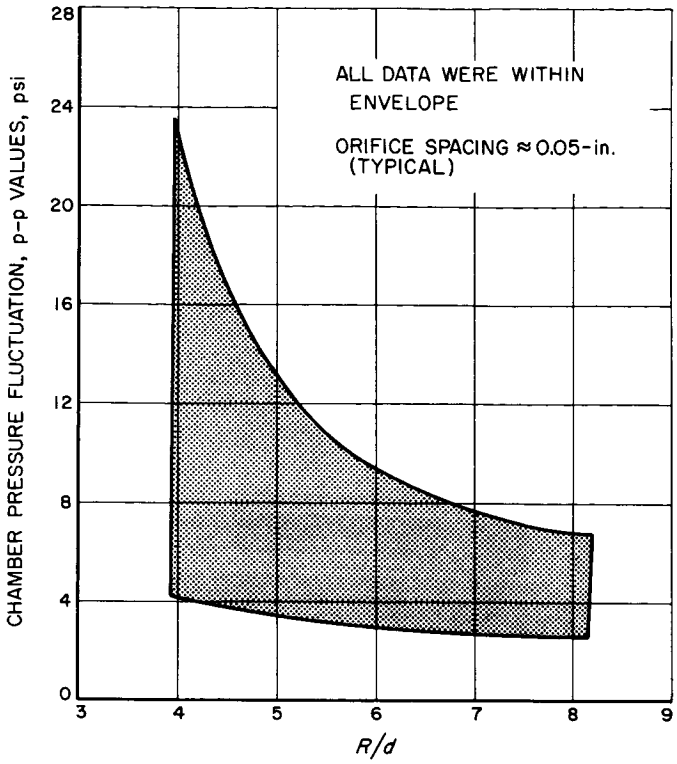


Fig. 7. Effect of radius-to-orifice diameter *R/d* on combustion roughness

(6) The element performance is affected little by variations in the parameters discussed above, thus manufacturing tolerances need not be closely controlled.

All testing to date has involved single-element firings only. A multiple-element 2000-lbf thrust injector is now being fabricated, and these test results will be reported subsequently.

4. *Hydrazine Sterilization Tests,* H. B. Stanford

The sterilization constraints (Ref. 1) for objects entering the atmosphere of Mars yield the following requirements:

1. The probability of landing one living microorganism on Mars shall be less than 10^{-4}.
2. The probability of contaminating Mars by an unsterile orbiting spacecraft due to orbit decay within a period of 50 yr (from the present) will not be greater than 10^{-4}.

In view of these stringent requirements, it has been deemed necessary to accomplish terminal heat sterilization of any spacecraft or capsule that may enter the Martian atmosphere or land on the surface of the planet. The sterilization heat cycle that is currently contemplated consists of a heat soak at a stabilized temperature of 135 ±2°C (275 ±3°F) for 24 hr although other time-temperature cycles are equally effective and are being considered. In order to demonstrate an adequate performance margin a Type-Approval heat cycle has been established which consists of three 36-hr soaks at 145 ±2°C (293 ±3°F).

If heat cycling of N_2O_4 is required, there appear to be no insurmountable problems. The vapor pressure of N_2O_4 is approximately 1000 psi at 290°F. Chemical experience indicates that in clean, compatible hardware this pressure will not be increased by chemical reaction or decomposition at this temperature. Sterilization pressures for given hardware can therefore be predicted and suitable tankage selected on this basis. Aside from the weight penalty of tanks heavy enough to stand these high pressures, subjecting N_2O_4 to sterilization temperature seems feasible.

The use of hydrazine and its derivatives both as a monopropellant and as a bipropellant fuel in sterilizable systems appears likely. The problems associated with sterilizing hydrazine are not so well defined as is the case with the oxidizer. The vapor pressure of N_2H_4 is only 35 psig at 290°F. Under certain conditions, however, hydrazine is known to decompose and build up pressure in an unvented container. As a result of this decomposition the pressure that will be generated in a tank of N_2H_4 heated to sterilization temperature for extended periods of time is not readily predictable. This is true because of the indefinite catalytic effect of heat and tank metal on the hydrazine decomposition rate.

To investigate the effect of sterilization temperatures on N_2H_4 a high-pressure "hot bomb" under development for other compatibility work was pressed into service. This stainless steel bomb is ceramic-lined to reduce by approximately 80% the amount of metal in direct contact with the N_2H_4 in order to reduce the amount of N_2H_4 decomposition by catalysis. In operation some difficulty was encountered in effecting a completely reliable seal between the body sections of the bomb in long-term high-temperature tests. For this reason the results obtained from the first tests are considered qualitative.

Preliminary tests were made by placing 10 cc of N_2H_4 in the bomb (total bomb volume is 20 cc) which was then immersed in a molten metal bath and maintained at

sterilization temperatures for the 36-hr time periods or until the test was terminated because of obvious leaks. Temperature and pressure were recorded as a function of the time of immersion. The test was repeated with 2½-in. samples of 6Al 4V titanium tank material added to the 10 cc of N_2H_4; one sample had oxidized surface and the other was free of oxidation. These samples were cleaned according to JPL's standard procedure for materials in hydrazine systems. The results of these tests are shown in Fig. 8.

From the information presented by these curves only limited conclusions should be drawn because of the inability to positively seal the bomb for extended periods of time. Certain indications are evident, however. The first test with only N_2H_4 in the bomb, and the second test with clean unoxidized titanium added, are reasonably close to each other in rate and degree of pressure rise for a comparable time of 20 hr. Pressure in the third test, using oxidized, cleaned titanium, although 30 to 40 psi higher at the 20-hr mark and an apparent 75 to 100 psi higher at the 30-hr mark by extrapolation, than the other two tests, was not considered to be excessive. These facts tend to indicate that contact with clean unoxidized titanium does not materially affect the decomposition rate of N_2H_4 and that contact with oxidized titanium does not create a catastrophic situation under the conditions of these tests. It was therefore concluded that excessive pressures probably would not be generated by sterilizing N_2H_4 in a titanium vessel. With this assurance it was decided to sterilize N_2H_4 in a flight-type titanium tank, but one capable of withstanding high pressure.

A Block 5 *Ranger* nitrogen tank made of Ti 6Al 4V titanium alloy was obtained. This tank is 6.518 in ID, has a wall thickness of 0.096 in., a design working pressure of 3600 psig, and a minimum burst pressure of 7920 psig. The tank volume is 2440 cc, and attendant plumbing increases this to 2490 cc. The tank was cleaned and passivated with N_2H_4 per flight specifications. A laboratory-type controlled-temperature oven was obtained and installed in a safe area at Edwards Test Station (Fig. 9).

All operations such as tank pressurization, venting, emergency drain, oven control, and data monitoring were controlled remotely. Propellant fill and final drain for sampling purposes were done directly. An addition of 3½ lb (1590 cc) of N_2H_4 to the tank left an ullage space

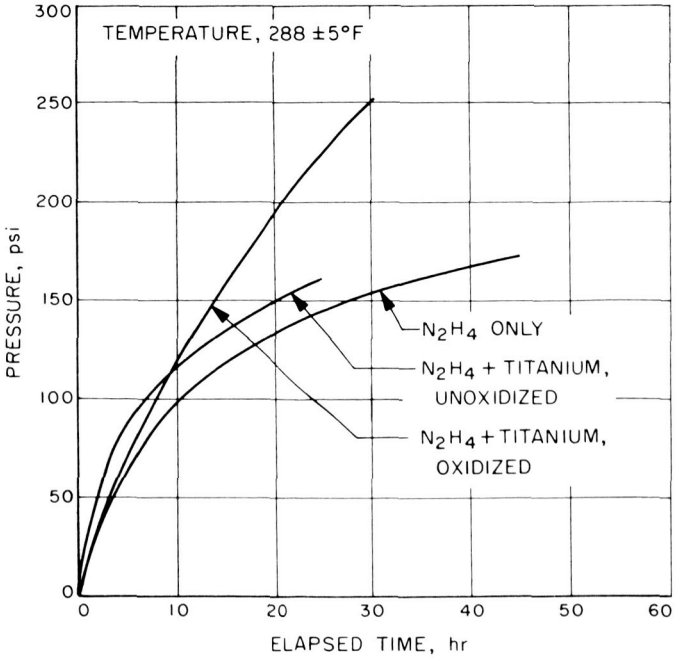


Fig. 8. Pressure histories obtained from N_2H_4 hot bomb tests

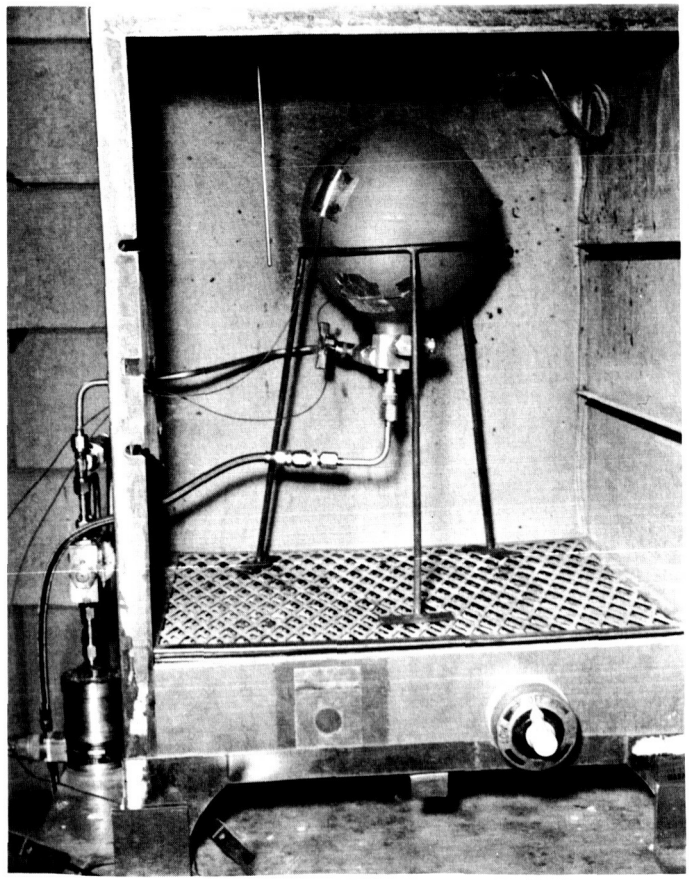


Fig. 9. Test setup simulating sterilization of N_2H_4 in a titanium tank

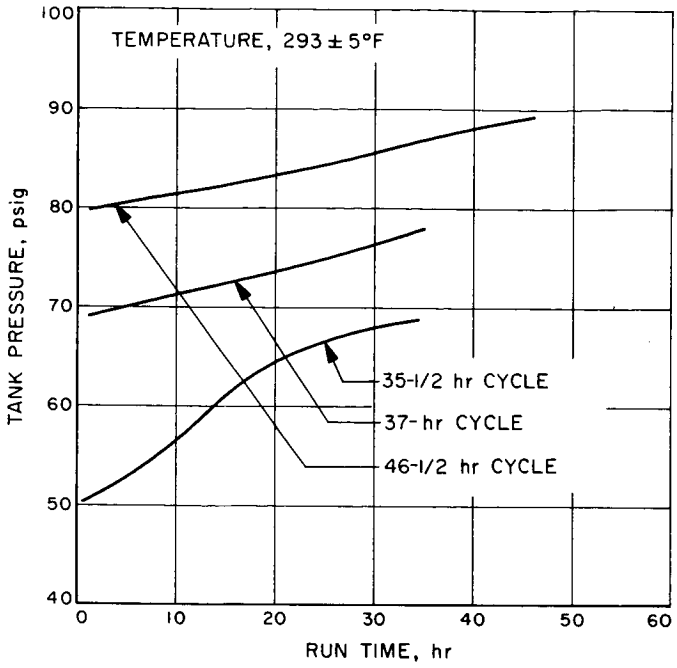


Fig. 10. Pressure histories obtained from N_2H_4 sterilization test in titanium tank

of 36% of tank volume. The ullage space was purged with nitrogen gas and the tank prepressurized to 7 psig. The tank was then subjected to three sterilization cycles consisting of 35½, 37, and 46½ hr duration, respectively, at the established temperature of 293°F. After the third cycle and cooldown to ambient temperature, a residual pressure of 23 psig was noted; this represented an increase of 16 psig over the tank pressure at the start of the test. Curves showing the results of these tests appear in Fig. 10. Chemical analysis of the propellant before test showed it to be 99.3% hydrazine, and post-test analysis indicated 98.9% hydrazine. This change, even if actual, is not considered sufficient to affect performance significantly as a propellant. The exact amount of hydrazine loss was not determined. The maximum pressure generated throughout the test series was 93 psig, which occurred during the third sterilization cycle, a figure much lower than originally anticipated. From these results, assuming a leak-tight system, it can be deduced that in adequately cleaned compatible hardware, hydrazine can be subjected to currently established heat-cycle sterilization procedures without detrimental effect to itself or the tank.

5. *Fill Valve Development,* W. F. MacGlashan

In *SPS 37-22,* Vol. IV, a manually operated *on-off* valve was described. This valve was developed for use in the ALPS system, where it is to be used to seal off securely the ports through which the fuel, oxidizer, monopropellant gas generant, and pressurizing gas are loaded. Several modifications have been made to the valve since the previous report. The new design, shown in Fig. 11, is described below.

The general description and operation of the new and the previous fill valve are identical. Three parts comprise the flight portion: a boss into which the seat is machined, a screw, and a ball which is retained in the end of the screw. The screw holds the ball against the seat to form the single seal in the valve when the valve is closed. The second portion of the valve, a pressure-balanced ground fitting, is attached to the boss and held in place by a multiball detent during filling and draining operations. The valve is opened by backing the ball off the seat. This is done by turning the screw with a special tool. The valve is closed by screwing the valve shut so that the ball is seated. During the time the ball is in contact with the seat, there is no scrubbing or relative motion between the ball and seat, because the ball pivots on the flat bearing surface of the screw. This is important, since it minimizes scratching, galling, and wear.

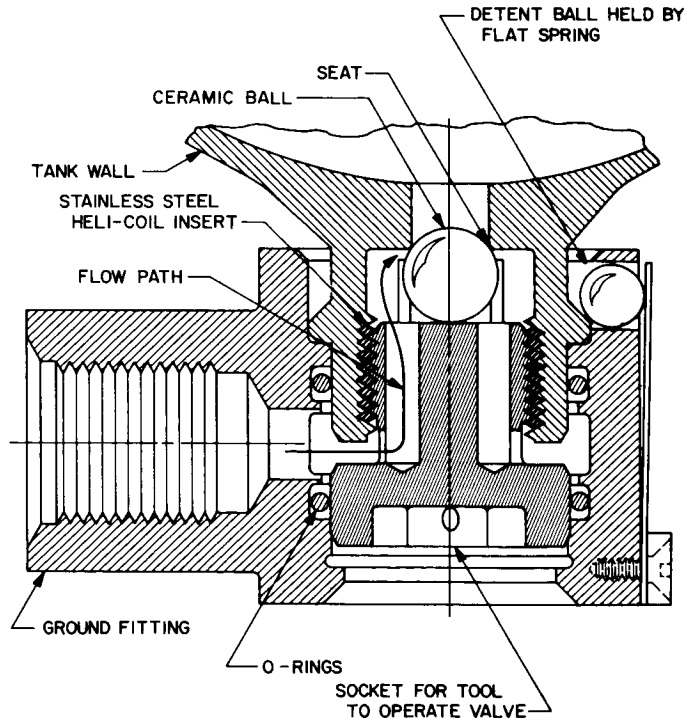


Fig. 11. Cross-section view of the fill valve shown in the closed position

The new fill valve is similar to the old fill valve, but it features several improvements. First, the axial flow passages, which were originally in the boss, were transferred to the screw in the new design. This provided space for increasing the screw thread from 0.375 to 0.438 in. while reducing the diameter of the screw head from 0.688 to 0.625 in. This smaller screw head diameter reduces the axial line pressure load on the threads when the ground coupling is engaged. Second, instead of swaging the seat ball in a socket, the new design retains the ball on the end of the screw with four spring fingers which are integral with the screw. This allows the ball to be readily removable for cleaning, burnishing, or replacement. Third, the seat diameter was changed from 0.188 to 0.156 in. to reduce the ball contact angle with the seat from 106 to 80 deg. This was done because smaller angles diminish the wedging action between ball and seat and reduce the tendency to smear seat metal on the ball. Fourth, a stainless steel Heli-coil thread insert was installed in the boss. It was noticed during the cycling tests with 1700-psi nitrogen gas that the valve with the heli-coil insert generated only about one half the amount of frictional torque that had been measured on valves without the insert. The hard-drawn corrosion-resistant steel Heli-coil not only reduces friction and burnishes the mating threads, but also acts as a magnetic collector by retaining, in the thread lubricant, any metal particles generated in the thread area. Retention of the magnetic particles was apparent during pressure cycling in hydrazine and nitrogen tetroxide. Prior to the use of the Heli-coil insert, examination after test always revealed grease and metal particle migration to the ball and seat. Also, this Heli-coil insert makes possible uniform thread reliability regardless of the boss material.

Fifth, the detent balls, for retaining the ground fitting in engagement, were changed from steel to sapphire or ceramic to eliminate the galvanic corrosion which occasionally froze the steel balls in their sockets. And finally, the ground fitting O-ring in contact with the screw head was changed to Teflon to reduce friction during high-pressure cycling.

Several features continue to be used unchanged. The lapped seat land was maintained at 0.002-in. width to give adequate unit seat loading with low torque. The use of Molykote microsize (MoS_2) burnished ceramic seat ball was continued in the new design. However, the Hercules Oxygen Thread Lubricant has been abandoned because it is possible for the hydrocarbons in it to react with the nitrogen tetroxide to form shock-sensitive compounds. At present, Kel-F No. 90 grease is considered to be the best lubricant for direct contact with nitrogen tetroxide. Apiezon L grease proved to be the best lubricant for lubricating fill valves used for nitrogen gas and hydrazine.

The test results using this modified fill valve design have been satisfactory.

After 100 cycles in hydrazine, the leak rate with 15 in.-lb closing torque was 1 bubble in 4 min; this was an alcohol bubble test using nitrogen gas at 1700 psi. When the closing torque was increased to 20 in.-lb, the leak rate using helium gas at 1700 psi was less than 1 cc (standard) per year according to measurements made on a mass spectrometer. This result typifies the relationship between the leak rate and closing torque.

Another type fill valve, similar to Fig. 11, except without the enlarged screw head, has been built but not tested. This version represents the simplest form for the flight portion. However, it requires a somewhat more complicated ground fitting. The ground fitting (not balanced) is retained by engaging external threads on the flight boss. The seat ball threads are completely relieved of axial load due to line pressure.

A JPL Technical Report describing all fill valves in this series, including the headless screw type just described, has been written and will soon be published.

B. Atomization and Injection Hydraulics

G. Jaivin

In earlier work performed at this Laboratory by Rupe (Refs. 2, 3) the simultaneous mass and mixture ratio distributions of two liquids in the spray resulting from the impingement of a pair of nonreactive liquid streams were measured experimentally. The relative efficiency of the mixing process was evaluated by the definition of a mixing factor that is dependent on the variation in local mixture ratio in the spray. The data obtained indicated that an approximate correlation did exist between certain spray characteristics and injector design parameters. It was concluded that the most uniform distribution of the weight ratio of the two components in the spray produced by a pair of impinging streams is attained when

the product of the velocity head ratio, and the diameter ratio for the two streams is unity. This relationship may be expressed as:

$$\frac{\rho_1 V_1^2 d_1}{\rho_2 V_2^2 d_2} = 1.0$$

where ρ, V, and d represent mass density, stream velocity, and stream diameter, respectively, and the subscripts 1 and 2 refer to the nonreactive fluids simulating a fuel and an oxidizer, respectively.

When simulating a particular propellant combination an attempt is usually made to match the density of the nonreactive liquid to the propellant which it represents. In past work involving studies of spray formed with nonreactive liquids, difficulty has been encountered in finding suitable test fluids, particularly in those instances where large differences in the propellant densities existed. Noting that the densities enter the correlation factor given above as a ratio, the possibility of obtaining similar results by simply matching the propellant density ratio became evident. Of course, identical results might be expected only if all other pertinent hydrodynamic parameters, such as the momentum ratio and the mixture ratio, could be held constant.

Note that the momentum ratio of a pair of liquid streams is given by

$$\frac{\mathbf{P}_2}{\mathbf{P}_1} = \frac{\rho_2 V_2^2 d_2^2}{\rho_1 V_1^2 d_1^2}$$

If the momentum ratio is set equal to 1.0, then one finds that

$$\frac{V_2}{V_1} = \frac{d_1}{d_2} (\rho_1/\rho_2)^{1/2}$$

The mixture ratio r, is given by

$$r = \frac{\rho_2 V_2 d_2^2}{\rho_1 V_1 d_1^2}$$

and by use of the above results

$$r = \frac{d_2}{d_1} (\rho_2/\rho_1)^{1/2}$$

The correlation factor $(\rho_2 V_2^2 d_2)/(\rho_1 V_1^2 d_1)$ becomes equal to d_1/d_2 under the above assumption as well. Thus, it is seen that if the momentum ratio and the diameter ratio are set equal to 1.0, the mixture ratio becomes a simple function of the density ratio and the correlation factor becomes equal to unity. Hence, under the conditions assumed above, if the density ratio is held constant, one would expect the spray characteristics to remain constant.

In order to test this premise two tests were made. An impinging stream doublet having a 60-deg impingement angle was used. Both orifices used had an ID of 0.121 in. and had a length/diameter ratio of 100, thus assuring fully developed turbulent flow conditions at the orifice exit. The test conditions are given in Table 1.

The factor η_m, listed in Table 1 and as defined in Ref. 2 (where it was identified as E_m), is a measure of the uniformity of the mixing occurring as a result of the impingement of the liquid streams. As seen from these data virtually identical results were obtained in the two tests that were made. Fig. 12 illustrates the mass distribution measured in the two test series. It is seen that similar

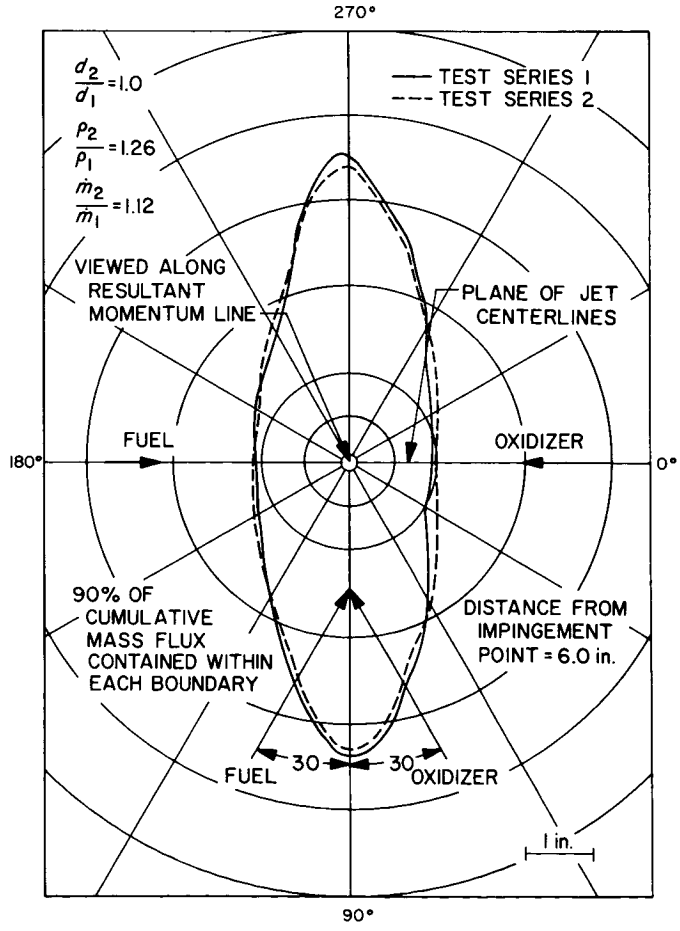


Fig. 12. Mass distributions of the sprays formed by an impinging stream doublet using nonreactive liquids having different absolute density values

Table 1. Parameters of tests to evaluate absolute density effects upon spray properties

Test Series	ρ_1, lbm/ft^3	ρ_2, lbm/ft^3	ρ_2/ρ_1	d_1, in.	d_2, in.	m_1, lb/sec	m_2, lb/sec	r	M_2/M_1	η_m
1	80.0	101.1	1.262	0.121	0.121	0.538	0.606	1.125	1.0	75.3
2	49.6	62.7	1.263	0.121	0.121	0.334	0.375	1.123	1.0	76.5

1. Mixtures of kerosene and perchlorethylene were used as test fluids in these tests.
2. The resultant momentum line of the spray passed through the center tube of the spray collector.
3. The free jet impingement distance was 0.484 in. (4 jet diameters).

distributions were observed, as would be expected, since the momentum ratio and flow geometry were identical in these tests.

Thus it is seen that, for the conditions described above, essentially identical results were obtained when the density ratio was held fixed but the absolute density levels of the nonreactive fluids were varied. If this result is generally applicable to other flow geometries, then a useful means of circumventing the problem of locating suitable test fluids to simulate specific propellant densities is at hand. Future work will be directed toward testing the validity of this result with varied operating parameters.

References

1. Ervin, G. F., "Voyager Capsule Sterilization — Part I. Status, Recommendations, Planning," Technical Memorandum No. 33-213, Jet Propulsion Laboratory, Pasadena, California, April 1, 1965.

2. Rupe, J. H., "The Liquid-Phase Mixing of a Pair of Impinging Streams," Progress Report No. 20-195, Jet Propulsion Laboratory, Pasadena, California, August 1953.

3. Rupe, J. H., "A Correlation Between the Dynamic Properties of a Pair of Impinging Streams and the Uniformity of Mixture Ratio Distribution in the Resulting Spray," Progress Report No. 20-209, Jet Propulsion Laboratory, Pasadena, California, March, 1956.

XIV. Research and Advanced Concepts

A. Thrust Comparisons for Swirling and Nonswirling Flows of Argon Through a Convergent-Divergent Nozzle as Determined From Wall Pressure Measurements

P. F. Massier

1. Introduction

A method of analyzing pressure and temperature measurements obtained at the boundaries of an axisymmetric convergent-divergent nozzle and the cylindrical upstream chamber for swirling flow of argon at approximately ambient stagnation temperature appears in Ref. 1. The nozzle has a convergent half-angle of 28 deg and a divergent half-angle of 5 deg. The ratio of the radius of curvature at the throat to throat radius is 4.0. A section drawing of the test apparatus is shown in Refs. 1 and 2. Distributions of flow variables such as velocities, stagnation and static temperatures, circulation, angular momentum, and swirl angle, which were obtained as a consequence of the measurements and the analysis, are also discussed in these references.

In the present discussion, comparisons are made of wall pressure measurements and thrust at a stagnation pressure of approximately 15 psia for flows with and without swirl. The experimental results are compared with plane one-dimensional flow theory. The experimental values of thrust were obtained by graphically integrating the product of measured wall pressures and areas, rather than by direct thrust measurements; hence, viscous drag and external pressure effects are not taken into account. For the test with swirl, the fluid was injected tangentially at a velocity of 136 ft/sec through four holes equally spaced circumferentially upstream of the nozzle. The tangential velocity near the wall at the nozzle inlet was 84 ft/sec, and the ratio of this tangential velocity near the wall to the axial velocity was 9.2.

2. Wall Pressure Measurements Along the Nozzle

a. Flow without swirl. Static pressure distributions in terms of p/p_t along the wall of the nozzle are shown in Fig. 1. In most of the convergent parts of the nozzle the static pressures for flow without swirl agree closely with the plane one-dimensional flow prediction. In the throat region, however, the wall pressures for this type of flow

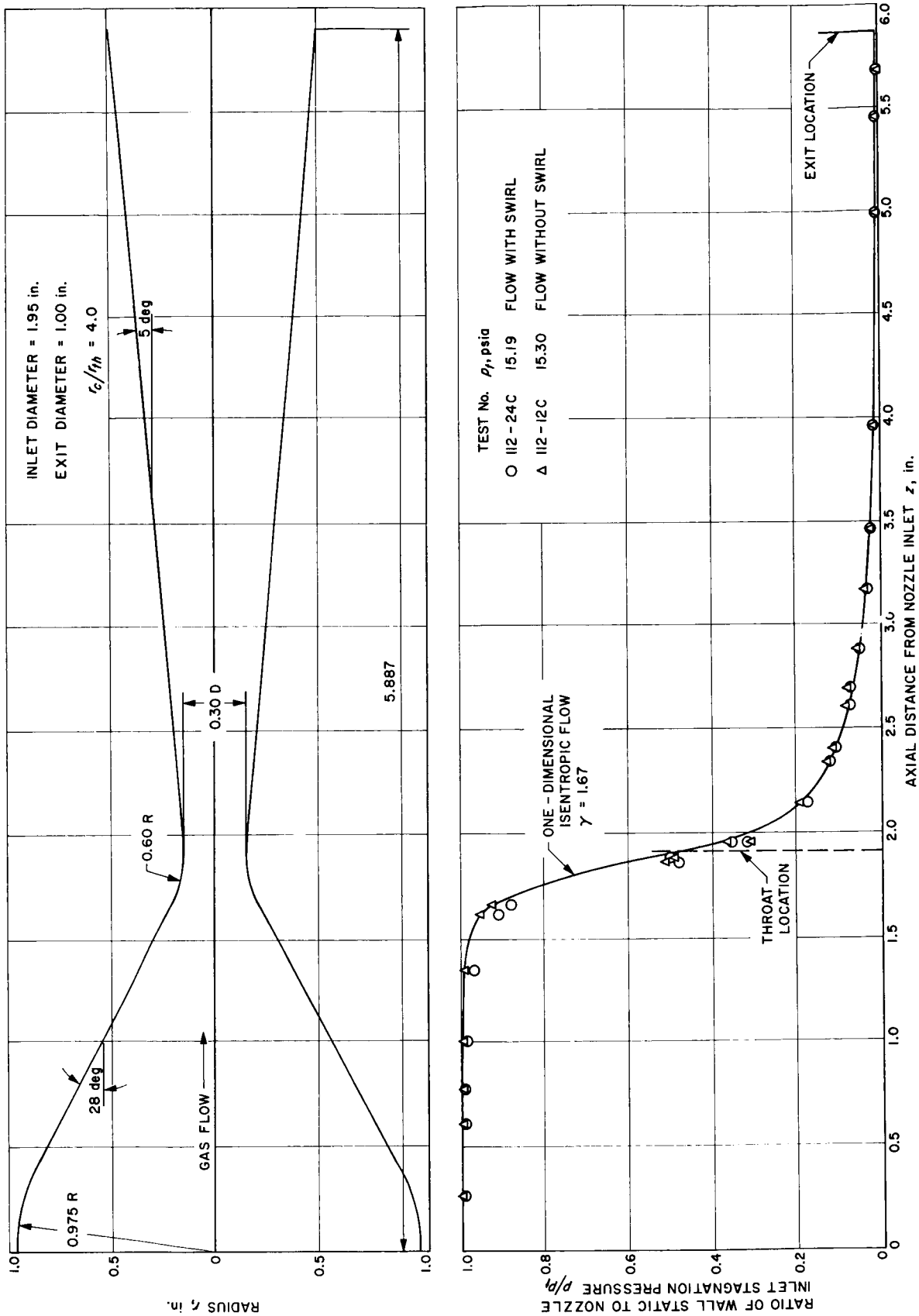


Fig. 1. Wall pressure distributions along a 28–5-deg convergent-divergent nozzle for argon flow with and without swirl

lie below the one-dimensional prediction and then rise above this line just downstream of the throat. Farther downstream, as near the inlet, the wall pressures are very close to the one-dimensional flow prediction. The deviations from plane one-dimensional flow theory are a consequence of the two-dimensional effect (radial velocity component) that results primarily from the throat radius of curvature (Ref. 3). These deviations are also influenced to some extent by the convergent and divergent walls.

The spread of the data in the vicinity of the throat is attributed primarily to the small size of the nozzle in this region. This presented difficulties in fabricating and installing the pressure taps accurately as well as subsequently locating the pressure taps with great precision.

The hole diameter of each pressure tap is approximately 0.019 in., although some of the holes may not be exactly circular in cross section at the nozzle wall. The axial locations of the holes were obtained by casting tool stone into the convergent and divergent parts of the nozzle separately, locating the pressure taps on the hardened stone by inserting drills through the pressure tap holes, removing the stone, and then measuring the distances to the centers of the holes in the stone from the imprints left by the flange faces. Shrinkage of the stone was very small.

To obtain the thrust results discussed in the next section, curves were faired through the data. The error associated with the location of the faired curve would, of course, influence the magnitude of the results. However, in the throat region where the fairing was done, the surface area was comparatively small; consequently, the error in the results introduced by the location of the curve is small.

b. Flow with swirl. In the convergent section of the nozzle and in a portion of the divergent section, wall static pressures for flow with swirl are generally lower than those for flow without swirl because of the comparatively large tangential velocity component associated with the swirling flow. In this case, the tangential velocity was initiated by injecting the argon tangentially into the cylindrical chamber upstream of the nozzle. The ratio of tangential to axial velocity for the swirling flow test was largest at the nozzle inlet, where it was 9.2 near the wall, and then diminished along the axial flow direction. Fig. 9 in Ref. 1 shows the magnitude of this ratio along the nozzle.

The static pressure distribution along the wall of a nozzle in which the flow has a swirl component is therefore influenced both by the two-dimensional effect (radial velocity component) that results from the wall contour and by the tangential velocity component of the swirling flow.

3. Thrust

The ratio of the axial thrust for flow with swirl to that for flow without swirl is shown in Fig. 2. This thrust ratio was determined by graphically integrating (on a greatly expanded scale) the product of the measured wall pressures and the radial component of the gas-side surface area; hence, the viscous drag and the external pressure were not taken into account. To accomplish the integrations, the wall static pressures were normalized to a constant experimental value of the inlet stagnation pressure for each of the tests.

For flow with swirl the stagnation pressure is a function of radius; however, as deduced from end-wall static pressure measurements, the stagnation pressure did not change much from the outer radius to the throat radius (Ref. 1, Fig. 4). Consequently, the stagnation pressure near the cylindrical wall at the nozzle inlet was selected as the normalizing pressure. For flow without swirl the stagnation pressure was assumed to be uniform, since no variations were observed in the end-wall static pressure measurements.

The ratio of thrust for flow with swirl to that for flow without swirl was obtained by the following equation:

$$\frac{F_{\text{with swirl}}}{F_{\text{without swirl}}} = \frac{\left\{p_t\left[\int \frac{p}{p_t}\,d\left(\frac{A_c}{A_{th}}\right)\right]_{ew} + \left[\int \frac{p}{p_t}\,d\left(\frac{A_c}{A_{th}}\right)\right]_{noz}\right\}_{\text{with swirl}}}{\left\{p_t\left[\frac{p}{p_t}\left(\frac{A_c}{A_{th}}\right)\right]_{ew} + \left[\int \frac{p}{p_t}\,d\left(\frac{A_c}{A_{th}}\right)\right]_{noz}\right\}_{\text{without swirl}}} \quad (1)$$

The decrease in thrust resulting from introduction of the previously specified amount of swirl into the flow is quite evident in Fig. 2. The axial thrust ratio for the entire nozzle assembly is the value that exists at the exit which is slightly below 0.80. The distribution along the nozzle indicates the extent to which the swirl influences the

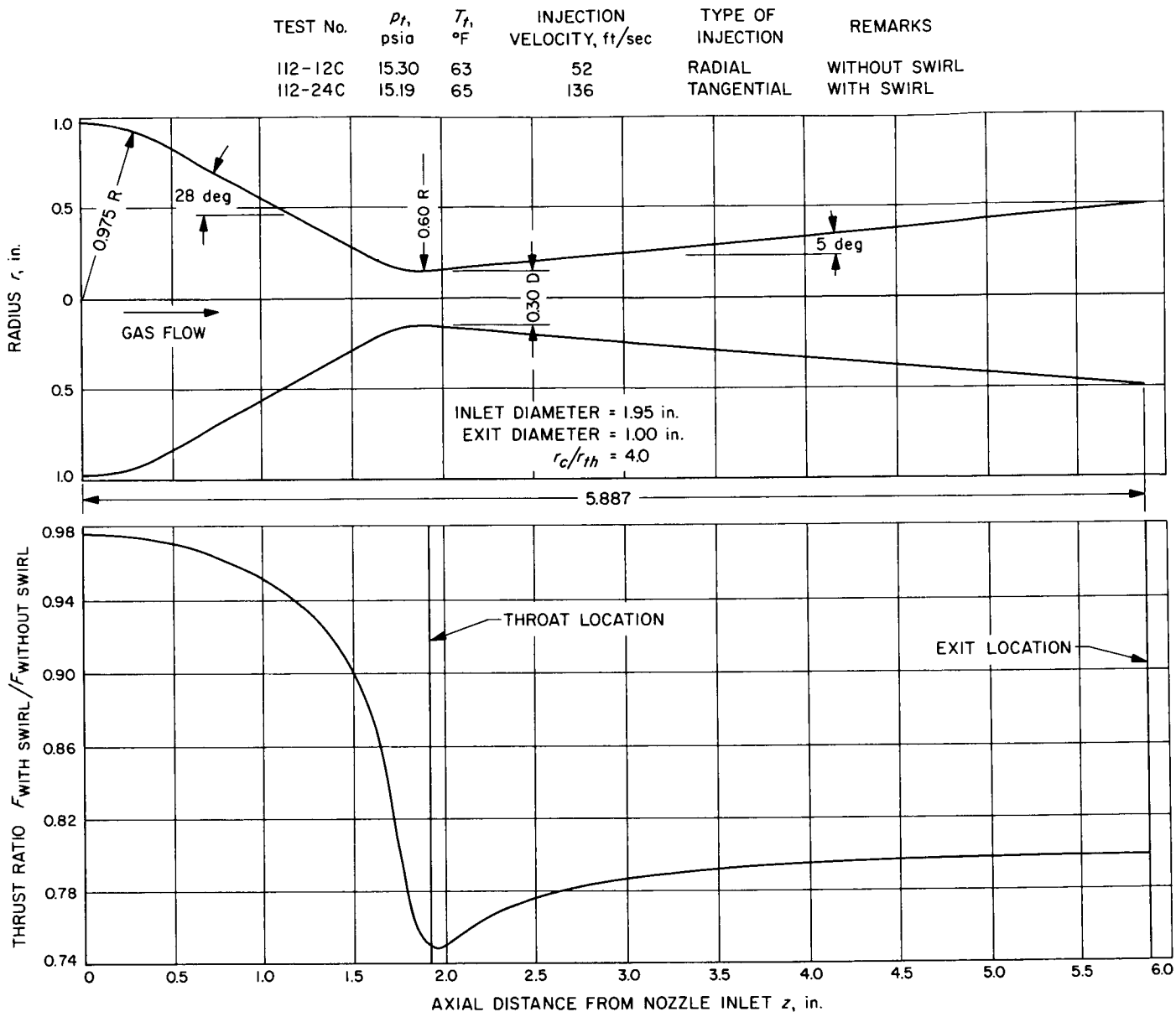


Fig. 2. Distribution of the ratio of thrust with swirl to thrust without swirl for argon flow along a 28–5-deg convergent-divergent nozzle

thrust from the inlet up to each location. A minimum value is attained in the vicinity of the throat. Presumably, convergent-divergent nozzles of this same configuration, but with shorter divergent sections, would have the same distribution up to their respective exits. Convergent nozzles, so short that they did not include the same throat, would undoubtedly have a different distribution.

The thrust ratio distribution would, of course, be different from that shown in Fig. 2 if the swirling flow field at the nozzle inlet were different from that of this test. This flow field would be altered, for example, by a change in injector configuration which could change the injection velocity and, hence, the tangential velocity distribution. The flow field would also be altered by a change in the mass flow rate in the existing apparatus. Numerous variables are needed to characterize the flow. Some of the important ones are v_w/u at the nozzle inlet, p_t, and T_t. Other influencing factors are the nozzle configuration and the fluid.

The thrust ratio at the nozzle inlet represents the ratio of forces on the end wall. The lower force at this location for flow with swirl is a consequence of the lower static pressures, particularly in the vicinity of the centerline, as shown in Fig. 4 of Ref. 1.

In the convergent part of the nozzle, the wall pressures for flow with swirl are also lower than those for flow without swirl, as may be seen in Fig. 1. The axial component of the force on the convergent part of the nozzle acts in the opposite direction of the thrust, and a reduction in pressure in this part would result in an increase in thrust for any given end-wall pressure. However, the lower pressures in this region for flow with swirl are not sufficiently low to cause an increase in the thrust, because the force on the end wall is not the same but is also lower. The net effect is that the thrust ratio decreases from the inlet to the throat.

In the divergent part of the nozzle, the axial component of the force acts in the same direction as the thrust and therefore contributes to the thrust. In Fig. 2 the thrust ratio rises gradually downstream of the throat because the static pressures for the two tests are very close together; therefore, because of the lower thrust at the end wall, the fraction of the contribution was greater for flow with swirl than for flow without swirl.

4. Summary and Conclusions

Experimental static-pressure distributions in terms of p/p_t along the wall of a convergent-divergent nozzle indicate lower pressures for swirling gas flow than for flow without swirl. Furthermore, compared with a plane one-dimensional flow prediction, this pressure distribution is influenced not only by the comparatively large tangential velocity component of the swirling flow but also, to some extent, by the two-dimensional effect (radial velocity component) that results from the wall contour.

The axial thrust produced by the flow of a gas through a convergent-divergent nozzle can be influenced significantly by the presence of swirl. For the experiment analyzed at a stagnation pressure of approximately 15 psia, a tangential injection velocity of 136 ft/sec, and a tangential velocity near the wall at the nozzle inlet of 84 ft/sec, the thrust was approximately 20% lower than when the gas was injected radially, rather than tangentially, upstream of the nozzle. The ratio of the tangential to axial velocity near the outer wall at the nozzle inlet was 9.2.

Terms used in this section are defined in the following table.

Definition of Terms

A_c	cross-sectional area
A_{th}	area of nozzle throat
F	thrust
p	static pressure
p_t	stagnation pressure
r	radius
r_c	radius of curvature at the nozzle throat
r_{th}	nozzle throat radius
T_t	stagnation temperature at nozzle inlet
u	axial component of velocity
v_w	tangential component of velocity near the wall
z	axial distance from nozzle inlet
γ	specific heat ratio
Subscripts	
$_{ew}$	end wall
$_{noz}$	nozzle

B. Liquid MHD Power Conversion

D. Elliot, D. Cerini, L. Hays, and E. Weinberg

The long lifetimes required of electric-propulsion powerplants make cycles without rotating components attractive. A nonrotating powerplant under investigation at JPL is the liquid metal magnetohydrodynamic (MHD) system shown schematically in Fig. 3. In this cycle a fluid, such as cesium, circulates in the vapor loop and causes a liquid metal, such as lithium, to circulate through an MHD generator in the liquid loop. The cesium leaves the radiator as a condensate, flows through an electromagnetic pump and regenerative heat exchanger to the nozzle, vaporizes on contact with the lithium, atomizes and accelerates the lithium in the nozzle, separates from the lithium in the separator, and returns to the radiator through the regenerative heat exchanger. The lithium

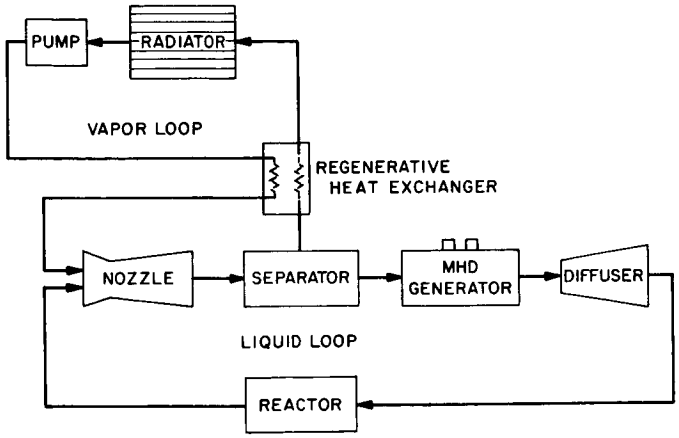


Fig. 3. Liquid MHD power conversion cycle

leaves the separator at high velocity (typically 500 ft/sec), decelerates through the production of electric power in the MHD generator, and leaves the generator with sufficient velocity (typically 300 ft/sec) to return through a diffuser to the reactor where the lithium is reheated.

Recent tests with cold NaK (Ref. 4) showed that the required closed-loop circulation of liquid metal can be accomplished, and analysis of AC liquid-metal generators (Ref. 5) indicated that efficient conversion of the fluid power to AC electric power is possible. Cycle analysis (Ref. 4), combining the liquid circulation results with the theoretically attainable generator performance, predicted 7% efficiency, 2.2 ft²/kw(e) prime radiator area, and 25 lb/kw(e) specific weight (unshielded) for a 300-kw(e) cesium-lithium system at 2000°F reactor outlet temperature. Confirmation of the generator analysis method is the object of forthcoming tests of a 5-kw AC generator using NaK (Ref. 5).

Fig. 4 shows the program currently being followed, with approximate target dates. The four main tasks are: (1) generator experiments, (2) NaK-nitrogen conversion system, (3) materials evaluation, and (4) cesium-lithium conversion system.

1. Generator Experiments

Fabrication of the experimental 5-kw AC generator is nearly complete. The NaK flow components have been assembled with dummy stator blocks for leak and flow testing. The stator coils were assembled on iron stator

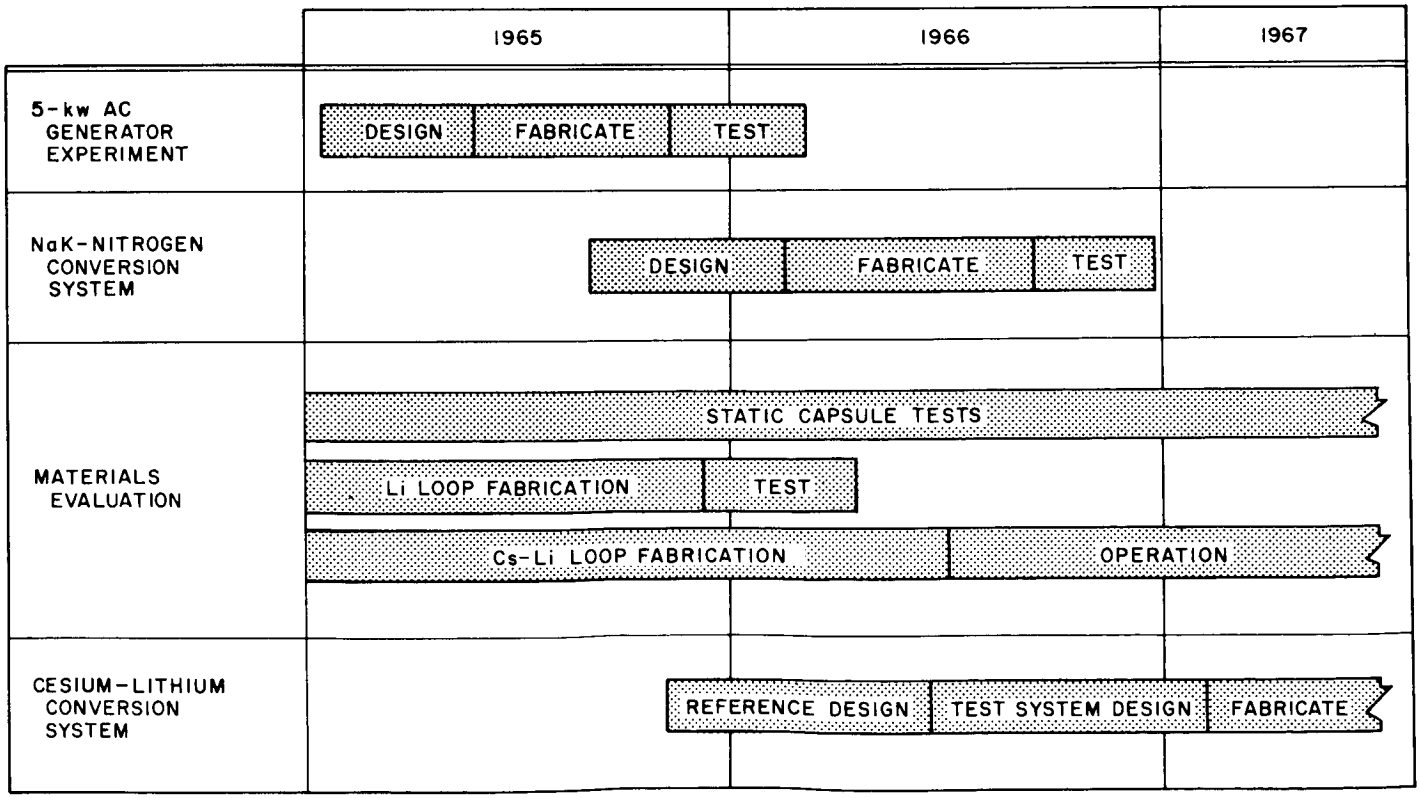


Fig. 4. Program schedule

blocks, and proper operation of the single-wavelength traveling wave and the compensating poles was verified with DC excitation. Laminated stator blocks were bonded from 8-mil Hiperco-50 and the slots were formed by electrical-discharge machining. Instrumentation, excitation, switching, and loading equipment are being assembled. Following bench tests to determine the empty-channel electrical characteristics of the generator, NaK tests will begin.

A 400-cycle AC MHD generator built by Atomics International (AI) as part of its liquid MHD program was tested in the JPL NaK facility, with good results. The data will be presented in an AI report.

2. NaK-Nitrogen Conversion System

This system will employ the nozzle used in the NaK-nitrogen closed-loop tests (Ref. 4), together with a two-dimensional separator and flat-channel AC generator, to produce a nominal 50-kw of three-phase AC power at a frequency of about 600 cps. The generator stators have been designed and iron mockups are being fabricated for magnetic field measurements with DC excitation. Design of the separator, generator housing, and diffuser is in progress.

3. Materials Evaluation

In a 2000°F cesium-lithium system, materials harder than columbium may be needed for the parts of the separator where the lithium initially impinges. Two materials of interest for this purpose, TiC and ZrC, were screened for lithium compatibility in static capsule tests. After exposure to 2000°F lithium for 1700 hr, the TiC sample was unchanged, except for a slight increase in surface porosity visible at 800 × magnification. The ZrC sample had developed a number of cracks (judged due to thermal shock rather than chemical attack) and showed a slightly greater increase in porosity than the TiC sample.

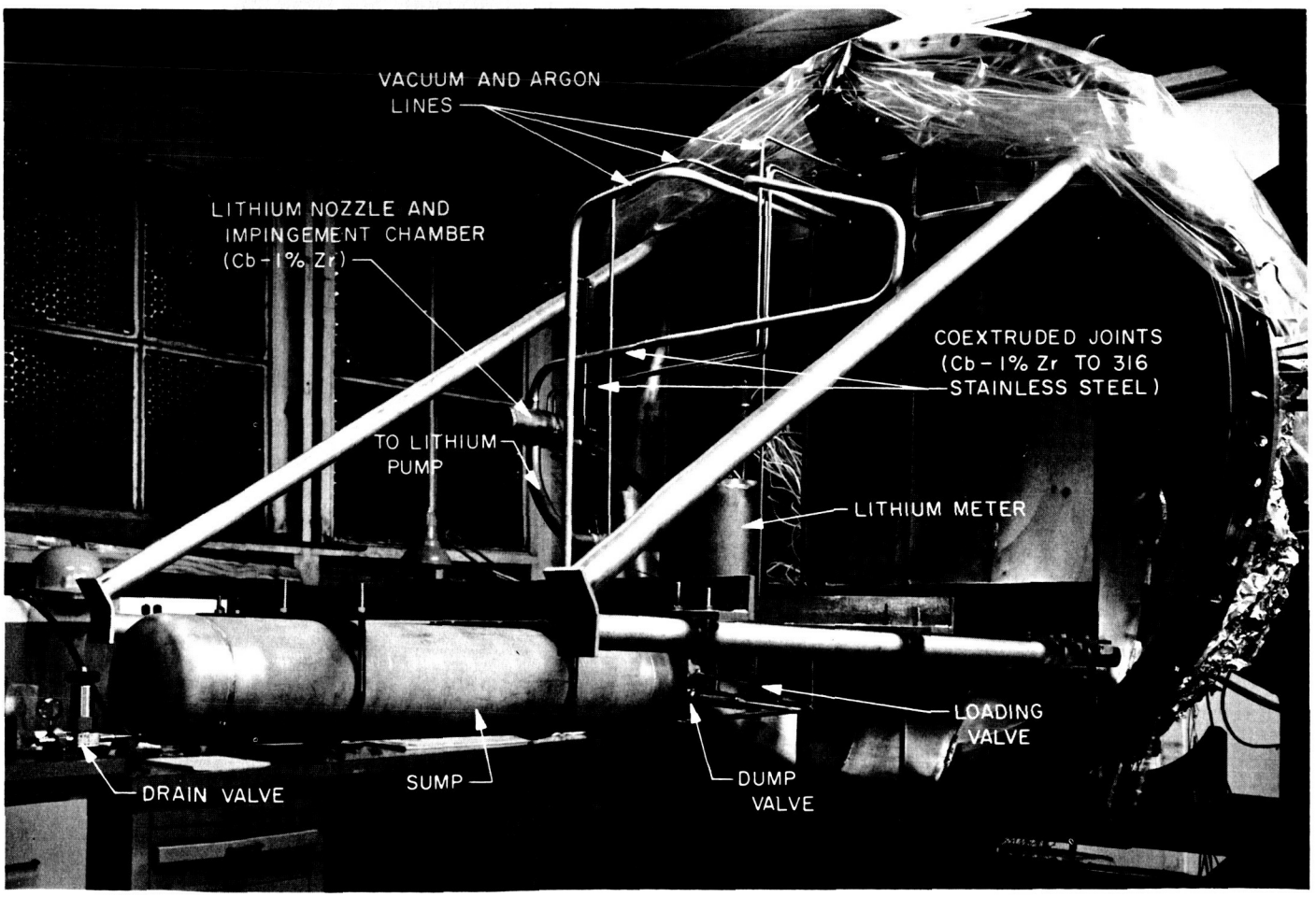


Fig. 5. Lithium erosion loop

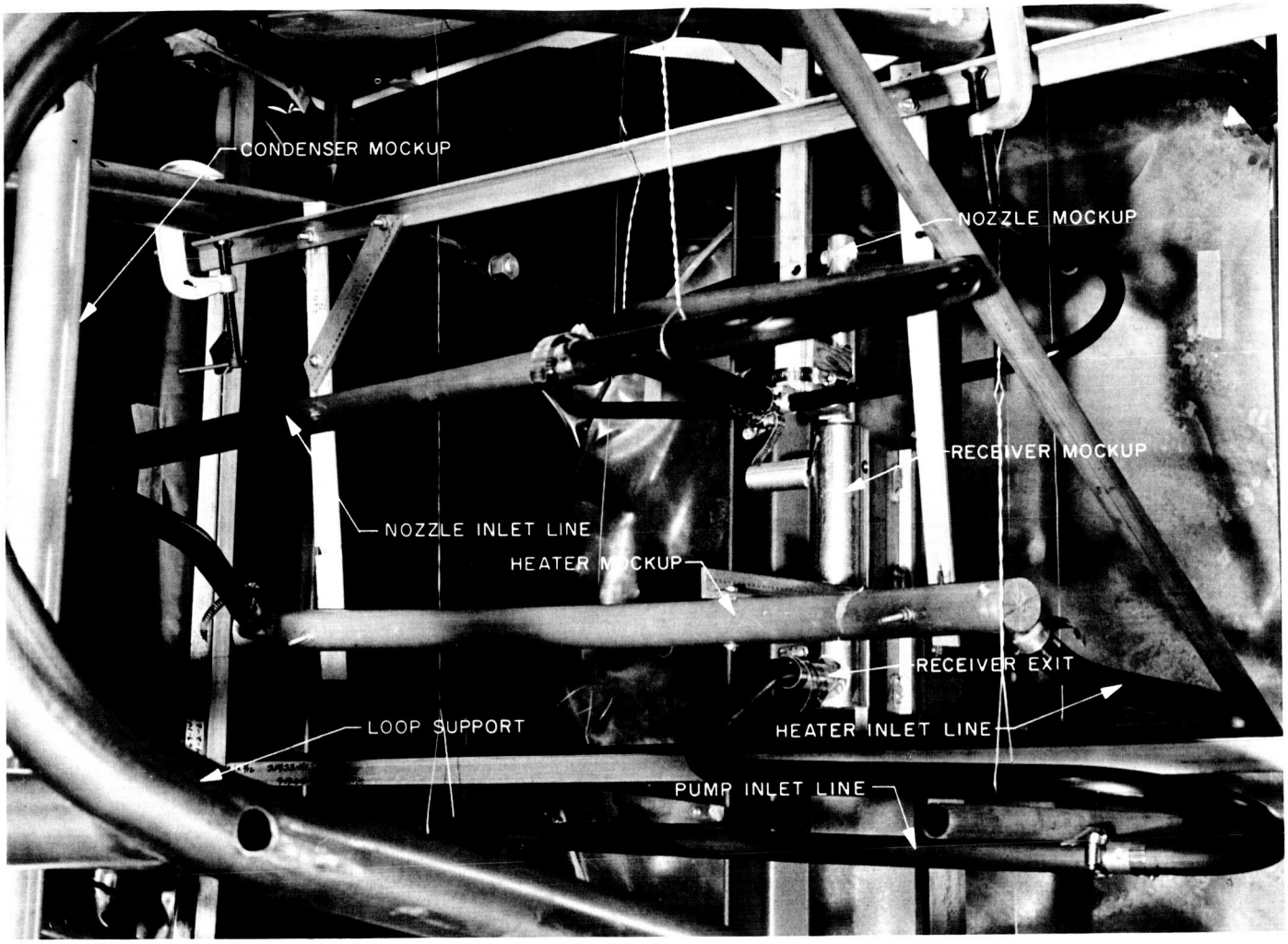


Fig. 6. Cesium-lithium erosion loop piping assembly

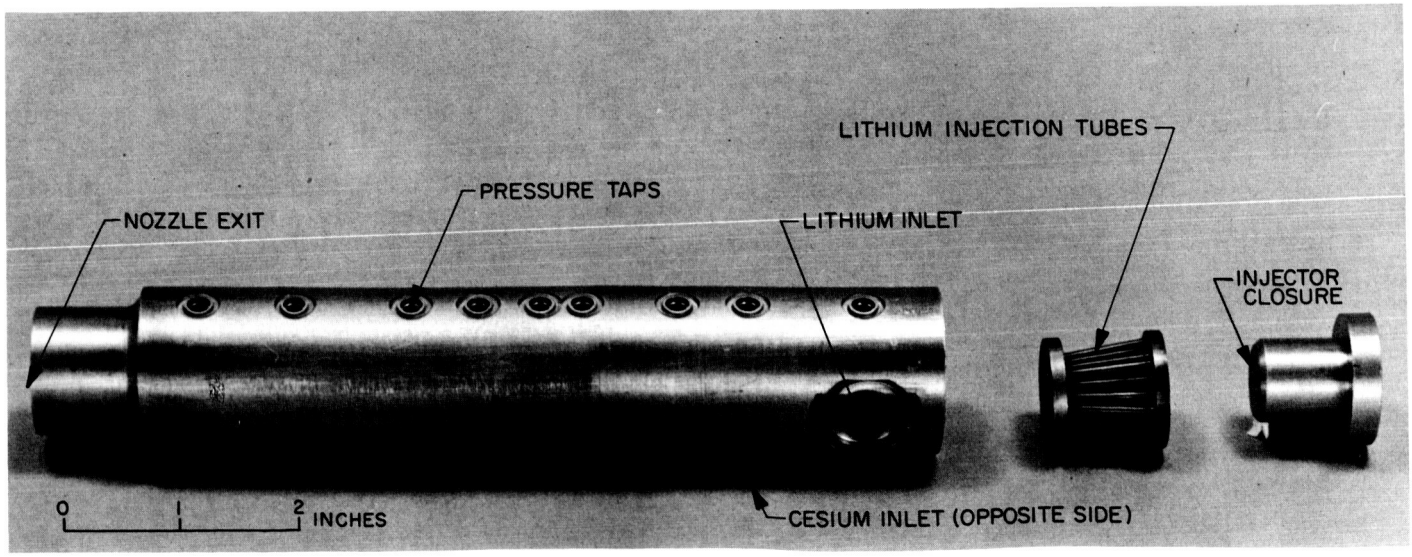


Fig. 7. Cesium-lithium nozzle for erosion loop

Capsule tests of hard materials, as well as of electrical insulators for the generator, are continuing.

A 2000°F cesium-lithium loop is being fabricated for evaluating the resistance of materials to impinging and flowing lithium at velocities up to 700 ft/sec. The initial version of the loop, circulating lithium only, will be used for performance testing of the helical induction lithium pump at 2000°F and to obtain preliminary data on lithium erosion of candidate separator and containment materials at velocities up to 250 ft/sec. In Fig. 5 this system is shown installed on the door of the ion-pumped vacuum system, prior to installation of thermocouples. All weldments (Cb-1% Zr and Type 316 stainless steel) have been leak-checked with a mass spectrometer and subjected to radiographic inspection. Critical Cb-1% Zr weldments were annealed at 2200°F for 1 hr. Final loop operations (lithium fill, checkout, operation) will be performed in the Energy Conversion Laboratory to be occupied in December.

Fabrication of the major Cb-1% Zr lines for the final form of the cesium-lithium erosion loop has been completed. Mockups of the components to be connected were fabricated of aluminum and mounted on the loop frame to provide anchor locations for the piping. The piping was then formed to the coordinates found satisfactory in the computer thermal stress calculations. This assembly is shown in Fig. 6. Supports for the lines and components have been designed and will be installed after completion of field welding.

The Cb-1% Zr mixer-nozzle assembly was completed by electron beam welding. Fig. 7 is a photograph of this unit prior to assembly. A group of 37 tubes in a hexagonal array is provided for lithium injection. Cesium enters the mixing section through the cusp areas between the tubes and vaporizes on mixing with the lithium. The lithium heater and cesium condenser are also approaching completion. The heater buses and lithium receiver are the remaining major loop components to be fabricated.

4. Cesium-Lithium Conversion System

A nominal 300 kw(e), 2000°F cesium-lithium conversion system is being designed, and methods of testing this system are being studied. The intended sequence is to prepare a reference flight-system design and then introduce minimum modifications in materials, layout, and operating conditions to provide a conversion system for laboratory testing with electrically heated lithium. Reference design studies are beginning with investigation of the generator insulation and cooling problem.

Lithium heater experiments are continuing with a heater using beryllia insulation. This assembly is shown in Fig. 8 prior to testing. The heat rejection in this test is increased by cylindrical fins of Cb-1% Zr which increase the radiant heat transfer area. The primary objective will be to determine the electrical resistivity and breakdown characteristics of the swaged beryllia at elevated temperatures for a geometry similar to that previously tested with alumina insulation. This heater will be tested in a vacuum system provided with diffusion pumping, in addition to ion pumping, to accommodate the large hydrogen release during outgassing of the Cb-1% Zr alloy.

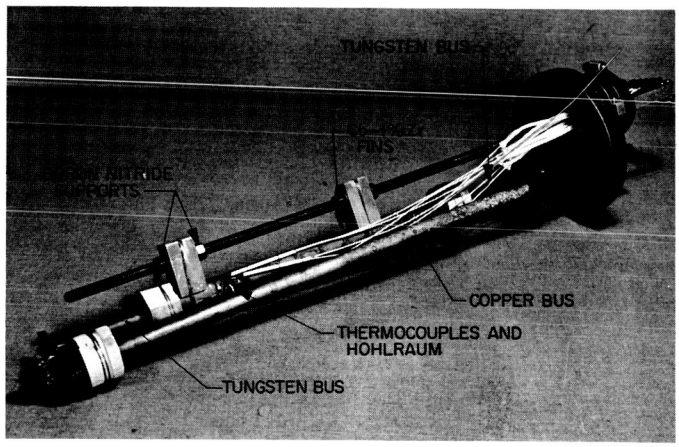


Fig. 8. Lithium heater test assembly

C. A Zero-Gravity Feed System for Mercury Ion Engines

T. D. Masek and D. J. Kerrisk

1. Introduction

This is a continuing report on the status of the zero-gravity feed system being built for test with an oxide-cathode mercury electron bombardment ion engine. The system, described in Ref. 6, consists of six major components, as shown schematically in Fig. 9. It is an expanding diaphragm system, using gas pressure behind a diaphragm to force liquid mercury to the vaporizer. A

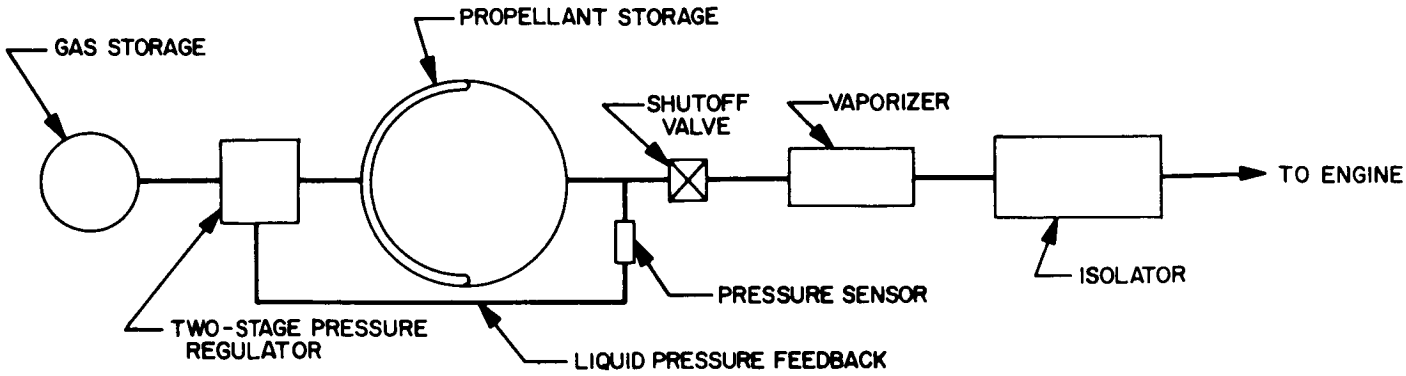


Fig. 9. Schematic of zero-gravity mercury propellant feed system

pressure feedback from the liquid expulsion line to the second stage of the two-stage pressure regulator is used to maintain a constant pressure in the liquid. Fig. 10 is a conceptual design of a propulsion system module using this feed system.

2. Component Development

Since the most recent report (Ref. 6), several design changes have been made in various components, and testing is being pursued in several areas.

a. Isolator. The isolator described in Ref. 6 was 10 in. long and had an ID of 1¼ in. Isolators of glass and boron nitride were being tested. These isolators required an auxiliary heating coil to prevent mercury condensation. A Paschen curve analysis had indicated that much smaller isolators could be used; subsequently, a 5 in. long, 1¼-in. ID isolator was fabricated of boron nitride. It has been found that this isolator performs quite satisfactorily; in addition, the shorter length has permitted the elimination of the auxiliary heating coil. The high thermal conductivity of boron nitride, coupled with this shorter length, permits sufficient heat to be conducted from the vaporizer and from the engine housing to maintain the temperature of the isolator at about 150°C, well above the condensation temperature for mercury at the exit pressure (18 to 48°C for the pressure range of 10^{-3} to 10^{-2} torr). Still smaller isolators seem feasible, and a further reduction in size and weight is planned for the final system design.

b. Vaporizer. The vaporizer described in Ref. 6 was a 50% dense porous stainless rod, 6 in. long, wrapped with a single heater coil. Tests had been conducted in which the liquid-vapor interface was inside the vaporizer and at the vaporizer inlet. It had been found that the former mode of operation was unstable. In recent tests, the vapor interface has been consistently maintained at the vaporizer inlet, at which point operation appears to be completely stable. By taking advantage of the large mercury surface tension forces (approximately 15 psi for the vaporizers used), fairly large variations in the liquid pressure head can be tolerated without influencing system operation in any way.

The 6-in. vaporizer has been replaced by a 3-in. vaporizer, also 50% dense porous stainless, but wound with two heating coils to permit separate control of the inlet temperature and the temperature gradient. The former, by establishing the temperature at which vaporization takes place, controls the vapor pressure which provides the driving force for the propellant; the latter controls the vapor flow impedence of the porous rod. Together they provide gross and fine controls on the propellant flow rate.

c. Shutoff valve. Since the last report, a shutoff valve has been incorporated in the liquid expulsion line. This valve, weighing 0.75 lb, is solenoid-actuated and requires 3.6 w in operation. This valve was designed by Hughes Research Laboratories (HRL) for their SERT I ion engine test flight, and has been space-qualified. In 4 mo of daily operation it has proven completely reliable, and is presently considered satisfactory for the final system.

d. Propellant storage tanks. All system tests to date have been performed with the bolted metal diaphragm tank described in Ref. 6. In addition, elastomeric diaphragm tanks are being investigated. The five elastomers tested for expulsion pressure (as reported in the reference cited above), have been soaked in mercury for 3000 hr. Visual inspection of the samples after this time has indicated some deterioration of all samples with the exception

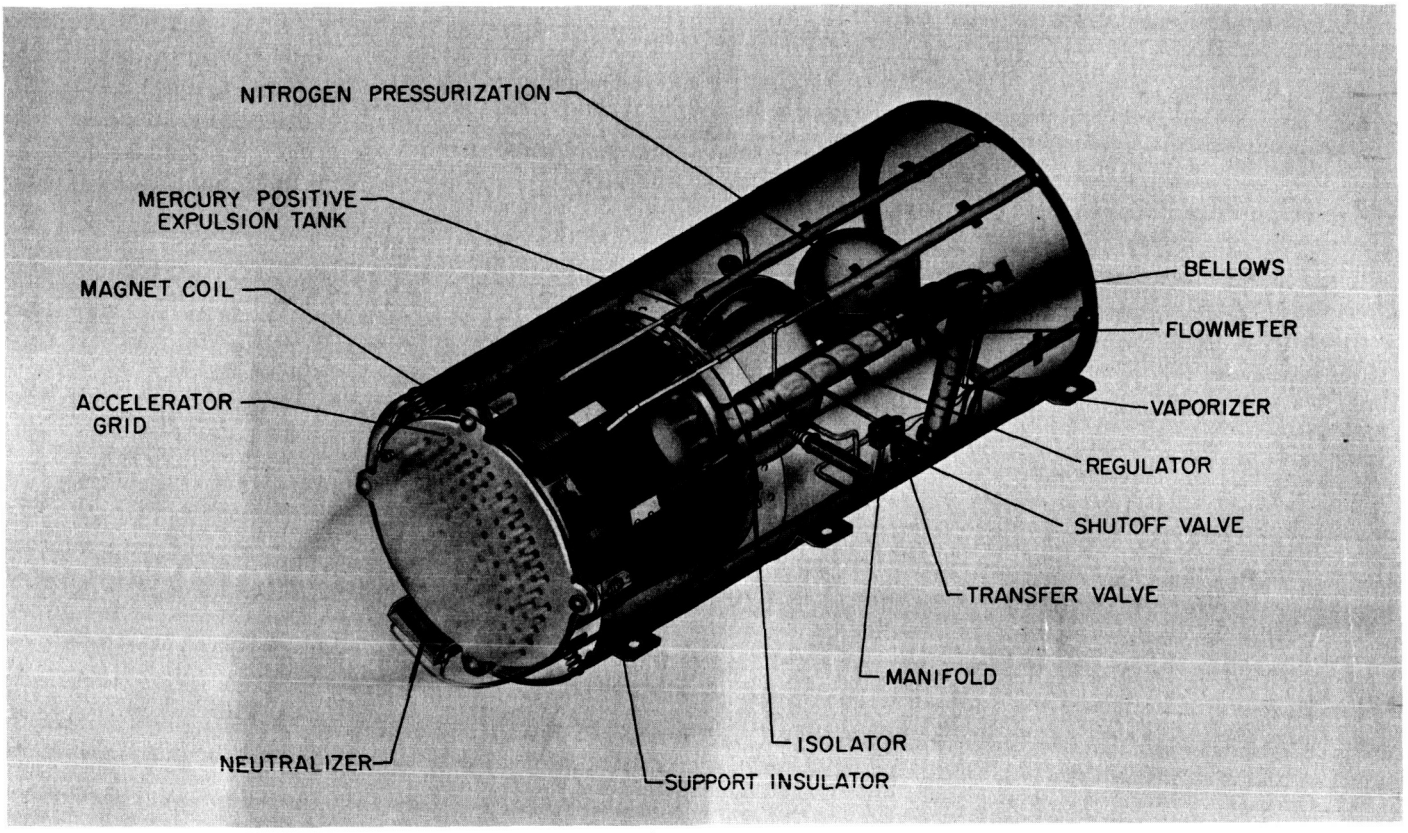


Fig. 10. Conceptual drawing of thrustor module using the described propellant feed system

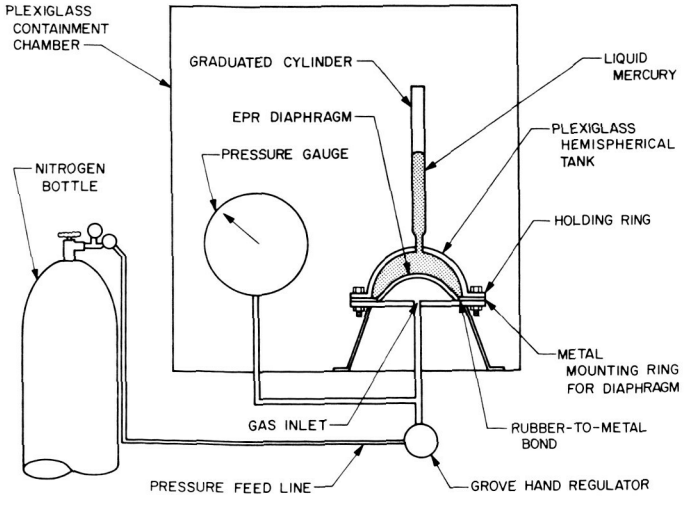


Fig. 11. Setup for determining the tensile properties of elastomers in contact with mercury

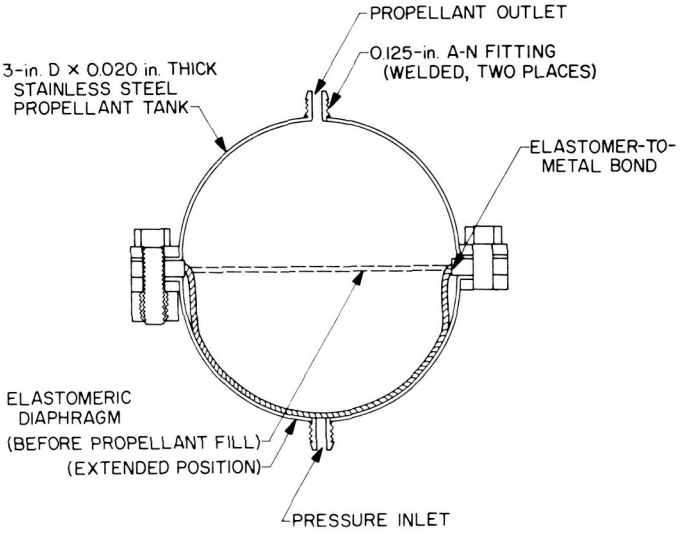


Fig. 12. Propellant storage tank for testing elastomeric diaphragms

of ethylene-propylene (EPR); the samples have been sent out for detailed chemical analysis. Meanwhile, a further test of EPR is being set up. A diaphragm of this material will be placed in a hemispherical test rig filled with mercury, and pressure applied to partially stretch the diaphragm. Changes in its tensile properties will be noted by observing how the displaced volume changes with

time with constantly applied pressure. The test setup is shown in Fig. 11.

Two designs of metal tanks with elastomeric diaphragms have been developed, and are shown in Figs. 12 and 13. The first, shown in Fig. 12, is designed primarily for evaluation of diaphragms in the test system. The diaphragm has been bonded to a metal ring for bolting or welding into the propellant tank. Bonded elastomeric diaphragms have been fabricated and are awaiting test to determine the integrity of the bond under stress. In the second design, the bond has been replaced by a mechanical seal, as shown in Fig. 13. The tank, which is a modification of the final tankage design for the ALPS program (Ref. 7), is 6 in. in diameter, and is to be made of 0.030 in. thick titanium. When fabricated, it is expected to weigh 0.55 lb, or approximately 1.4% of the propellant weight, and represents a realistic flight weight system. This will be incorporated into the final system sometime next year.

e. Pressure regulator. The pressure regulator will be built in two stages: the first will reduce the pressure from 3000 to 150 psi, and the second will perform the final reduction. For the first stage, a slightly modified commercial design will be used which, it is estimated, will weigh

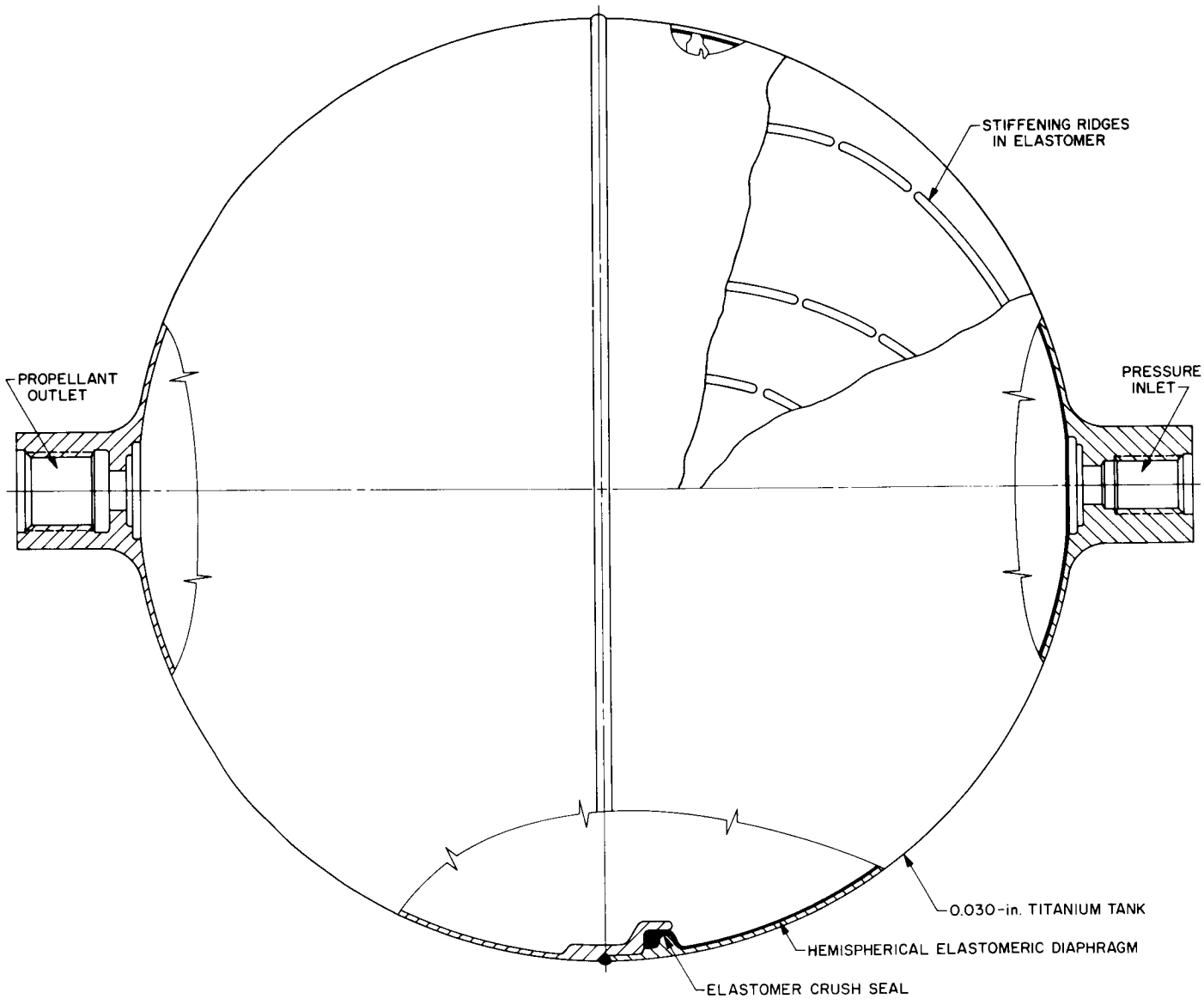


Fig. 13. Design of propellant tank for final feed system

approximately 0.8 lb. The second stage must accept a pressure feedback from the liquid line to regulate its output pressure. Such a stage has been designed and fabricated, and is shown in Fig. 14 It is made of aluminum and weighs 0.85 lb. This weight will be reduced considerably in the final system when the bolted construction is replaced by a welded construction. The regulator is currently being tested; initial results indicate some modifications are necessary before it can be incorporated into the test system.

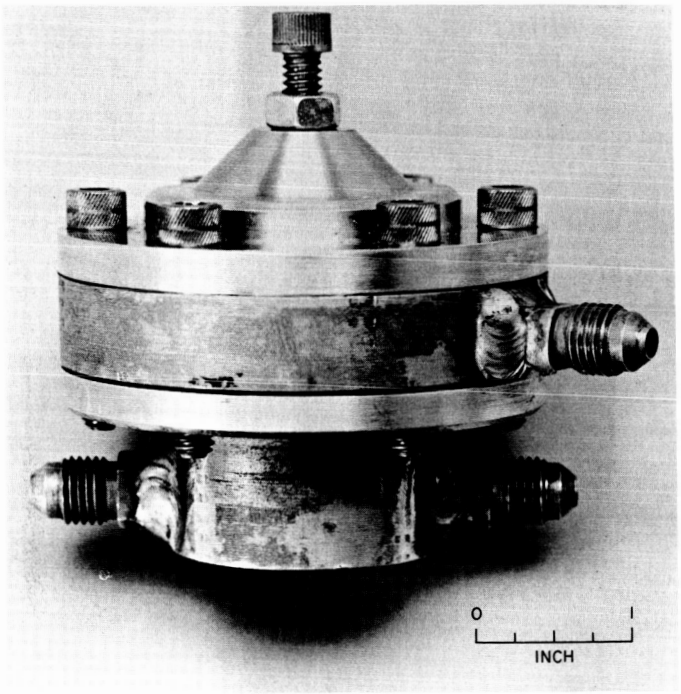


Fig. 14. Second-stage pressure regulator

f. Pressure storage tanks. The fiberglass pressurization tanks described in Ref. 6 are still to be used in the final system. However, until the regulator has been completely checked out, no attempt will be made to incorporate the tanks into the test system.

3. Test System Operation

The present test system includes the 5-in. unheated isolator, the 3-in. vaporizer with two heating coils, the HRL solenoid valve, and the bolted metal diaphragm storage tank. This system is shown mounted on the vacuum tank header in Fig. 15. Expulsion pressure is supplied from an external source, and a pressure transducer in the liquid expulsion line monitors the liquid pressure.

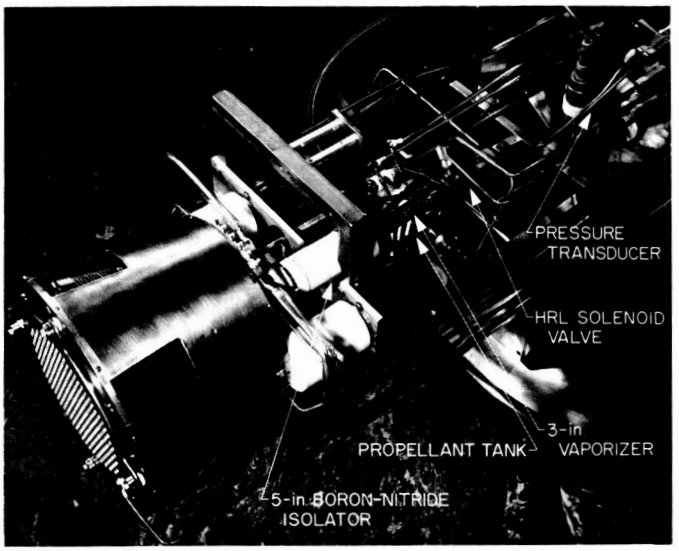


Fig. 15. Test feed system and thrustor mounted on vacuum tank header

To date, the feed system has been operated for about 200 hr in its present configuration. At the levels of current operation, this means that about 15% of the total propellant inventory has been used. No problems have been observed from pressure surges due to diaphragm "popping"; apparently the expansion process is proceeding in a relatively smooth manner. The liquid pressure head has been maintained between 4 and 6 psi to assure that the liquid-vapor interface will remain at the vaporizer inlet. This has required expulsion pressures ranging from 4.5 to 14 psia, with the value rising slowly, but nonuniformly, with time. This is presumed to be due to the gradual expansion of the diaphragms, with the amount of pressure transmitted through the diaphragm changing as its configuration changes. It is anticipated that with elastomeric diaphragms this variation would be a much smoother process; however, this must still be tested.

In all testing, the liquid-vapor interface has been maintained at the vaporizer inlet. As expected, flow rate is primarily controlled by the vaporizer inlet temperature; the calibration curve for the current vaporizer is shown in Fig. 16. It has been verified that small flow-rate changes could be controlled by the vaporizer temperature gradient; however, irregularities in the engine performance have prevented a real performance mapping of this control function. These irregularities, traced to the cathode, have been causing a random $\pm 10\%$ variation in the beam current. This variation is sufficient to completely mask the 1 to 2% control expected to be achieved by the temperature gradient.

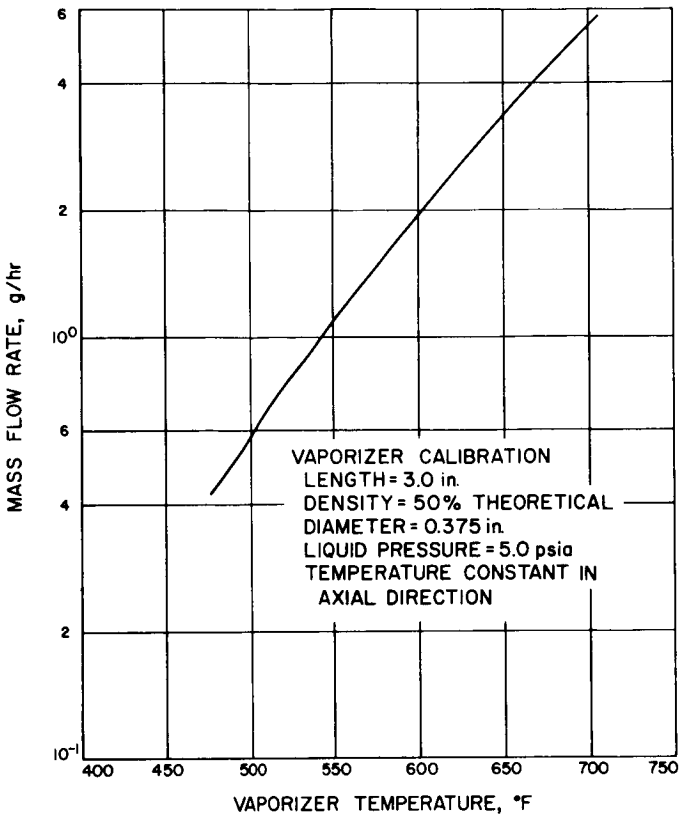


Fig. 16. Calibration curve for 3-in. vaporizer

During the entire test period no high-voltage breakdowns have been observed in the isolator. Tests have been performed over flow rates ranging from 0 to 7 g/hr, or up to about twice the design flow rate for the engine.

Table 1. Component weights and power requirements based on an engine operating at 4000 sec I_s with a power input of 2 kw

Component	Present test system		Final system	
	Weight, lb	Power, w	Weight, lb	Power, w
Isolator	0.24	0	0.18	0
Vaporizer	0.44	23	0.44	23
Solenoid valve	0.75	1.0	0.75	1.0
Propellant tank	5.0	0	0.85	0
Regulator	1.65	0	1.20	0
Pressure tank	1.12	0	1.12	0
Total	9.20	24	4.54	24
Specific weight	$4.60 \frac{lb}{kw}$		$2.27 \frac{lb}{kw}$	

In summary, the feed system has behaved as expected over about 200 hr of system operation. The metal diaphragms have given no indication of large pressure surges, as initially feared; however, on the basis of the observed expulsion pressure fluctuations, it would still seem desirable to use the elastomeric diaphragms if a suitable material can be found. Power requirements and weights are modest, as indicated in Table 1, thus assuring that the design goals of a specific weight less than 2.5 lb/kw into the engine and a power requirement less than 2% that of the total propulsion system will be met. It is expected that a completely self-contained test system will be tested by the middle of October.

D. Spectroscopic Measurement of Plasma Electron Temperature Using the Relative Intensities of Argon II Lines

N. Nerheim

1. Introduction

To fully describe the state of the plasma, one of the properties that must be known is the electron temperature. Knowledge of the electron temperature is of particular importance in some applications, such as power generation (Ref. 8) and plasma acceleration (Ref. 9), in which the necessary ionization is produced by collisions between neutral particles and highly energetic electrons. In the experiments described in a previous report (Ref. 10), the electron temperature of an argon plasma produced by a magnetoplasmadynamic (MPD) source was measured by two well-known methods. A Langmuir probe was used in the cooler, downstream portion of the exhaust plume, and a spectroscopic method, i.e., the relative line intensity method, was used for temperature determination near the arc exit. Because the latter measurements were made along a single line of sight, the measured quantity was an integrated light intensity from zones of different local emission; therefore, the reported temperature was an average of some kind. In the work described in this report, spectral measurements of the emission from a number of Ar II lines were made along a diameter of the exhaust plume of the MPD source operating at the conditions reported in Ref. 10. The observed intensity

profiles were then inverted by using Abel's transformation to obtain radial distributions of the local emission, and the radial temperature profiles were determined from the local emission.

2. Determination of Radial Intensities

A cross section of the plume is shown in Fig. 17 with the plume centerline coincident with the z axis and the direction of observation parallel to the y axis. The problem is to determine the local emission from each point along the line of sight from the measured intensity. For a cylindrically symmetric and optically thin plasma, the measured intensity may be expressed by

$$N(x) = 2\int_0^{(r_0^2-x^2)^{1/2}} \epsilon(r)\,dy = 2\int_x^{r_0} \frac{\epsilon(r)\,r\,dr}{(r^2-x^2)^{1/2}} \qquad (1)$$

where $N(x)$, the observed quantity, is the spectral radiance in the y direction at the position x, and $\epsilon(r)$, which is a function of r only because of the cylindrical symmetry, is the desired local emission coefficient at the radius r. Eq. (1) is an integral equation which may be solved in terms of $\epsilon(r)$ using Abel's transformation (Ref. 11):

$$\epsilon(r) = -\frac{1}{\pi}\int_r^{r_0} \frac{N'(x)\,dy}{(x^2-r^2)^{1/2}} \qquad (2)$$

A number of methods are available to obtain $\epsilon(r)$ from either Eq. (1) (Ref. 12) or Eq. (2) (Refs. 13 and 14). The method used in the present work is described by Bockasten (Ref. 15). The procedure is to divide the x axis into n equal intervals and to approximate the experimental curve, $N(x)$, over each interval by a third-degree polynomial, $P(x)$. Integration is then carried out in a piecewise fashion according to

$$\epsilon_j = -\frac{n}{\pi r_0}\sum_{k=j}^{n-1}\int_k^{k+1} \frac{P_k'(t)\,dt}{(t^2-j^2)^{1/2}} \qquad (3)$$

where

- ϵ_j is the value of $\epsilon(r)$ at $r_j = j\,r_0/n$ (j = 0,1,2, $\cdots$ n)
- k is the index on $x(x_k = k\,r_0/n;\ k = 0,1,2, \cdots n)$
- r_0 is the radius of the luminous plume
- t is a variable introduced by the substitution $x = t\,r_0/n$
- $P_k'(t)$ is the derivative of the polynomial $P_k(t)$ with respect to t.

$P_k(x)$ is Lagrange's form of the polynomial, substituted for the experimental curve over the interval x_k to x_{k+1}, that fits exactly the points N_{k-1}, N_k, N_{k+1}, and N_{k+2}. The results of the integration are cast in the form

$$\epsilon_j = r_0^{-1}\sum_k a_{jk}\,N_k \qquad (4)$$

where the summation is from $k = j-1$ to $n-1$ for $j > 0$, and from $k = 0$ to $n-1$ for $j = 0$.

Complete sets of the coefficients a_{jk} are tabulated in Ref. 15 for $n = 10$ and $n = 20$, and a partial set, for values of $j > 29$, is tabulated for $n = 40$. Bockasten has published a formula for most of the coefficients as a function of n. Because it may be convenient to divide the radius into intervals other than 10, 20, or 40, the equations for the rest of the coefficients have been determined.

The general coefficient a_{jk}, given by Bockasten, is

$$a_{jk} = n(12\pi)^{-1}\left[f(j,k+2) - 4f(j,k+1) + 6f(j,k) - 4f(j,k-1) + f(j,k-2)\right] \qquad (5)$$

where $j+1 < k < n-2;\ k > 2$.

$$f(j,k) = (3j^2+6k^2-2)\ln\left[k+(k^2-j^2)^{1/2}\right] - 9k(k^2-j^2)^{1/2} \qquad (6)$$

where $j \le k \le n-1;\ k > 0;\ j < n-1$.

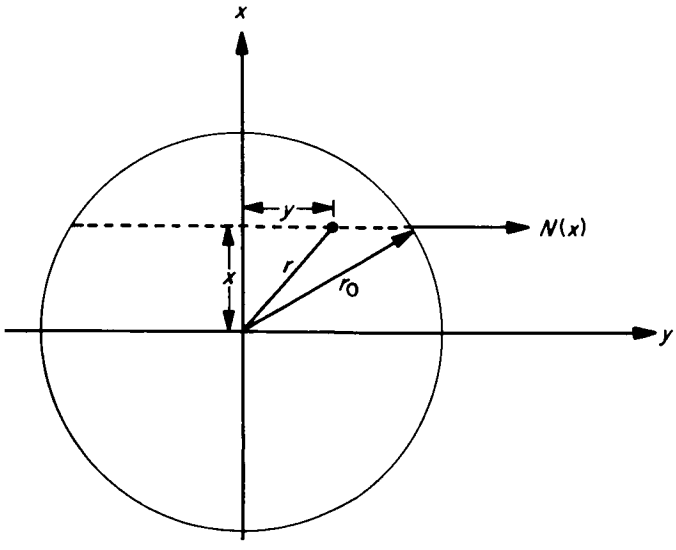


Fig. 17. Schematic view of plume cross section

Because special polynomials must be used to approximate the ends of the $N(x)$ curve, the above formula is not applicable for k less than 3 or greater than $n-3$. $P_0(x)$, the polynomial used for the interval from the centerline to x_1, is formed from N_0, N_1, and N_2 and the condition $P_0'(0) = 0$; $P_{n-1}(x)$ is a second-degree polynomial formed from the points N_{n-2}, N_{n-1}, and zero for N_n. Special formulas must also be determined for $a_{j,j-1}$, $a_{j,j}$, and $a_{j,j+1}$, because the general coefficient a_{jk} is formed from the sum of four integrals (there are five terms because there are five limits of integration), each of which include the experimental reading N_k, provided k is greater than $j+1$. For $k = j, j+1$, and $j-1$, the number of integrals including N_k are three, two, and one, respectively. The equations for the special cases are given below in terms of $A_{j,k}$, where $A_{jk} = \left(\frac{12\pi}{n}\right) a_{jk}$. The auxiliary function $f(j,k)$ is given by Eq. (6).

For calculations involving large values of n, a word of caution is in order. Because the coefficients are very small numbers that result from summing individual terms that increase rapidly in magnitude as n increases, it is necessary to carry a large number of significant figures in the calculation. It should also be noted that, because the method used involves the slope of the $N(x)$ curve, it is necessary to approximate the experimental curve with a very smooth curve to prevent fluctuations of the slope that may result from small errors in the experimental readings. The seven-point, third-degree smoothing formula, suggested by Bockasten, was used in this work.

$$A_{0,0} = 13.5 + 22 \ln 2$$

$$A_{0,1} = -18 - 36 \ln 2 + 52 \ln 1.5$$

$$A_{0,2} = 4.5 + 18 \ln 2 - 114 \ln 1.5 + 94 \ln (4/3)$$

$$A_{j,j-1} = f(j,j+1) - (9j^2 + 12j + 4) \ln j; \; (j < n-1)$$

$$A_{j,j} = f(j,j+2) - 4(j,j+1) + 3(9j^2 + 8j - 2) \ln j; \; (j < n-2)$$

$$A_{j,j+1} = f(j,j+3) - 4f(j,j+2) + 6f(j,j+1) - 3(9j^2 + 4j - 4) \ln j; \; (j < n-3)$$

$$A_{j,n-2} = -3f(j,n-1) + 6f(j,n-2) - 4f(j,n-3) + f(j,n-4) - 12\,(n^2 - j^2)^{1/2} + 6(2n-1) \ln \left[n + (n^2 - j^2)^{1/2}\right]; \; (j < n-3)$$

$$A_{j,n-1} = 3f(j,n-1) - 4f(j,n-2) + f(j,n-3) + 24\,(n^2 - j^2)^{1/2} - 24(n-1) \ln \left[n + (n^2 - j^2)^{1/2}\right]; \; (j < n-2)$$

$$A_{n-3,n-2} = -12\,(6n-9)^{1/2} + 6(2n-1) \ln \left[n + (6n-9)^{1/2}\right] - 3f(n-3,n-1) + 6f(n-3,n-2) - 3(9n^2 - 50n + 65) \ln (n-3)$$

$$A_{n-2,n-2} = -24\,(n-1)^{1/2} + 6(2n-1) \ln \left[n + 2\,(n-1)^{1/2}\right] - 3f(n-2,n-1) + 3(9n^2 - 28n + 18) \ln (n-2)$$

$$A_{n-2,n-1} = 48\,(n-1)^{1/2} - 24(n-1) \ln \left[n + 2\,(n-1)^{1/2}\right] + 3f(n-2,n-1) - 3(9n^2 - 32n + 24) \ln (n-2)$$

$$A_{n-1,n-2} = -12\,(2n-1)^{1/2} + 6(2n-1) \ln \left\{\left[n + (2n-1)^{1/2}\right]/(n-1)\right\}$$

$$A_{n-1,n-1} = 24\,(2n-1)^{1/2} - 24(n-1) \ln \left\{\left[n + (2n-1)^{1/2}\right]/(n-1)\right\}$$

3. Relative Line Intensity Method

Thorough discussions of the relative line intensity method may be found in a number of standard texts (Refs. 16 and 17), and will be covered very briefly here. The emission coefficient of a transition from electronic level m to level n, where the energy levels have a Boltzmann distribution, may be expressed as

$$\epsilon = \frac{K A_{mn} g_m e^{-E_m/kT}}{\lambda} \tag{7}$$

where

A_{mn} is the transition probability

g_m is the statistical weight of the upper level

λ is the wavelength of the spectral line

E_m is the energy of the upper level

T is the temperature that characterizes the Boltzmann distribution of energy levels.

Eq. (7) may be rearranged as follows:

$$\ln\left(\frac{\epsilon\lambda}{A_{mn} g_m}\right) = -E_m/kT + \text{const.}$$

T is then determined from the slope, $\frac{-1}{kT}$, of a plot of

$$\ln\left(\frac{\epsilon\lambda}{A_{mn} g_m}\right) \text{ vs } E_m.$$

4. Experimental Methods

The plasma source was an MPD arc that was operated at 1000 amp with an argon flow rate of 0.3 g/sec and at an ambient pressure of about 400 μ. A more complete description of the source is given in Ref. 10. Spectroscopic measurements of the line intensities were made with a Jarrell-Ash 0.5-m, Ebert-mounted, plane grating monochromator and a RCA 7265 photomultiplier (PM) tube. The grating–PM tube combination was calibrated with a tungsten strip lamp. The measurements were made at positions not greater than 5 cm downstream of the arc exit because the SNR at positions farther downstream was too small for accurate measurement. To obtain intensity profiles across the plume diameter, an image of the plume was sought across the entrance slit of the monochromator by moving a mirror, which was part of the optical train, parallel to the plume diameter, as shown in Fig. 18. A second mirror, which moved at one-half the speed of the first, was included to maintain the length of the optical path from the plume to the entrance slit at a constant value. The mirror movement was measured with

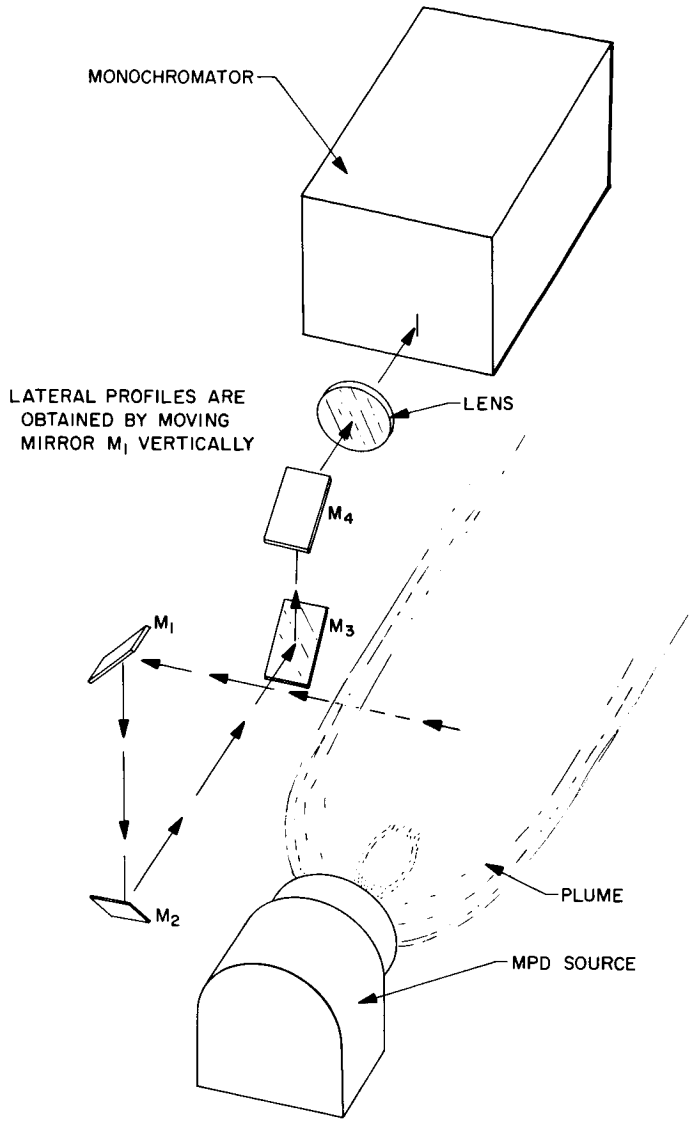


Fig. 18. Optical path from the plasma to the monochromator

a rectilinear potentiometer; the output of the potentiometer was connected to one axis of an x-y recorder, and the signal from the PM tube was connected to the other axis. Each profile was folded about the axis of symmetry, and a smooth curve was drawn through the two curves. The difference between the two curves was not greater than 2% at all radial positions where the measured intensity exceeded 5% of the centerline intensity.

5. Results and Discussion

The spectral lines selected for this work were 14 Ar II lines for which Olsen (Ref. 18) has published experimental values of the transition probability. The ion lines were

used in preference to the atom lines because of their greater intensities in the spectral region where the response of the 7265 PM tube is high.

Lateral and inverted profiles of the emission from the λ4806 line at two axial positions, 1 cm and 5 cm downstream of the arc exit, were superimposed on the same plot for each axial position and are shown in Figs. 19 and 20, respectively. Twenty-five data points were read from the experimental curve, and the smoothed values that resulted from the least-squares fit of the seven-point, third-degree smoothing formula were divided by the centerline value. These are shown in the figures as the solid curves. Straight line segments connect the points. The inverted profiles, which were normalized by dividing each point by the centerline value, are shown as circles with no connecting line. The position of the anode wall and the plume radius, r_0, are shown in each figure.

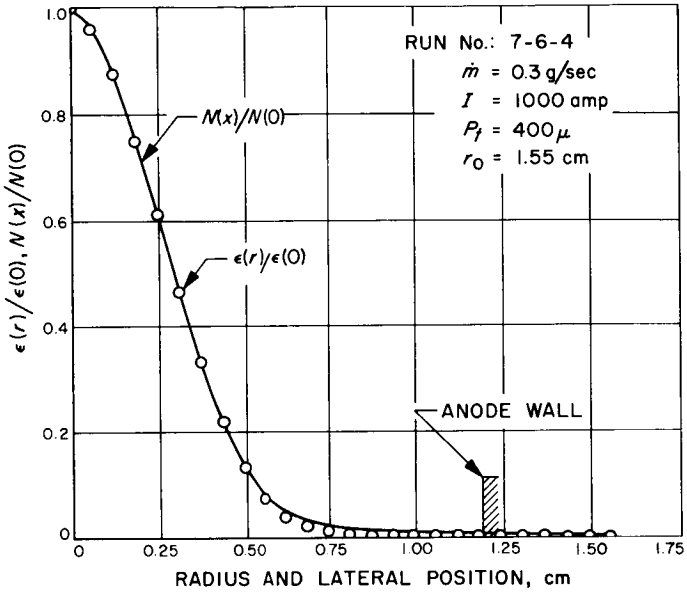


Fig. 19. Radial and lateral profiles of λ4806 Ar II line intensity of MPD source at 1 cm downstream

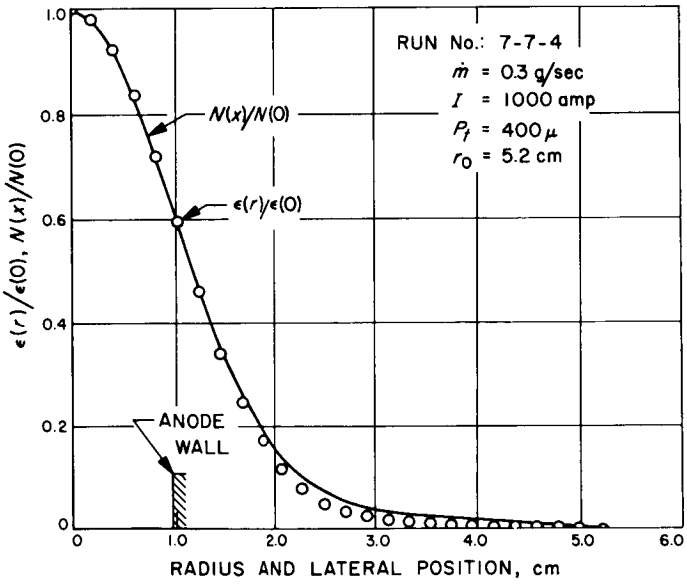


Fig. 20. Radial and lateral profiles of λ4806 Ar II line intensity of MPD source at 5 cm downstream

Similar profiles were obtained for each of the 14 Ar II lines. Only 13 of these lines were used for the temperature determination because the intensity of one, the λ4847 line, was too large by more than a factor of ten. A reason for this discrepancy has now been found.[1] The determination of the centerline temperature at the 1-cm position is shown in Fig. 21. Because relative intensities of the lines are used, the ordinate contains a constant to give the numerical values shown in the figure. The straight line drawn

[1]Olsen, H. N., Private communication. The experimental and theoretical values of the transition probability of the λ4847 line, given in Ref. 18, should read 9.5×10^7 sec^{-1} and 8.1×10^7 sec^{-1}, respectively.

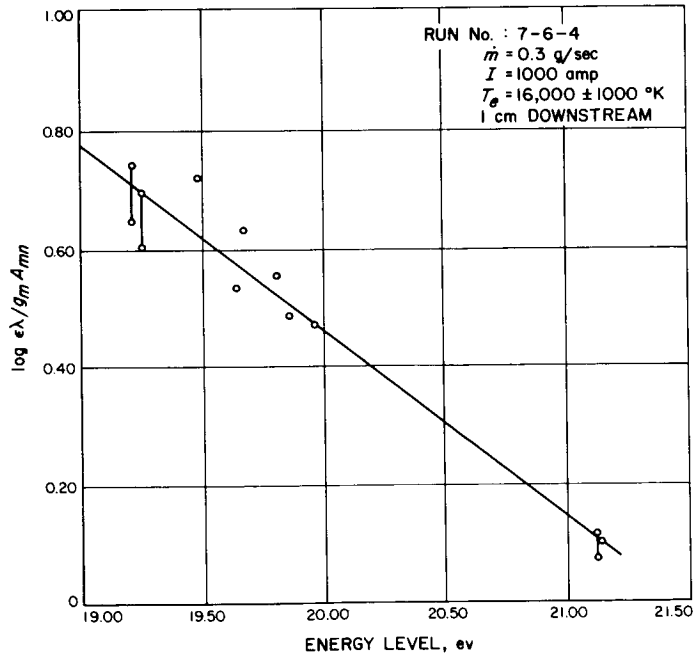


Fig. 21. Electron temperature at plume centerline determined from the relative intensities of Ar II lines

through the data was found by a least-square fit, and the error associated with the temperature, which was found from the slope of the line, was calculated by the normal procedure of the least-squares method (Ref. 19).

Temperature profiles at the two axial positions are shown in Figs. 22 and 23 as shaded areas that, for each of the 25 radial positions, are centered on the calculated temperature and cover the error limits. The "average" temperature, determined from the single centerline intensity of each line, is also shown. Because of the flatness of the temperature profiles, the average temperature that was reported in Ref. 10 is a valid temperature. The very large errors at radial positions greater than the fifteenth subdivision result from readings that are only a few percent of the centerline reading and consequently are of low accuracy.

The theory of the relative line intensity method assumes that the distribution of the different electronic energy levels involved in the transitions is Maxwellian. For this to be true, the reaction rates that determine the population densities of the excited levels must be dominated by collisions with electrons that have a Maxwellian velocity distribution. It is not necessary that local thermodynamic equilibrium (LTE) exist for these conditions to be satisfied. One example is the two-temperature model proposed by Kerrebrock (Ref. 20). In this model, all of the excited levels, the degree of ionization, and the free electron velocity distribution are in equilibrium at the electron temperature. It is only because the atoms are at a lower kinetic energy that the model departs from LTE. The criteria for the use of the relative line intensity method are less restrictive than either LTE or the two-temperature model described above. For instance, in work on electron-ion recombination rates, Byron, et al. (Ref. 21), have shown that, for hydrogen-like atoms, the excited states above a certain minimum level are all close to the equilibrium populations as given by the Saha equation.

Calculations of the instantaneous population densities of the excited levels of hydrogen-like ions as a function of the free electron density and temperature have been made by McWhirter and Hearn (Ref. 22). These calculations do not apply to this study because all 14 Ar II lines result from transitions occurring between levels that have the same principal quantum number, i.e., the ion is not hydrogen-like. McWhirter and Hearn make the assumption that the sublevels of each level, specified by the principal quantum number, are populated according to their statistical weights by electron collisions.

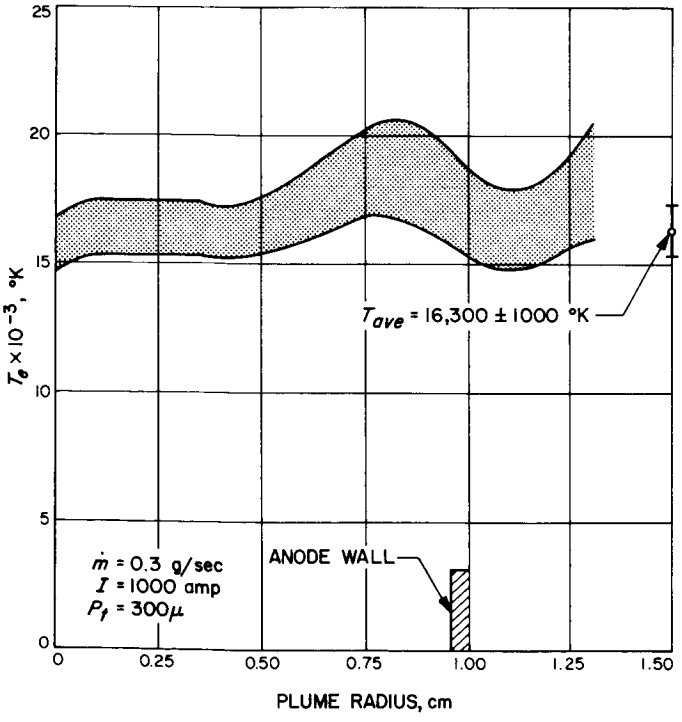


Fig. 22. Radial electron temperature profile of argon MPD source at 1 cm downstream

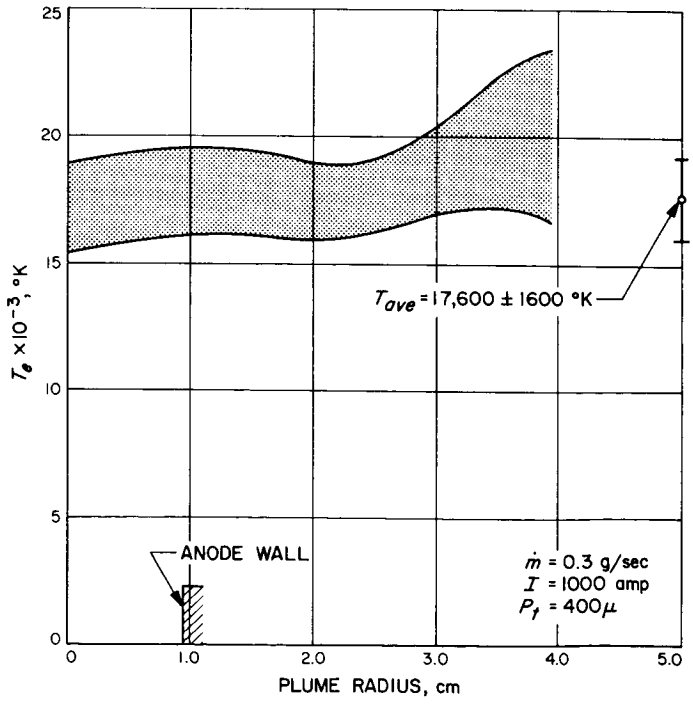


Fig. 23. Radial electron temperature profile of argon MPD source at 5 cm downstream

The average enthalpy of the plasma, determined from measured values of input power, cooling losses, and the argon flow rate, was too low for the degree of ionization to be in equilibrium with the measured electron temperature. This is to be expected at the low plasma pressure. Thus, there is no *a priori* reason to expect a Maxwellian distribution of the energy levels. It is believed that the free electrons do have a near-Maxwellian velocity distribution because electrons easily thermalize through electron-electron collisions. Because no independent measurements of the electron temperature were made at the same axial position, the only justification that can be made for equating the measured temperature to the electron temperature, aside from the reasonableness of the value, is how well the data points fall on a straight line. The scatter of the data shown in Fig. 21, which does not include the λ4847 line, results in a temperature error of about 6.5%. Of the 13 lines used, there are three pairs each of which has a common upper energy level. For example, both members of the pair λ4806 and λ5009 have the common upper level of 19.22 ev. The data points from the three pairs have been joined by vertical lines in the figure. Because each pair has a common upper level, the separation of the two points of each pair must be caused by deficiencies in the experimental technique and not by departures from a Maxwellian distribution of the population densities of the excited states.

That the measured temperature is a reasonable value is demonstrated, in part, by the results obtained with a Langmuir probe (Ref. 10) at positions farther downstream. The electron temperature dropped rapidly from about 9000°K at the 20-cm position to about 5000°K at 100 cm. The electric energy is imparted directly to the electrons in the arc; then, by inelastic and elastic collisions, the electrons transfer the kinetic energy gained in the electric field to the heavy particles. Elastic losses would result in a rapidly decreasing electron temperature in the field-free region.

To adequately demonstrate the applicability of the relative line intensity method to temperature measurements of this study, it will be necessary to reduce the experimental errors and to obtain an independent measurement of the electron temperature. In the future, a mechanical light chopper, in conjunction with a lock-in amplifier, will be used to improve the SNR. This should also permit more reliable measurements at positions farther downstream. In addition, an effort will be made to obtain Langmuir probe measurements closer to the arc exit. This can be done either by cooling the probe or by rapidly swinging the probe through the plasma.

E. Analog Study of Thermionic Reactor Stability

H. Gronroos and G. Kikin

1. Introduction

The prediction of the transient behavior and general stability characteristics of a fast liquid metal cooled in-pile thermionic reactor for space powerplant application presents novel problems not encountered in other reactor types. These problems have not been investigated to date, although some general thoughts on the subject have been reported (Ref. 23). With the development of better thermionic diodes and various reactor concepts, a dynamic analysis has become important in order to introduce further design constraints and provide an extended basis for screening the proposed concepts. A general kinetic study is presently being pursued at JPL by the Energy Sources Group of the Propulsion Research and Advanced Concepts Section. The purpose of this study is to provide information on the problem in general and to determine where simplifications may be applied with a minimal sacrifice in accuracy.

The highly nonlinear character of the system equations makes a linearized analytical description valid over only a very small range. Such an analysis also may introduce oscillatory modes of behavior not physically existing or may fail to predict existing modes. As a first approach, a lumped parameter model using the analog computer with time as an independent variable is most practical for this type of nonlinear problem. Omission of the spatial dependence is disturbing, particularly for an in-pile thermionic reactor, since the conversion performance is highly sensitive to the heat and temperature inputs. Power flattening, employing fuel zoning and reflectors, is necessary to obtain an optimized system (Ref. 24), and it also helps to justify a lumped parameter model. Introduction of spatial variables and the analytical treatment of the nonlinear equations are being studied concurrently with the analog simulation. These studies will be reported at a later date.

2. Problem Definition

Fig. 24 shows the general layout of a thermionic reactor with the thermionic diodes stacked in series—the so-called "flashlight" concept. The diode module shown, which is taken from Douglas Aircraft Co. Report SM-46275, represents one of many similar proposed designs.

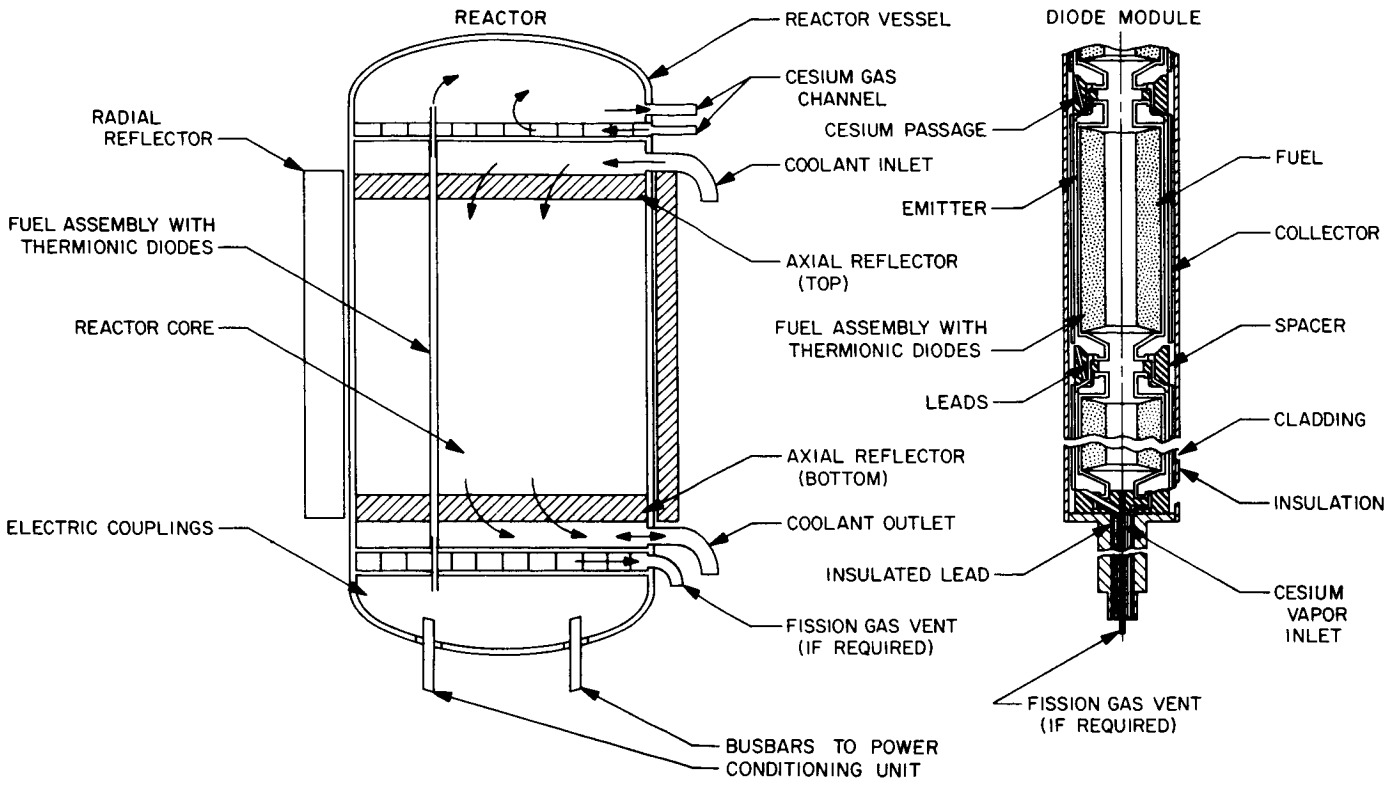


Fig. 24. Design of thermionic reactor with series stacked diodes

The direct coupling between the fuel and electrical power and the loose coupling to the coolant make the load-following characteristics of the plant dependent primarily on the fuel and emitter temperature coefficients. The fuel-to-coolant heat transfer mechanism is unique, and adjacent thermionic elements are almost decoupled thermally. A change in the electrical power demand immediately influences the emitter and then the nuclear fuel. The coolant is not leading as in a conventional system, but is lagging. The system, therefore, exhibits two distinct parallel loops (Fig. 25). The temperature distribution within the reactor changes with the load impedance, and some form of electrical output–nuclear reactor coupled control must be employed. Depending on the temperature coefficients, the reactor may or may not be inherently stable against load perturbations. However, the performance of the power conditioner and coolant loops further complicates the situation, and conditions may be present for the existence of stable centers and limit cycles.

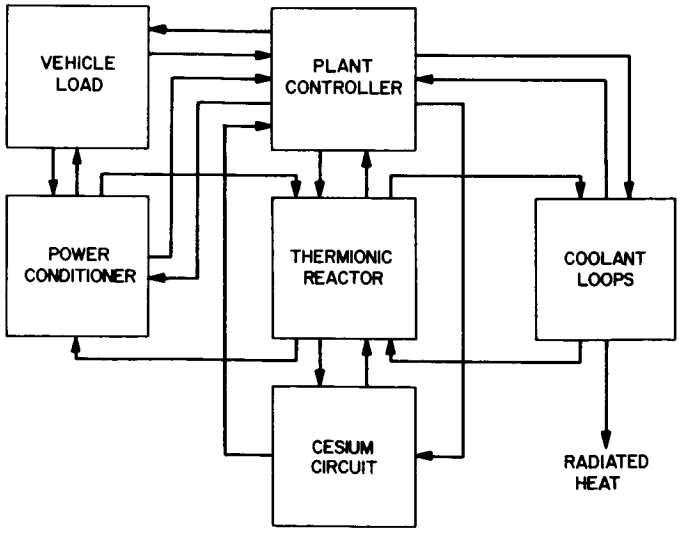


Fig. 25. Main components in thermionic space reactor powerplant

Fig. 26 shows a block diagram of the system. Incomplete preliminary results have been obtained, and the evaluation of this phase of the investigation is continuing. The plant controller is not simulated since it can best be defined after the steady state and transient behavior is established.

3. *Neutron Kinetics*

Fig. 27 illustrates the general neutronic situation in a typical thermionic reactor, and pertinent parameters are

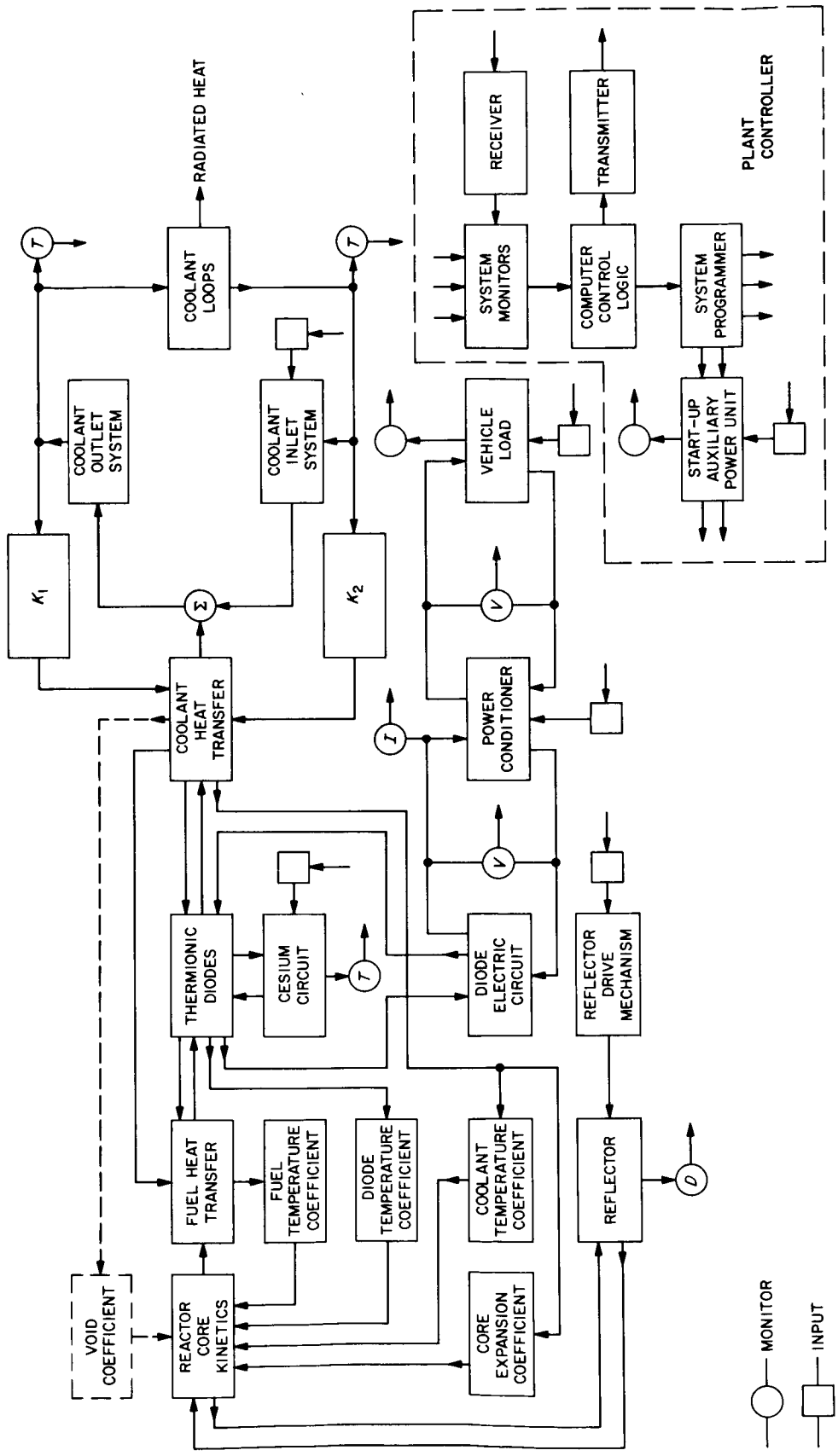


Fig. 26. Simplified diagram for thermionic space powerplant stability analysis

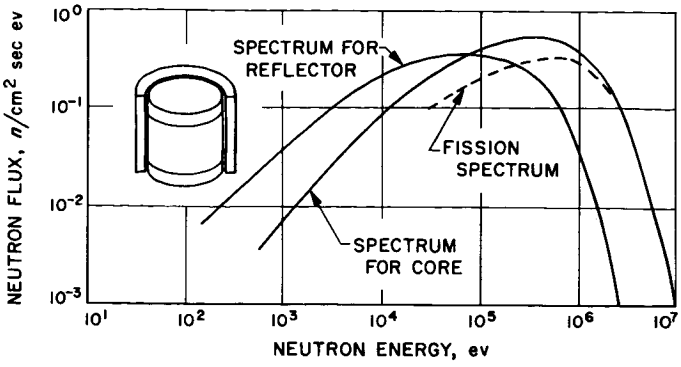


Fig. 27. Neutron spectra in a thermionic reactor

shown in Table 2. For a small, fast neutron system such as this, the point kinetics approximation, suitably augmented to take into consideration reflector control (Ref. 25), seems applicable. However, for prompt criticality cases this model may be inadequate (Ref. 26) and must be further qualified. Neglecting delayed neutron diffusion to the reflector and generated photo neutrons there, the point kinetics equations are:

$$\frac{dn_c}{dt} = \frac{[k_c(1-\bar{\beta})-1]}{l_c} n_c + \frac{n_r k_{rc}}{l_r} + \sum_{i=1}^{6} \lambda_i C_i \quad (1)$$

$$\frac{dn_r}{dt} = -\frac{n_r}{l_r} + \frac{n_c k_{cr}}{l_c} \quad (2)$$

$$\frac{dC_i}{dt} = \frac{n_c k_c \beta_i}{l_c} - \lambda_i C_i \quad (3)$$

Table 2. Parameters for neutron spectra in a thermionic reactor system

Component	Material	Temperature
Nuclear fuel	UC, UO_2, cermets	2000–3200°F
Coolant	NaK, Li-7	1200–1800°F
Construction material	Cb-1% Zr, W, Ta, alloys	2500°F
Reflectors	Be, BeO	
Parameter	**Location**	**Dimension**
Core diameter		1–3 ft
Core height		1–3 ft
Fuel volume fraction		0.3–0.6
Neutron flux level	In core	$\lesssim 10^{13}\ n\bar{v}$
	In reflectors	$< 10^{13}\ n\bar{v}$
Thermal power		0.5–30 Mw
Median neutron energy	In core	0.2–0.7 Mev
Reactor weight	Reactor core only	1000–10,000 lb

where

n_c = neutron density in reactor core

n_r = neutron density in reflector

l_c = neutron lifetime in reactor core

l_r = neutron lifetime in reflector

k_c = neutron reproduction in core

k_{cr} = probability that a disappearing core neutron will reappear in the reflector

k_{rc} = probability that a disappearing reflector neutron will reappear in the core

$\bar{\beta} = \sum_{i=1}^{6} \beta_i$; total fraction of delayed neutrons

β_i = fraction of delayed neutrons belonging to group i

C_i = percursor concentration in group i

It is seen that the reflector adds an extra delayed group which is characterized by both the reciprocal of the prompt neutron lifetime in the reflector and the fraction of neutrons returned to the core. The steady state criticality condition is determined by

$$k = k_c + k_{cr} k_{rc} = 1 \quad (4)$$

k_c may be expressed in terms of the reactivity coefficients for the core as follows:

$$k_c = (1 - \rho_c \bar{\beta})^{-1} \quad (5)$$

$$\rho_c = \rho_{co} - \alpha_j (T_j - T_{jo}) \quad (6)$$

where

ρ_{co} = reactivity in \$ units in core for criticality at equilibrium temperature distribution

T_{jo} = mean temperature in region j at equilibrium

α_j = temperature coefficient of reactivity for region j

For a reflector controlled reactor, it is reasonable to postulate

$$\frac{k_{rc}}{k_{cr}} = \text{constant less than } 0.5 \quad (7)$$

For the purpose of the analog simulation reactor, invariants are defined from the relations

$$\Lambda_c = \frac{l_c}{k_c} = \text{constant} \quad (8)$$

and

$$\Lambda_r = \frac{l_r}{k_{cr} k_{rc}} = \text{constant} \quad (9)$$

Eqs. (8) and (9) are somewhat arbitrary and inherently assume the separability of the space and time dependence of the neutron flux. Numerical values of $\Lambda_c = 10^{-6} - 10^{-7}$ sec and $\Lambda_r = 10^{-5} - 10^{-6}$ sec are suitable averages for the system discussed. Appropriate starting values for k_c, k_{cr}, and k_{rc} are 0.950, 0.350, and 0.143, respectively.

The temperature coefficients α_j are small, $\sim -10^{-6}$ $(\delta k/k)°K^{-1}$, except for the over-all delayed core expansion coefficient which is of the order -10^{-5} $(\delta k/k)°K^{-1}$.

4. Equations for the Thermionic Process

Although the thermionic conversion process is well understood, it is difficult to analytically reproduce experimental *I-V* characteristics. Aside from adequately describing the plasma in the interelectrode gap, factors such as emitter patchiness and thermal leakage through support structures introduce uncertainties. A simple analytical model is employed that reasonably accounts for the cesium pressure dependence, and, with the aid of empirical factors, the thermionic process and thermal leakage are adjusted to give appropriate *I-V* characteristics and efficiencies. Fig. 28 shows a simplified drawing of the thermionic module and the heat balance for emitter and collector. The equations are as follows:

a. Nuclear fuel to emitter

$$V_1\rho_1 c_{p_1}\, dT_1/dt + 8\pi L_{12}k_1(T_1 - T_2) = V_1\gamma_1 n_c \tag{10}$$

where

V_1 = fuel volume

ρ_1 = fuel density

c_{p_1} = fuel heat capacity

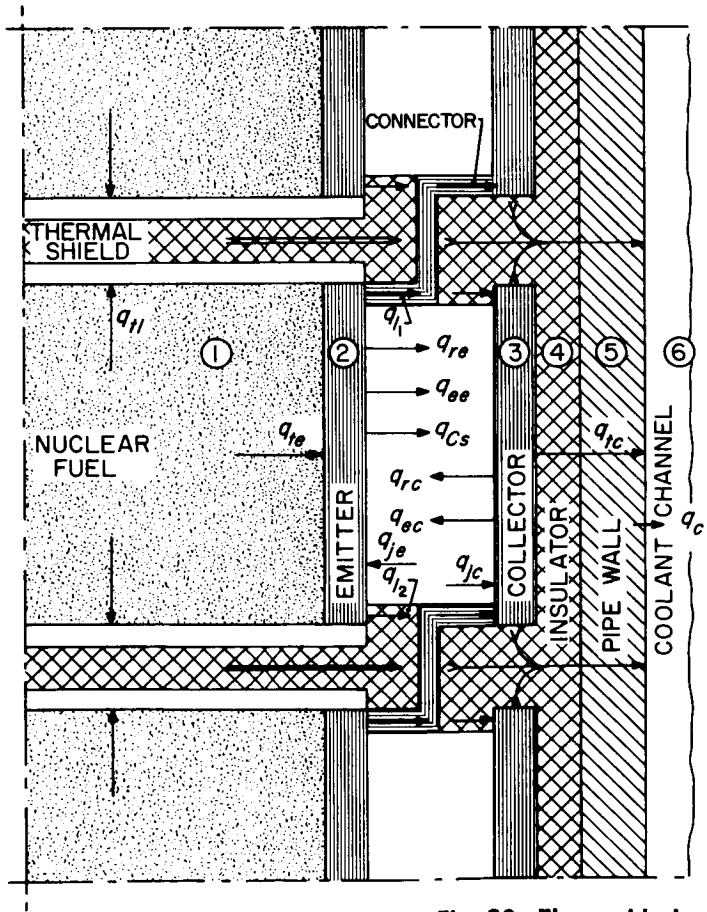
T = average temperature

L_{12} = fuel and emitter length

k_1 = fuel thermal conductivity

γ_1 = conversion coefficient for fission heat generation

n_c = core neutron density from the solution of Eq. (1)



q_{tc} = HEAT FLUX FROM COLLECTOR
q_{tf} = HEAT FLUX THROUGH FUEL END
q_{te} = HEAT FLUX INTO EMITTER
q_l = CONDUCTION HEAT LOSS
q_r = RADIATION HEAT LOSS
q_e = ELECTRON COOLING
q_{Cs} = CESIUM CONDUCTION LOSS
q_j = JOULE HEATING

BALANCE EQUATION FOR THE EMITTER:

$$q_{te} = q_{re} + q_{ee} + q_{Cs} + q_{l_1} + q_{l_2} - q_{rc} - q_{ec} - q_{je}$$

BALANCE EQUATION FOR THE COLLECTOR:

$$q_{tc} = q_{te} + q_{je} + q_{jc} - \beta q_{Cs} - \alpha\left(2q_{tf} + q_{l_1} + q_{l_2}\right)$$

α, β = FRACTION OF HEAT FLUX
e = SUBSCRIPT FOR EMITTER
c = SUBSCRIPT FOR COLLECTOR

Fig. 28. Thermal balance in a thermionic diode

b. Emitter heat balance

$$-V_2\rho_2 c_{p_2}\,dT_2/dt + 8\pi L_{12}k_1(T_1 - T_2) + I^2R_2/3 = \pi D_2 L_{12} q_2'' + k_{23}(T_2 - T_3) \quad (11)$$

where

I = total current

R_2 = total emitter resistance

k_{23} = leakage heat loss factor for emitter to collector

q_2'' = emitter heat loss to interelectrode gap

D_2 = emitter diameter

$$q_2'' = \frac{I}{\pi D_2 L_{12}}\Phi_2 + \frac{2k_B}{e}(J_2T_2 - J_3T_3) + \frac{\sigma}{1/\varepsilon_2 + 1/\varepsilon_3 - 1}(T_2^4 - T_3^4) + \frac{k_{Cs}(T_2 - T_3)}{t_{23}} \quad (12)$$

where

Φ_2 = effective emitter work function

k_B = Boltzmann constant

e = electron charge

σ = Stefan-Boltzmann constant

ε = emissivity

k_{Cs} = cesium thermal conductivity

t_{23} = interelectrode gap

J = current density

The main problem is to account for emitter work function dependence on temperature, plasma resistance, and plasma leakage resistance back to the emitter. Because several other uncertainties exist and since the qualitative behavior is of particular interest, the following simple correlations are used:

$$\phi_2 = \phi_{20} + \frac{T_2}{T_{Cs}} \cdot a_2 \quad (13)$$

$$R_{Cs} = a_{Cs}\,t_{23}\,P_{Cs}^{1/2} \quad (14)$$

where

a_2, a_{Cs} = empirical parameters

T_{Cs}, P_{Cs} = cesium temperature and pressure, respectively

The effective emitter work function, Φ_2, becomes

$$\Phi_2 = \phi_2 + k_v\,[V_o - (\phi_2 - \phi_3)] \quad (15)$$

where

$k_v = 0$ for $V_o < \phi_2 - \phi_3$ (downhill mode)

$k_v = 1$ for $V_o > \phi_2 - \phi_3$ (uphill mode)

V_o = output voltage

The expression for the net current is, therefore,

$$\frac{I}{\pi D_2 L_{12}} = J_2 - J_3 = A f_2 T_2^2\, e^{-e\phi_2/k_BT_2} - A\, e^{-e(\phi_3 - V_o)/k_BT_3} \quad (16)$$

where

A = Richardson-Dushman constant

f_2 = correction factor for emitter patchiness (Ref. 27)

Φ_3 = collector work function

In the above development, it is assumed that the thermionic-electrical behavior instantaneously follows the emitter and collector temperatures.

The equations for the electrical connections follow from summing the voltages in the series units and adding currents from parallel connections. For identical thermionic diodes (Fig. 29):

$$P_e = I_t V_t = N_p I \cdot N_s V = I_t^2 R_L \quad (17)$$

$$V = V_o - \frac{I^2(R_2 + R_3)}{3} - I^2 R_{Cs} \quad (18)$$

where

P_e = electrical power into power conditioner

I_t = current into power conditioner

V_t = voltage across power conditioner

R_L = resistive load

R_3 = collector resistance

It is intended that the electrical loop at R_L be closed. In accord with Ref. 23, a simple extension to include payload coupling can be accomplished as follows: The output regulation is characterized by

$$V' = V_o'(1 - KI'/I_o') \quad (19)$$

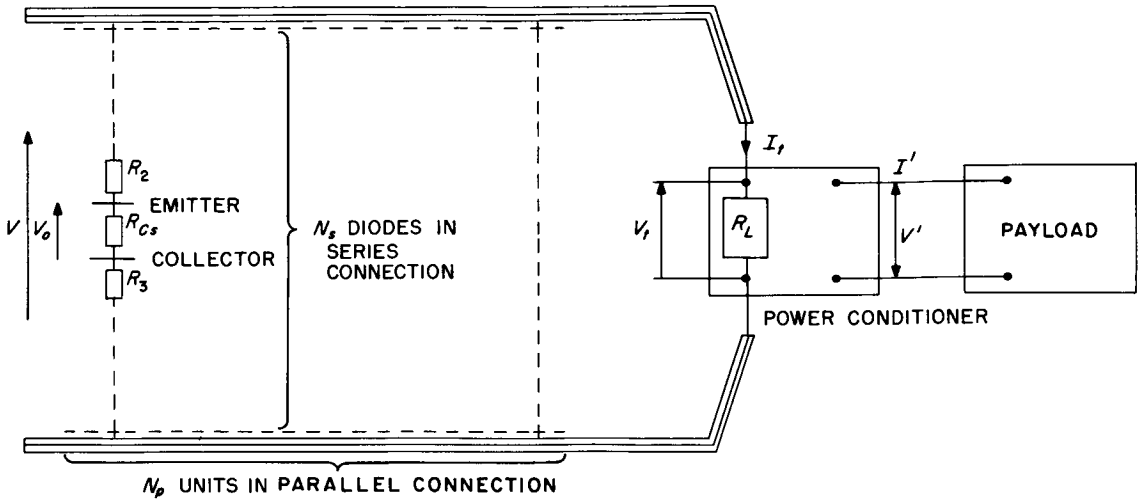


Fig. 29. Electrical connections for in-pile thermionic system

where V_o' is the open circuit voltage, I_o' is the full load current, and K the fractional regulation. K is considered as a function of the input

$$K = K_o - \mu \overline{dV/dI} \tag{20}$$

where K_o is a limiting output regulation and μ a scale factor. The average derivative is determined from the I-V characteristic.

5. Cesium Circuit

The cesium temperature (pressure) is regulated by an outside circuit shown schematically in Fig. 30. Employing the notation indicated on Fig. 30 and the equations of state, continuity, and motion, the following expressions are obtained (Ref. 23):

a. The thermionic cell and pipe

$$\frac{d^2}{dt^2} \ln p + \frac{k}{A\rho_p} \frac{d}{dt} \ln p + \frac{A}{\rho_p V_{cs} l} p = \frac{A}{\rho_p V_{cs} l} p_w + \frac{d^2}{dt^2} \ln T_p + \frac{K}{A\rho_p} \frac{d}{dt} \ln T_p \tag{21}$$

b. Continuity for equal mass flow at the two ends of the pipe

$$-\frac{V_{cs}}{RT_p} \frac{dp}{dt} + \frac{V_{cs} p}{RT_p^2} \frac{dT_p}{dt} = \frac{V_w}{RT_{cs}} \frac{dp_w}{dt} + \lambda (p_w - p_{cs}) \tag{22}$$

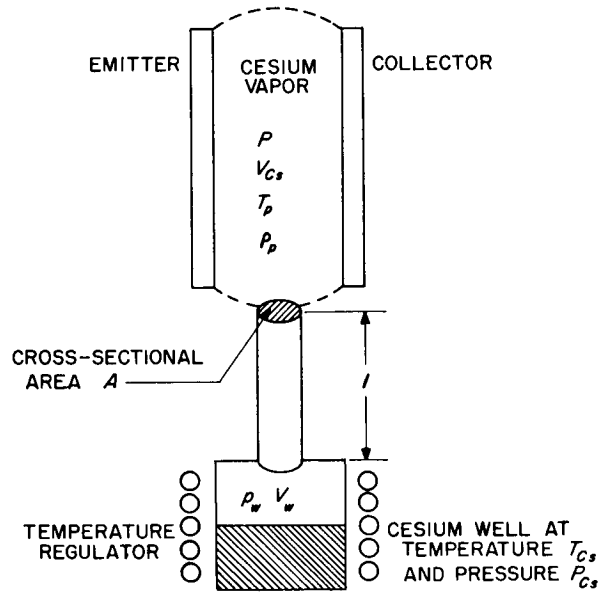


Fig. 30. Schematic of cesium temperature regulation circuit — adapted from General Atomic Report GA-3540 (Ref. 23)

The assumptions leading to Eqs. (21) and (22) are:

(1) Velocity and density within pipe are constant

(2) Constant well temperature

(3) Condensation rate assumed proportional to $P_w - P_{Cs}$

(4) Pipe pressure drop corresponds to laminar flow.

Attempts have recently been made to build thermionic diodes charged with the proper amount of cesium, thus eliminating the need for a separate circuit. If successfully developed, this will considerably simplify the system.

6. Coolant Loops

About 90% of the heat generated by the reactor must be dumped. For this analog study a representative heat dump system was designed. Each region is characterized by average temperatures. Fig. 31 shows a schematic of the system and, together with Fig. 28, shows the employed notation for the following equations:

$$- V_3 \rho_3 c_{p_3} dT_3/dt + \pi D_3 L_{12} q_2'' + L_{12} k_{23} \gamma_3 (T_2 - T_3) - IV_o + I^2 R_3/3 = \pi D_4 L_{12} (t_3/k_3 + t_4/2k_4)^{-1} (T_3 - T_4) \quad (23)$$

$$- V_4 \rho_4 c_{p_3} dT_4/dt + \pi D_4 L_{12} (t_3/k_3 + t_4/2k_4)^{-1} (T_2 - T_3) + L_{12} k_{23} (1 - \gamma_3) (T_2 - T_3) = 2\pi D_5 L_{12} (T_4/k_4 + t_5/k_5)^{-1} (T_4 - T_5) \quad (24)$$

$$- V_5 \rho_5 c_{p_5} dT_5/dt + 2\pi D_5 L_{12} (t_4/k_4 + t_5/k_5)^{-1} (T_4 - T_5) = h_{56} \pi D_6 L_{12} (T_5 - T_6) \quad (25)$$

$$N_s V_6 \rho_6 c_{p_6} dT_6/dt + w_6 c_{p_6} (T_{6o} - T_{6i}) = h_{56} \pi D_6 L_{12} (T_5 - T_6) \cdot N_s \quad (26)$$

$$T_6 = \tfrac{1}{2}(T_{6i} + T_{6o}) = T_H \quad (27)$$

a. Reactor to heat exchanger

$$T_{6o} = T_{Hi} (t + \tau_1)$$
$$T_{6i} = T_{Ho} (t - \tau_2) \quad (28)$$

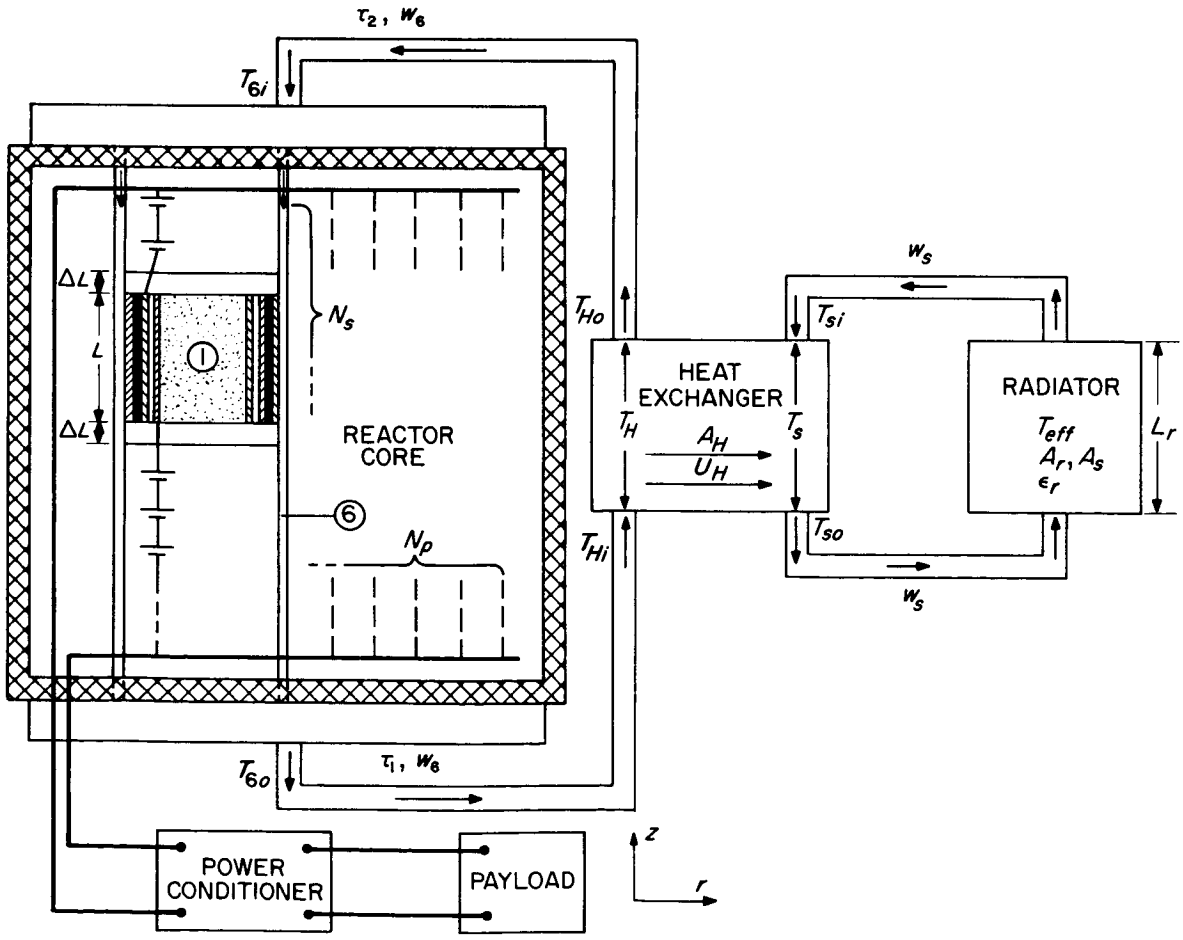


Fig. 31. Coolant loops for in-pile thermionic space reactor

$$N_p w_6 c_{p_6} (T_{Ho} - T_{Hi}) = U_H A_H (T_H - T_s) \quad (29)$$

$$T_s = \tfrac{1}{2}(T_{si} + T_{so}) \quad (30)$$

b. Heat exchanger to radiator

$$M_s c_{p_s} dT_s/dt = U_H A_H (T_H - T_s) - w_s c_{p_s} (T_{so} - T_{si}) \quad (31)$$

$$- L_r A_s \rho_s c_{p_s} dT_{eff}/dt + w_s c_{p_s} (T_{so} - T_{si}) = \varepsilon_r \sigma A_r T_{eff}^4 \quad (32)$$

Compared with the true relationships, Eqs. (23) to (32) are, of course, very simplified, but can be expected to yield a reasonable description of the over-all behavior. The proper determination of the various parameters has required an extensive effort which is continuing.

The foregoing equations can be simulated on the analog facility. An extension to several spatial regions requires a prohibitively large number of amplifiers. Conversion to the digital facility is simplified by making use of a translation program that directly digitalizes the original analog simulation. With the employment of this program, it is anticipated that the spatial dependence can be included with a minimum of coding effort.

References

1. Massier, P. F., "Axisymmetric Steady Flow of a Swirling Compressible Fluid Through a Convergent-Divergent Nozzle Without External Heat Transfer," *SPS 37-33*, Vol. IV, pp. 133–141, Jet Propulsion Laboratory, Pasadena, California, June 30, 1965.

2. Massier, P. F., "Swirling Flow of Argon Through an Axisymmetric Convergent-Divergent Nozzle," *SPS 37-34*, Vol. IV, pp. 149–157, Jet Propulsion Laboratory, Pasadena, California, August 31, 1965.

3. Back, L. H., Massier, P. F., and Gier, H. L., "Comparisons of Experimental With Predicted Wall Static-Pressure Distributions in Conical Supersonic Nozzles," Technical Report No. 32-654, Jet Propulsion Laboratory, Pasadena, California, October 15, 1964 (Also appears in *AIAA Journal*, Vol. 3, No. 9, September 1965, pp. 1606–1614).

4. Elliott, D., Cerini, D., Hays, L., Weinberg, E., and O'Conner, D., "Liquid MHD Power Conversion," *SPS 37-34*, Vol. IV, pp. 163–173, Jet Propulsion Laboratory, Pasadena, California, August 31, 1965.

5. Elliott, D., Cerini, D., Eddington, R., Hays, L., and Weinberg, E., "Liquid MHD Power Conversion," *SPS 37-33*, Vol. IV, pp. 142–150, Jet Propulsion Laboratory, Pasadena, California, June 30, 1965.

6. Kerrisk, D. J., and Masek, T. D., "A Zero-Gravity Feed System for Mercury Ion Engines," *SPS 37-33*, Vol. IV, pp. 150–159, Jet Propulsion Laboratory, Pasadena, California, June 30, 1965.

7. Porter, R. N., Stanford, H. B., Tyler, W. H., Kocsis, J. D., and Keller, O. F., "Advanced Liquid Propulsion Systems," *SPS 37-31*, Vol. IV, pp. 192–201, Jet Propulsion Laboratory, Pasadena, California, February 28, 1965.

References (Cont'd)

8. Hurwitz, H., Sutton, G. W., and Tamor, S., "Electron Heating in Magnetohydrodynamic Power Generators," *ARS Journal*, Vol. 32, August 1962, p. 1237.

9. Russell, G. R., Byron, S., and Bortz, P. I., "Performance and Analysis of a Non-Equilibrium Crossed Field Accelerator," AIAA Electric Propulsion Conference, Colorado Springs, Colo., March 1963, AIAA Paper No. 63-005.

10. Kelly, A. J., Nerheim, N. M., and Gardner, J. A., "Electron Density and Temperature Measurements in the Exhaust of a MPD Source," AIAA Second Annual Meeting, San Francisco, Calif., July 1965, AIAA Paper No. 65-398.

11. Tricomi, F. G., *Integral Equations*, Interscience Publishers, Inc., New York, 1957.

12. Pearce, W. J., "Calculation of the Radial Distribution of Photon Emitters in Symmetric Sources," *Conference on Extremely High Temperatures, Boston, Mass., March 1958*, H. Fisher and L. C. Mansur, Editors, John Wiley & Sons, Inc., New York, 1958, p. 123.

13. Nestor, O. H., and Olsen, H. N., "Numerical Methods for Reducing Line and Surface Probe Data," *SIAM Review*, Vol. 2, July 1960, p. 200.

14. Freeman, M. P., and Katz, S., "Determination of the Radial Distribution of Brightness in a Cylindrical Luminous Medium with Self-Absorption," *Journal of the Optical Society of America*, Vol. 50, August 1960, p. 826.

15. Bockasten, K., "Transformation of Observed Radiance into Radial Distributions of the Emission of a Plasma," *Journal of the Optical Society of America*, Vol. 51, September 1961, p. 943.

16. Dieke, G. H., "Spectroscopy of Combustion," *Physical Measurements in Gas Dynamics and Combustion*, R. W. Ladenburg, B. Lewis, et al., Editors, Princeton University Press, Princeton, N. J., 1954, p. 491.

17. Griem, H. R., *Plasma Spectroscopy*, McGraw-Hill Book Company, Inc., New York, 1964.

18. Olsen, H. N., "The Electric Arc as a Light Source for Quantitative Spectroscopy," *Journal of Quantitative Spectroscopy and Radiative Transfer*, Vol. 3, 1963, p. 305.

19. Hildebrand, F. B., *Introduction to Numerical Analysis*, McGraw-Hill Book Company, Inc., New York, 1956.

20. Kerrebrock, J. L., "Nonequilibrium Ionization Due to Electron Heating: I. Theory," *AIAA Journal*, Vol. 2, June 1964, p. 1072.

21. Byron, S., Stabler, R. C., and Bortz, P. I., "Electron-Ion Recombination By Collisional and Radiative Processes," *Physical Review Letters*, Vol. 8, May 1962, p. 376.

22. McWhirter, R. W. P., and Hearn, A. G., "A Calculation of the Instantaneous Population Densities of the Excited Levels of Hydrogen-Like Ions in a Plasma," *Proceedings of the Physical Society*, Vol. 82, 1963, p. 641.

23. Wright, W. B., Jr., "Parametric Study of Direct Conversion Reactor Power Systems," Final Technical Report, Report GA-3540, General Dynamics Corporation, General Atomic Division, San Diego, Calif., October 31, 1962.

Reference (Cont'd)

24. Hill, P. R., and Giertzen, A. J., "Analytical Study of Power Matching in Thermionic Reactor Cores," *Transactions of the American Nuclear Society,* Vol. 7, No. 2, November 1964, pp. 378–379.

25. Cohn, C. E., "Reflected Reactor Kinetics," *Nuclear Science and Engineering,* Vol. 13, 1962, pp. 12–17.

26. Yasinsky, J. B., and Henry, A. F., "Some Numerical Experiments Concerning Space-Time Reactor Kinetics Behavior," *Nuclear Science and Engineering,* Vol. 22, 1965, pp. 171–181.

27. Weissman, I., "Improved Thermionic Emitter: Thorium on Grain-Oriented Polycrystalline Tungsten," *Journal of Applied Physics,* Vol. 36, No. 2, February 1965, pp. 406–411.

SPACE SCIENCES DIVISION

XV. Lunar and Planetary Sciences

A. A Pattern Range for Small-Aperture Microwave Antennas

M. L. Kellner

1. Introduction

The need for a pattern range for spacecraft microwave instruments of small aperture became evident during the *Mariner II* microwave radiometer program. Frequency ranges to be considered were 8 to 12, 12 to 18, 18 to 26, and 26 to 40 Gc. After investigation, suitable space was located at the Mesa area of JPL. An area providing a 150-ft base line between the antenna pedestal and the exciter was chosen. Between these two points, there is an existing deep gully, which is a desirable condition for an antenna range. The equipment and procedures used and the data obtained will be described here.

2. Implementation

Two 10- × 10-ft concrete pads, 150 ft apart, were constructed on the Mesa. Between these pads is the existing deep gully. This is the current arrangement for "far field" operation, and this range could be used in the future with other radiometers with similar base-line requirements.

An over-all view of the positioner is shown in Fig. 1. Fig. 2 is a view looking from the window of the transmitter building. In the background the radiometer, mounted on the positioner, can be distinguished. Behind the positioner is a panel covered with microwave absorber, which is used to cut down the effects of reflection from other nearby equipment. Located on the stand that supports the transmitting parabola is a telescope with which visual collimation of the output beam can be accomplished. A precision mirror is fixed to an adjustable mirror mount located on the boresighting stand. This adjustable mirror mount assembly can be used to accurately calibrate the horizontal and vertical planes of the output beam to within 0.1 deg.

The equipment inside the transmitter building is shown in Fig. 3. The lefthand side shows the boresighting bench mounted in position on its stable stand. Each boresighting bench consists of a complete transmitter system containing a signal source and all microwave components for precision antenna patterns. These boresighting benches are broadband systems and cover the entire band of the

following frequency ranges: 8.2 to 12.4, 12.4 to 18.0, 18.0 to 26.5, and 26.5 to 40.0 Gc. The center portion of Fig. 3 shows the equipment rack for the antenna patterns recorder and positioner control unit. The bottom portion of the rack contains the voltage regulator for the entire system. To the right of the equipment rack is the power supply for the backward wave oscillators. The wide-range microwave receiver is located on top of the power supply. The combination of the precision microwave receiver and the antenna pattern recorder provides the antenna range with the capability of a linear dynamic range of 60 db. The various-frequency boresighting benches are stored on the rack at the right-hand side of Fig. 3.

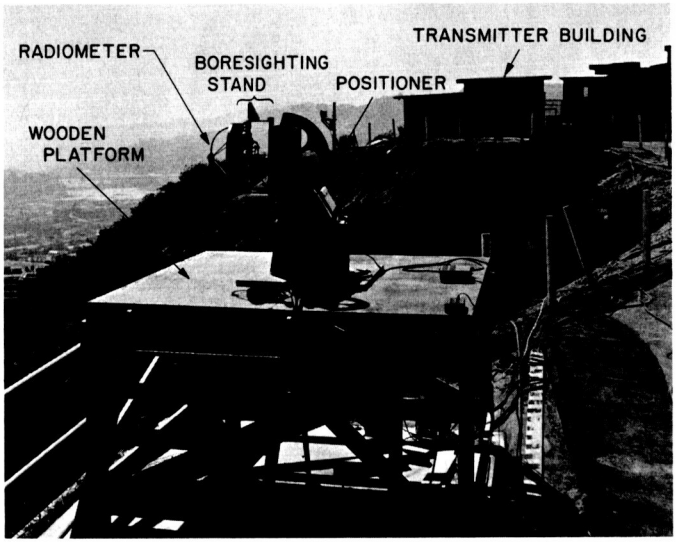


Fig. 1. Over-all view of positioner assembly

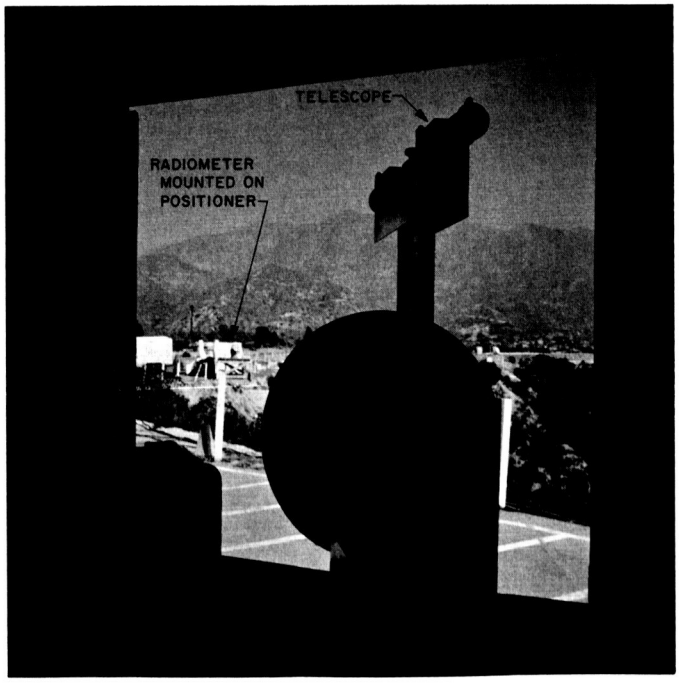


Fig. 2. View of positioner assembly from transmitter building

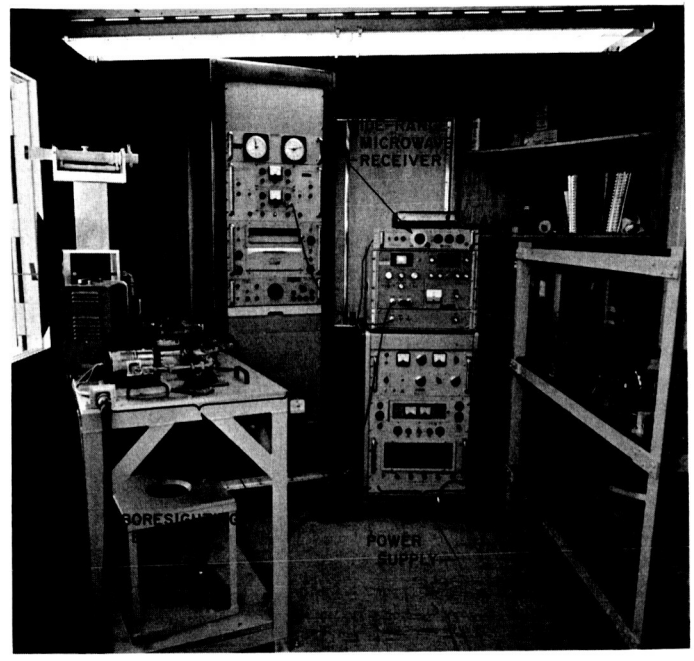


Fig. 3. Equipment inside transmitter building

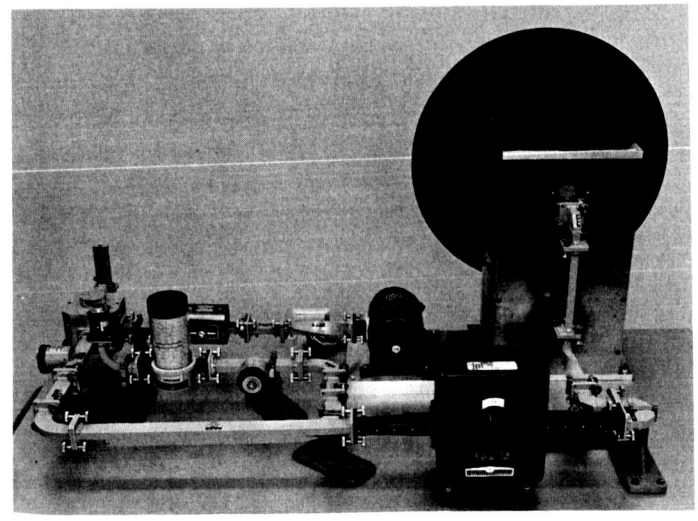


Fig. 4. X-band boresighting bench

Fig. 4 shows an X-band boresighting bench. In this unit, a backward-wave oscillator with a maximum output of approximately 100 mw is used as an RF signal source. It is followed by an isolator, a level-setting attenuator, a waveguide switch that permits the power to be routed into a load when desired, a directional coupler that feeds a precision wavemeter and temperature-compensating thermistor for frequency and power monitoring, two attenuators, and a transmitting parabola. One of the two attenuators before the parabola is a precision attenuator to permit accurate output power level changes for linearity calibration purposes. A rotating step-twist is incorporated into the transmitting parabola mounting structure to allow rapid polarization changes of the transmitted signal. A telescope is also mounted on the parabola mounting structure for visual collimation of the output beam. A block diagram of this system may be seen in Fig. 5.

All of the components are mounted on a ½-in.-thick aluminum plate for rigidity. This plate is mounted to an extremely rigid table, which is optically positioned with regard to the center of the instruments to be tested and is securely anchored to the concrete floor. By using this arrangement, very accurate pattern measurements are possible.

The wide-range microwave receiver is the instrument which provides the range with a 60-db linear dynamic range capability. The frequency range of this unit is 1.95 to 100 Gc. One coaxial cable from the antenna to the receiver eliminates lossy waveguides and rotary joints. A crystal mount is located at the terminal of the receiving antenna, where fundamental or harmonic mixing takes place. The local oscillator is tuned such that its fundamental or appropriate harmonic frequency produces a 65-Mc heterodyne frequency with the received signal. This intermediate-frequency (IF) signal is coupled back through the same coaxial cable and frequency-selective tee to a 65-Mc IF amplifier.

The 60-db dynamic range is achieved by a unique feedback arrangement utilizing a shaped potentiometer coupled to the logarithmic potentiometer of the rectangular recorder. The logarithmic potentiometer is used to modify

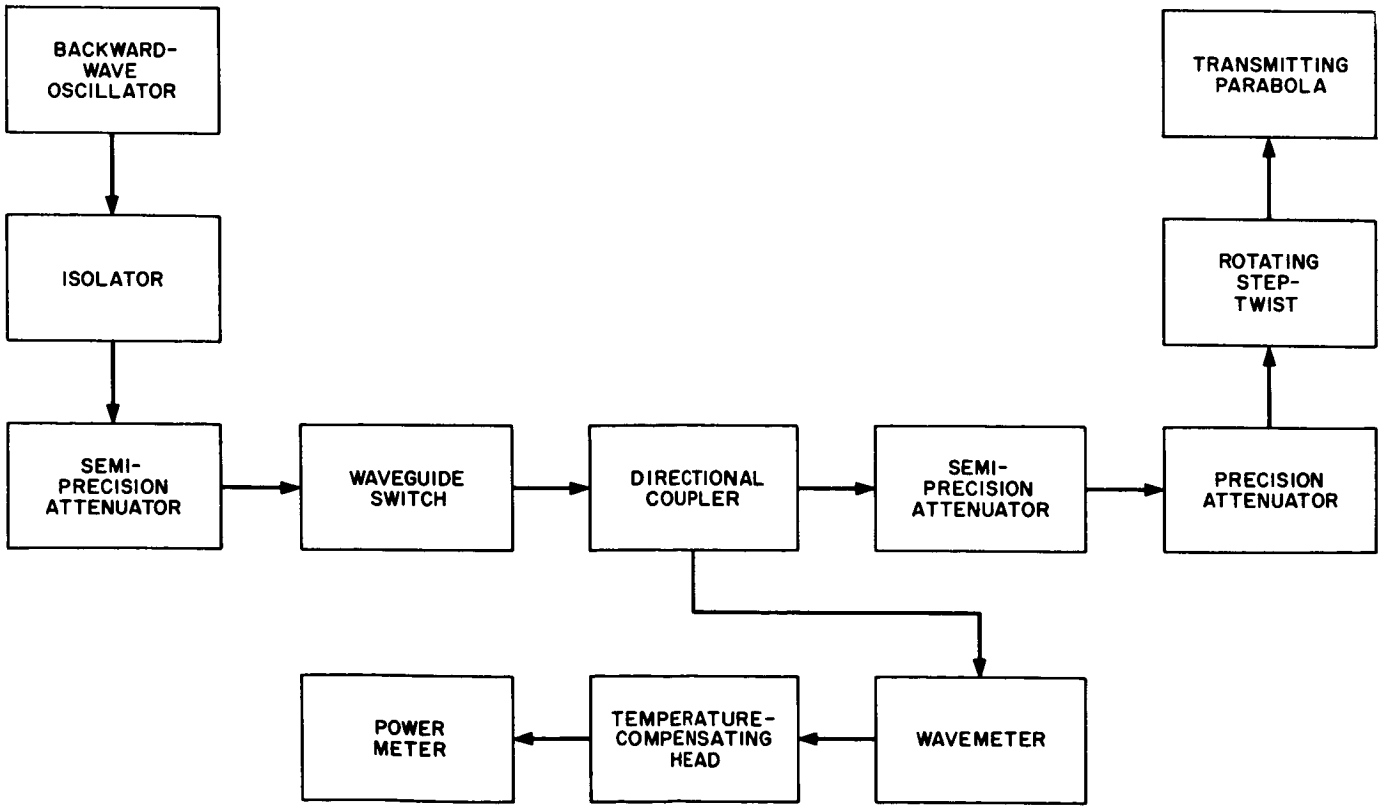


Fig. 5. Block diagram of X-band boresighting bench

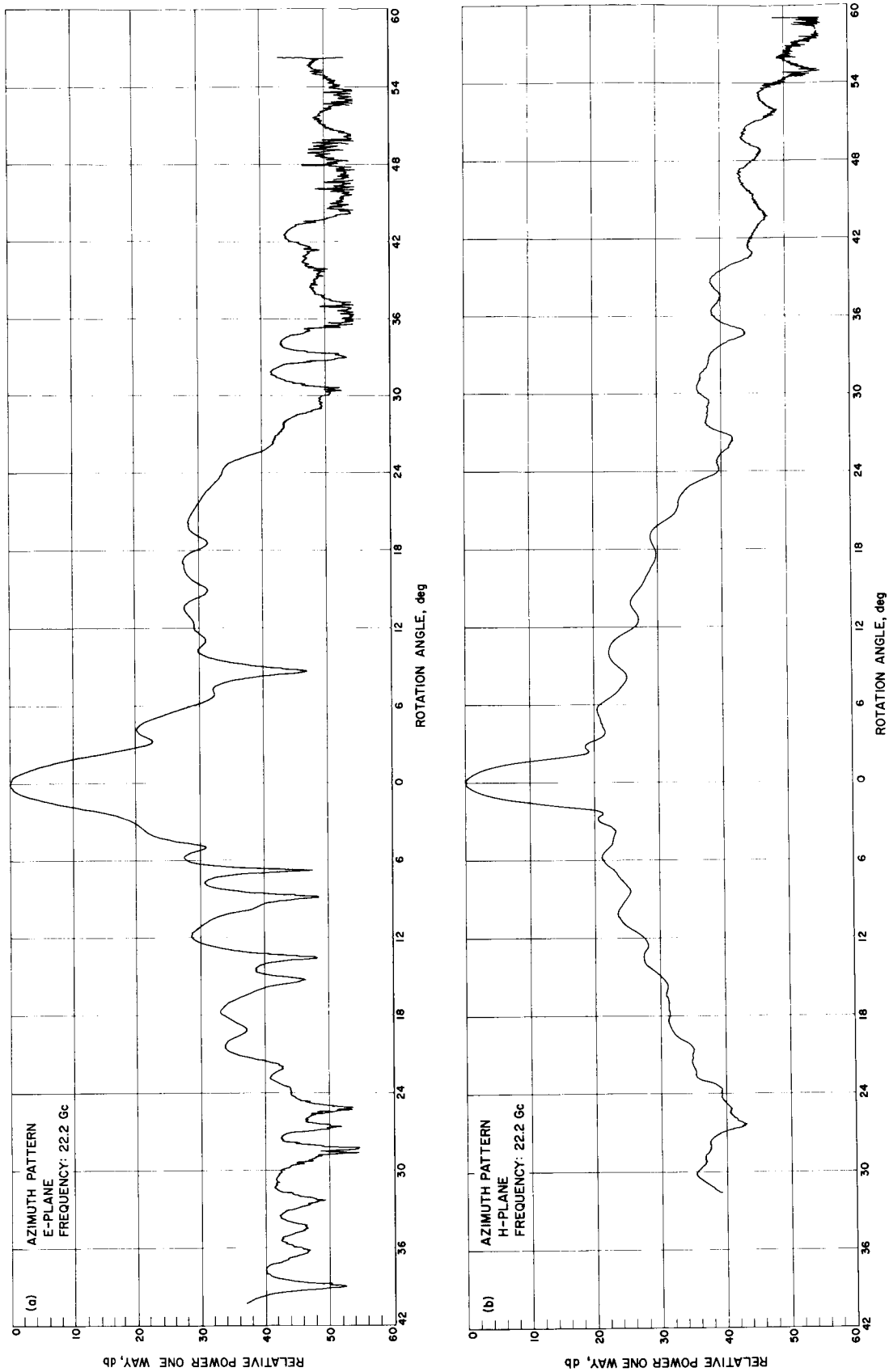


Fig. 6. Patterns obtained at 22.2 Gc

the gain of the receiver IF amplifier to expand the recorder scale from 40 to 60 db. The system is extremely stable and is linear to within ±0.25 db over the entire range.

3. Typical Procedure and Results

The first step in the *Mariner II* radiometer antenna measurements was to properly focus the diplexer on the antenna. Since this is a two-channel system consisting of 15.8 and 22.2 Gc, all diplexer focusing was done at 22.2 Gc. This was accomplished by observing the on-axis diplexer position yielding the maximum output consistent with lowest sidelobes and symmetry when the antenna was pointed at the transmitting parabola.

The patterns obtained at 22.2 Gc are shown in Fig. 6. The main characteristics at this frequency are summarized below:

Scan plane	3-db beam-width, deg	6-db beam-width, deg	First side-lobe, db
Azimuth E-plane	2.0	2.0	−20.0
Azimuth H-plane	2.0	2.6	−19.0

B. Infrared Spectrometry of Mars

H. T. Bull

1. Introduction

The spectrometry of Mars between 3–4 μ has been sought after for a number of years; e.g., Ref. 1, 2, and 3. Very low resolution spectra have been obtained, but the low resolution of 60 to 90 cm^{-1} severely limits their value. It was therefore decided to attempt to obtain about an order of magnitude improvement in resolution by using a Michelson interferometer. A working breadboard instrument had been made and used at JPL (*SPS 37-23*, Vol. IV, pp. 254–258). The only telescope which was available around the time of the Mars opposition was the solar telescope at Kitt Peak National Observatory. There, a generous amount of observing and preparation time was made available. However, because of the relatively small collecting area the scan times were made very long, greater than two hours, when trying for the 5 cm^{-1} resolution. This aggravated the problems with manual guiding and cooling the PbS detectors; therefore, rather lower resolution than had been hoped for was obtained in this attempt.

2. Preliminary Results

Spatial division of the planet could not be obtained because of the low signal/noise ratio and poor seeing. The Moon and Sun were used as a reference and some variations due to H_2O, HDO, and different air masses are evident (Fig. 7). However, the solar spectrum (Fig. 8) was sufficiently reproducible so that large departures should be detectable, if present in the Martian spectrum. A much more elaborate calibration is obviously desirable, but circumstances at this time did not permit this. There are few celestial objects which are bright enough to serve as calibration sources in the region between 3–4 μ. The noise in the Martian spectra limited the resolution to a minimum of 12.5 cm^{-1}. Fig. 9 shows the average of the four Mars spectra. RMS deviations have been computed, based on just the four individual spectra; the statistics are therefore not good. Arrowhead-to-arrowhead on the graphs indicate 2σ. Under the usual interpretation of probable errors, this would be too inaccurate to definitely establish departures from the solar spectrum. However, when the Martian and solar spectra are inspected for correspondence, close agreement is evident between 3000 and 3200 cm^{-1}. There are comparatively large departures around 2850 and 2950 cm^{-1}, and these could possibly be real. The photometric accuracy is not sufficient to definitely establish this, but these results do suggest there may be something new in this region.

3. Future Work

One of the most interesting things to look for is evidence of C–H bonds. This reveals itself in terrestrial plants in a very striking way (see Fig. 10, or for more complete data, *SPS 37-31*, Vol. IV, pp. 261–272). It is therefore proposed to obtain reference spectra of our terrain and also resume developing this technique in conjunction with observing Mars at a larger telescope. It should be possible to determine whether or not any large departures exist between the solar and Martian spectra before the next opposition of Mars in 1967. Also obtaining the spectra at widely-spaced times may well help in any interpretation of features suspected of seasonal variations.

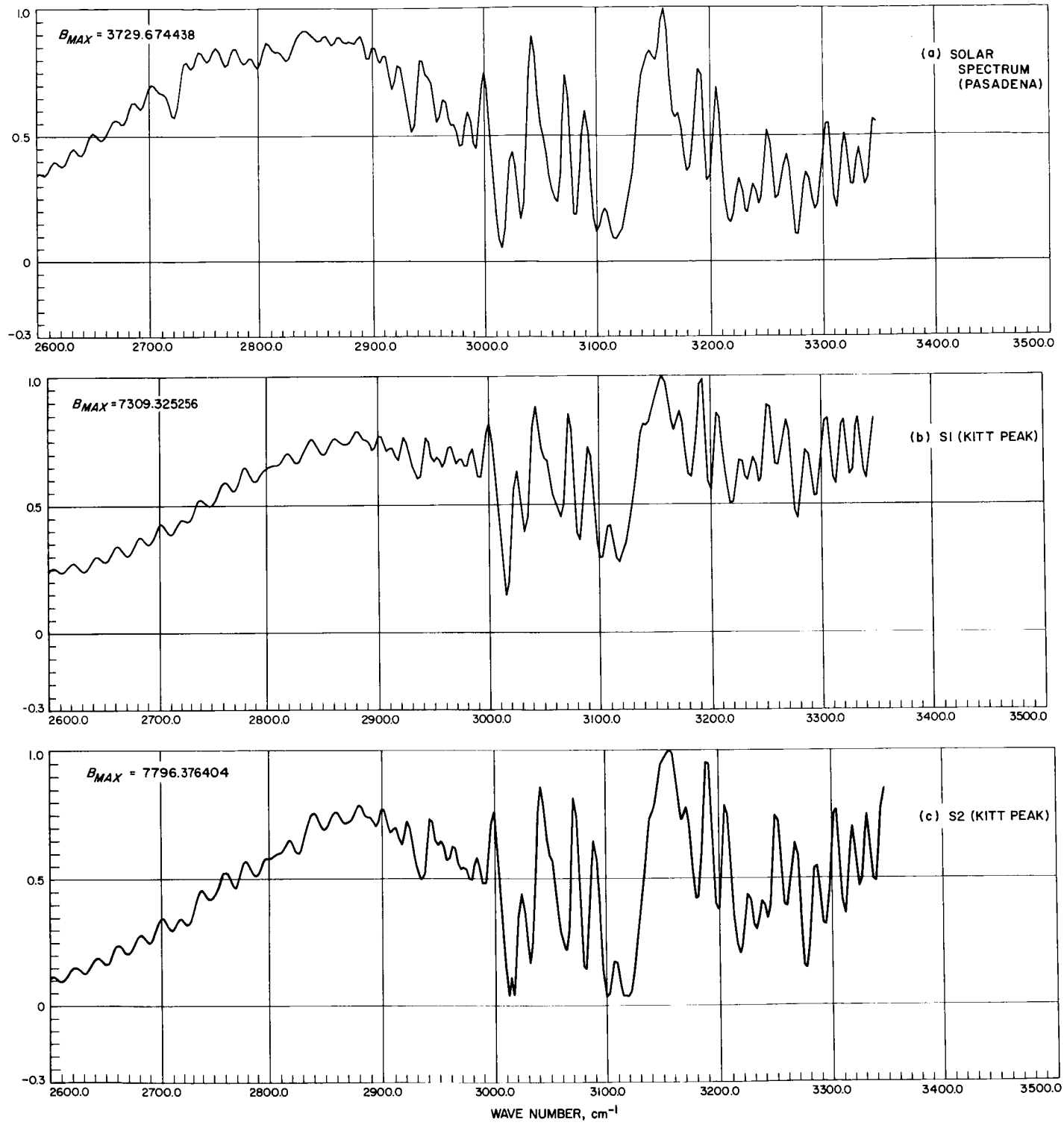


Fig. 7. Solar spectra (5 cm^{-1})

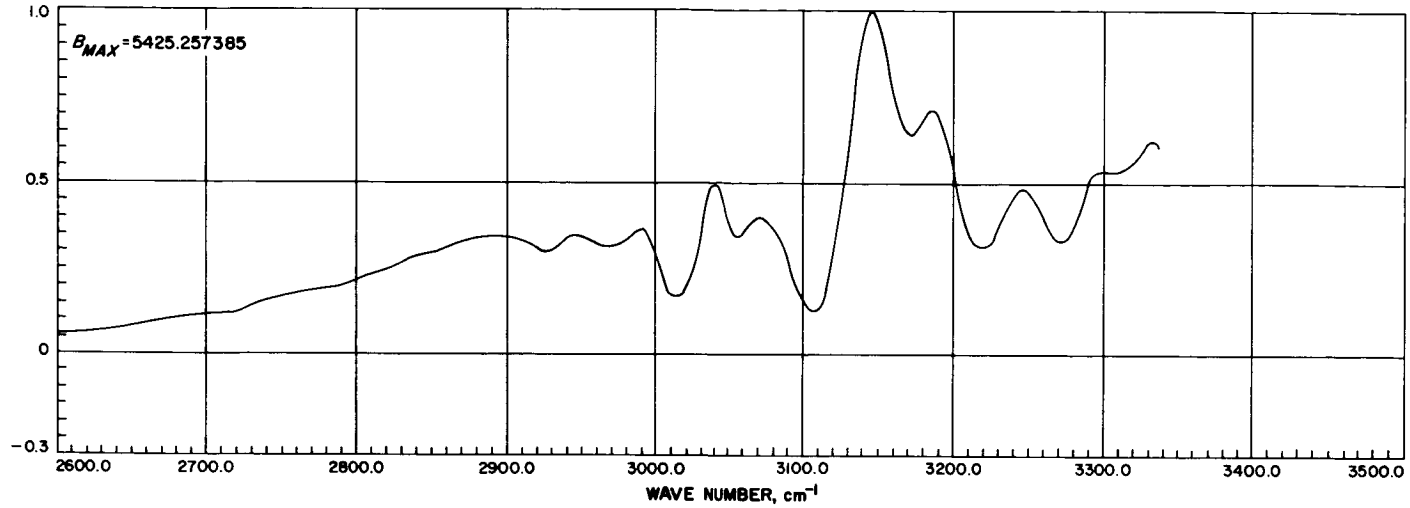


Fig. 8. Solar spectrum (12.5 cm^{-1})

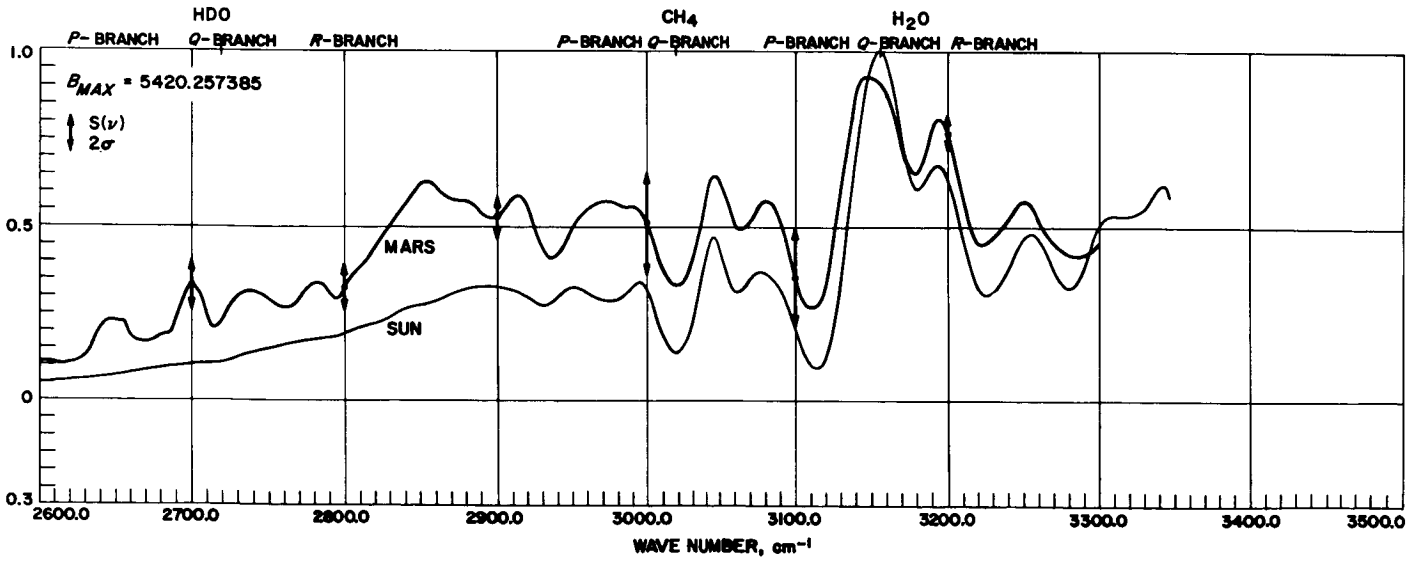


Fig. 9. Mars spectrum vs solar spectrum (12.5 cm^{-1})

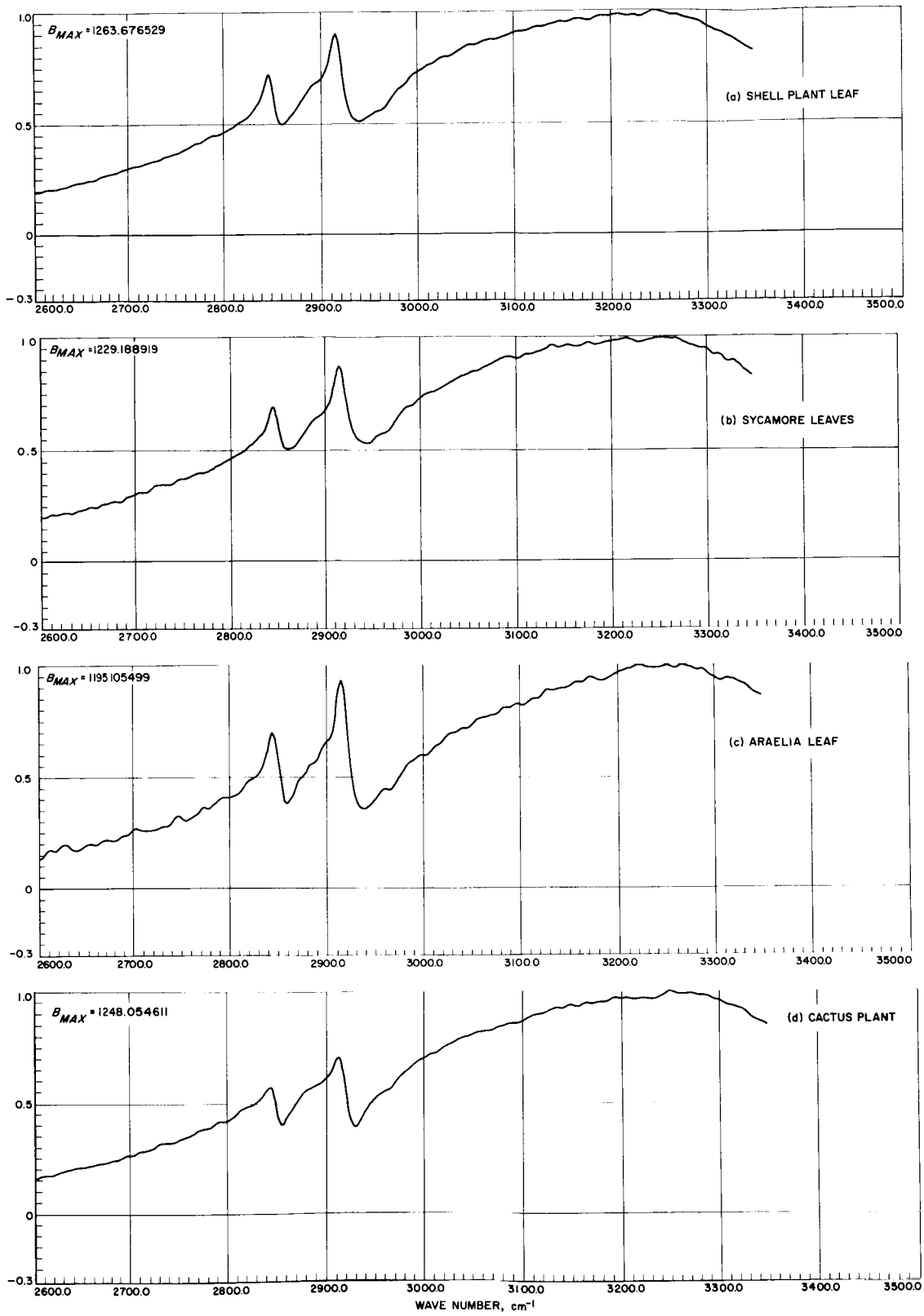


Fig. 10. Laboratory reflection spectra of vegetation, optically smooth specimens

C. Scanning Mode Options for a Lunar Diffraction Experiment

J. A. Dunne

1. Introduction

This paper addresses the problem of the total power-time product of a lunar X-ray diffraction experiment from the point of view of goniometer scanning mode. It is assumed that a scan range encompassing the major peaks of the common rock-forming and accessory minerals is a first priority requirement. Therefore, scanning mode options must operate over a minimum range of 4–45 deg 2θ (using CuK_α radiation). The objective is the reduction of the power-time profile by an order of magnitude. The present instrument requires approximately 100 w-hr for a standard-sample–unknown-sample pair.

2. Scan Modes

The present lunar diffractometer employs a continuous scan of 0.5 deg 2θ per min. The instrument readout is in the form of direct or binary divided counts transmitted in real time. Thus, the question of the optimum ratemeter time constant for graphical presentation of the diffraction data is not pertinent to this discussion. Power-time (PT) profile product reduction can take one of two forms in this type of scan mode. One can either run the goniometer at a higher speed or reduce X-ray tube (XRT) power input. The net result of either of these techniques is the loss of statistical precision in the measurement of X-ray intensity. For the small angular interval $d\theta$, the precision of X-ray intensity measurement is given by

$$\sigma_{d\theta} = \left(\frac{Rd\theta}{\theta'}\right)^{1/2}$$

where R is the counting rate and θ' the scan speed.

The choice as to constant power or constant time reduction of power-time product will be discussed in a succeeding section. It is clear, however, that reduction in PT using the continuous scan mode is accompanied by a loss in peak intensity precision, which will decrease in proportion to the square root of the PT gain.

A second scan mode, which has been used for many years in achieving high precision at low diffraction intensities is the step scan mode, wherein the goniometer is moved in fixed increments and the X-ray intensity at each position is integrated for a given time (fixed time) or until a preset number of counts are accumulated (fixed count). This technique has been proposed by D. Spreng of JPL for the Mark II lunar diffractometer in order to meet bandwidth constraints imposed by the spacecraft telemetry system. The proposed stepping increment of 0.01 deg 2θ and count-integration time of 1 sec plus 250 msec for readout and advance result in a scan speed roughly equivalent to the present 0.5 deg/min. Since diffraction peak widths obtained using the lunar instrument are many times the proposed 0.01-deg increment size (0.2 deg 2θ at half height), it may be possible to obtain significant PT savings by increasing the increment, with little loss in resolution and peak intensity. Of course, the accuracy to which peak angular position can be measured will be no better than the size of the selected increment. Fig. 11 shows the "five peaks" pattern of quartz obtained on a standard diffractometer with increments of 0.05 and 0.10 deg 2θ. A continuous scan of the same pattern is given for comparison. Increments of 0.05 and 0.10 deg 2θ represent PT reductions by factors of 6 and 12, respectively, compared to the continuous mode, which scans at a rate of 0.5 deg/min. Fig. 12 presents a similar comparison for very weak diffraction peaks. A standard Norelco diffractometer with 1-deg divergence and 0.006-in. receiving slits was used in collecting the data for Figs. 11 and 12.

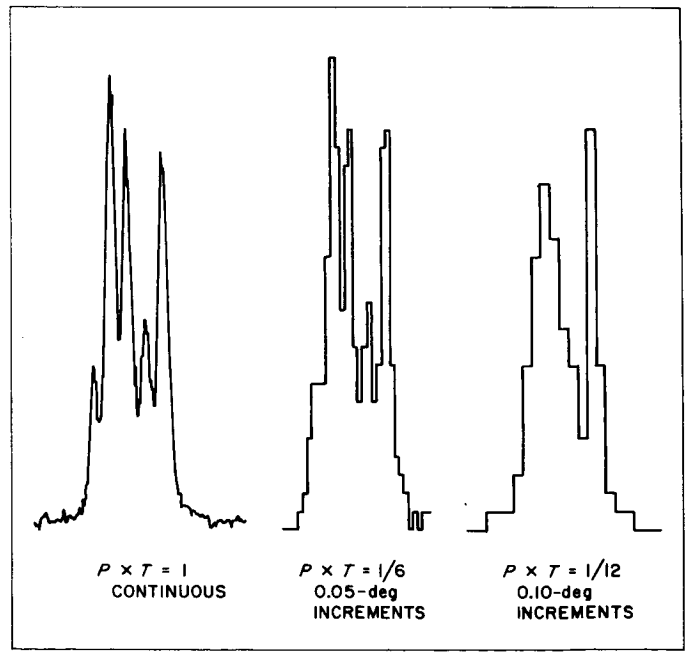


Fig. 11. Quartz resolution standard scans, using continuous and step-scan modes

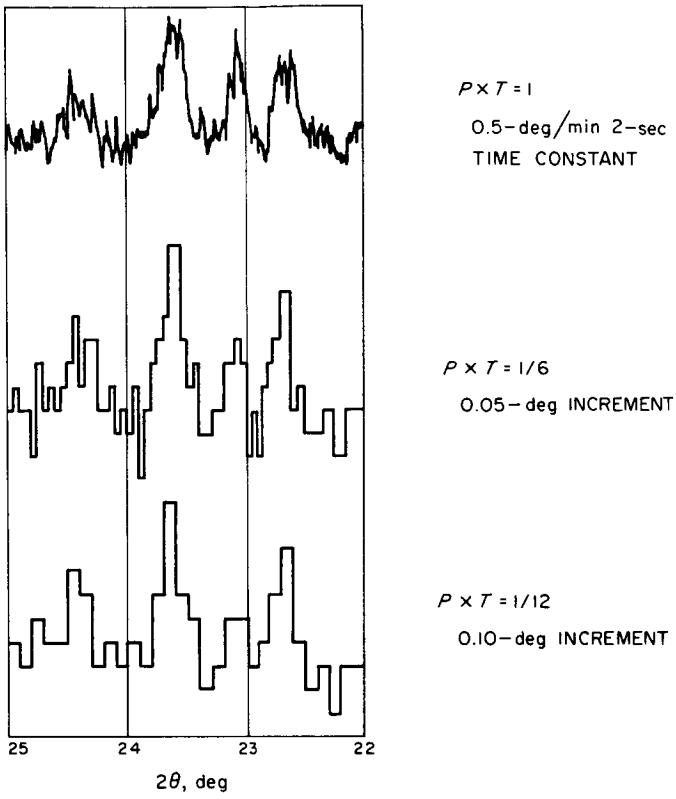


Fig. 12. Scans of relatively weak diffraction lines, using continuous and step-scan modes

Further savings in power or time can be accomplished using the step scan mode through the application of another technique proposed by Spreng. This consists of skipping an increment when the count on the measured increment is less than some predetermined "background" level. Fig. 13 shows the pattern obtained using this technique compared to the diagram obtained from the same specimen without any rejection. In this example, a PT saving of 37% is achieved. The problem in the application of this method is that it requires an accurate foreknowledge of background levels. An alternative procedure which might prove more effective in practice would be to program the goniometer to step in larger increments in 2θ regions known to be sparsely populated with peaks in most rock diffraction diagrams. For example, using a scan range of 4–45 deg 2θ with CuK_α radiation, one could elect to scan the first 16 and the last 5 deg in 0.10-deg increments, using the 0.05-deg increments in the heavily populated 20–40 deg 2θ region, with a PT saving of 46%, which compares favorably to the maximum achievable saving of 50% for the "skip alternate" technique outlined above.

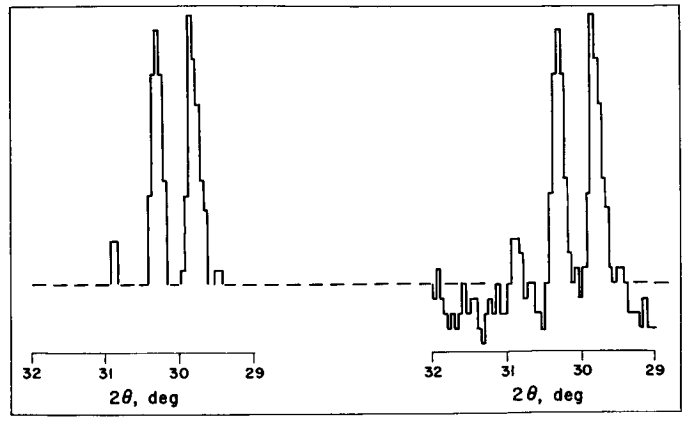


Fig. 13. Step scan of a portion of a rock-diffraction diagram, showing background rejection by the "skip-alternate" technique

One important difference between the step scan and continuous mode approaches to PT reductions is that the former, by accumulating counts at fixed angular increments, provides peak intensity precision in the raw data comparable to that obtained in a much higher PT continuous scan. By selecting the increment employed with due regard to the instrument line breadth, a significant reduction in PT can be accomplished with very little loss in peak intensity and line resolution. It can be shown that, for a diffraction-line half-width of breadth of 0.2 deg 2θ, a 0.1 deg 2θ increment step scan will integrate at least once at a point within approximately 75% of the maximum peak intensity. It should be noted, however, that the total number of counts observed in traversing a diffraction peak (integrated intensity) in either the continuous or step scan modes will be approximately the same at any given PT, and peak intensity precision can be recovered from the continuous scan mode data by summing over a time interval equivalent to an appropriate step scan interval.

The differences between the two scan modes, then, reduce essentially to engineering considerations such as mechanical execution, electronic complexity, required telemetry bandwidth, and programming flexibility. The step scan method, for example, requires the use of a digital register (scaler) for the accumulation and storage of the X-ray counts observed at each angular increment, while the continuous mode necessitates real-time transmission at reasonably high pulse rates. Absolute angular indication using a digital shaft encoder is much easier to implement using the step scan method. This feature, although adding to the complexity of the instrument, offers the advantages of accurate knowledge of angular

position without "open loop" error. In addition, the step scan mode, with angular indication and appropriate comparator circuits, can be readily programmed to slew through angular regions of the standard-sample pattern, saving a large fraction of the time presently required for standardization. If four peaks at wide angular separation were used for standardization, for example, suitable slew programming could provide a standard run in approximately one-tenth the time required for running an unknown specimen.

Using this standard sample programming concept, a basic stepping increment of 0.05 deg 2θ, and the large increment (0.10 deg 2θ) option for the first 16 and last 4 deg of the 4–45 deg 2θ scan range, the total PT of a standard sample—unknown sample pair—would be approximately 8 w-hr (at 40-w peak power), compared to the 100 w-hr figure required by the present instrument.

3. Experimental Implications

Assuming that significant PT savings can be effected without serious losses in data quality for a lunar diffraction experiment, one is faced with the question of whether power or time should be reduced. It is clear that this decision will be determined principally by mission constraints, particularly in the area of available telemetry bandwidth, transmitter time, and the science data storage capacity (a 4–45-deg diffraction diagram scanned at 0.05 deg 2θ increment with absolute angle indication will require around 2×10^4 bits to be transmitted). There are, however, certain aspects of these alternatives that should be examined carefully. The fixed-power approach offers a number of potential advantages. The chief of these is the lower total PT of an experiment. The X-ray tube target dissipation and related efficiency losses account for approximately 90% of the total power consumed. The remaining 10% would not be reduced in the course of a fivefold XRT-target power saving. Thus, an experiment conducted for ⅕ hr at full XRT power would save approximately 40% more power-time product than one conducted for 1 hr at one-fifth the XRT target dissipation. In the presence of high non-instrument background levels (natural or artificially generated background flux not produced in the XRT), the constant-power approach would provide a signal-to-noise ratio that is higher than that of the constant-time method by a factor equal to the PT gain.

The principal advantage of the constant-time, lower-power approach results from the reduced target loading, which would in principle allow the utilization of a focal spot that is smaller than the present one by a factor equal to the reduction in power. Thus, a geometric scale-down of the entire goniometer would be possible, which would result in significant volume and weight reductions. Efficiency increases may accompany scale-down, since peak intensities vary inversely with goniometer radius if the Soller collimator angular apertures are maintained. It is not clear, however, how much intensity increase can be obtained without degradation in performance due to geometric aberrations resulting from the corresponding increase in effective axial acceptance angles.

Power reductions, if employed, will most probably be affected by beam current decrease. Reductions in XRT accelerating potential, although perhaps attractive from an engineering point of view, would result in serious efficiency (not to mention spectral line-to-background) losses. This effect is shown in Fig. 14, which is a constant power curve for a Cu target XRT taken on both standard commercial (constant potential) and lunar prototype diffractometer systems. As can be seen in this diagram, 25 kv itself represents an efficiency-high voltage tradeoff, and, unless a significant engineering breakpoint is found close

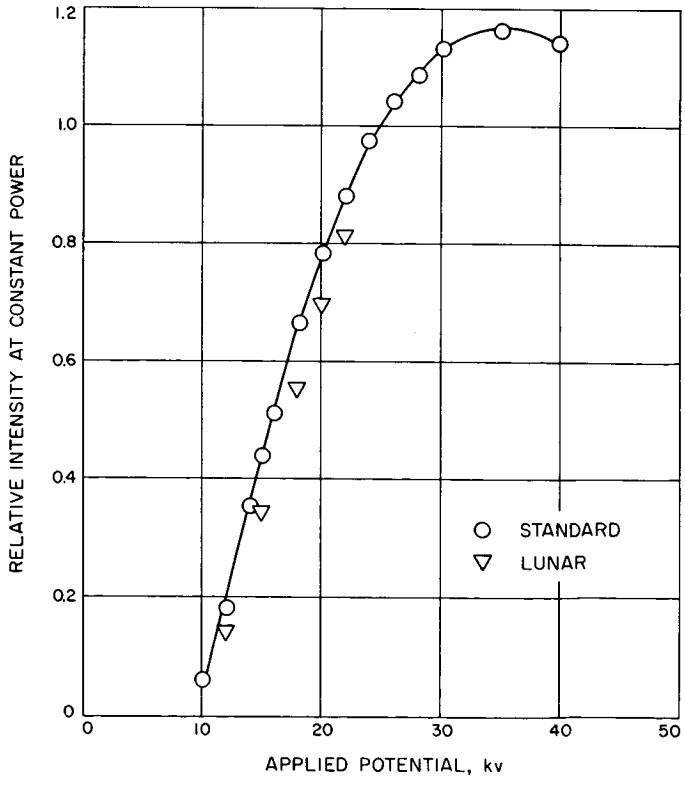


Fig. 14. Relative intensity at constant power vs applied potential

to that voltage, power reduction would be most efficiently achieved by reducing beam current.

4. Areas Under Investigation

The scan modes discussed above will be investigated using a lunar-type XRT in order to delineate performance-PT tradeoffs. Particular attention will be directed to resolution and peak position accuracy degradation as a function of step increment size or scanning speed. The instrument performance will be evaluated in terms of analytical objectives such as phase composition determinations.

D. Crater Densities in Mare Tranquillitatis and Alphonsus in the 1-km to 10-m Diameter Range

J. A. Dunne

Crater counts in the < 1-km-diameter range made on *Ranger VIII* (Mare Tranquillitatis) and *Ranger IX* (Alphonsus) photographs exhibit a slope change between 30–100 m which is interpreted by the writer as the manifestation of the operation of an erosional or lateral redistribution mechanism on the lunar surface. The resulting crater destruction rate closely approaches the rate of crater formation for craters < 100 m in diameter. The size-frequency distribution observed for craters of < 100-m diameter expresses a dynamic balance between formation and destruction rates, and, as such, may be expected to characterize all but the most recent areas of the surface.

1. Data

Crater-size frequency distribution observed in *Ranger VIII* (◇) and *Ranger IX* (○) photographs are plotted in Fig. 15. The data are presented in terms of numbers of craters per square meter versus crater diameter. Each data point represents the number of craters in an interval given by $D + 0.1\,D$ where D is the crater diameter. The vertical lines drawn through each data point are intended to indicate the uncertainty associated with that point. The length of the vertical line is $\pm 2N^{1/2}/A$ where N is the number of craters counted in an area of A square meters. The selection of $\pm 2N^{1/2}$ as the confidence interval is based on the assumption that craters of a given size are distributed randomly over large areas of surface, and that the probability that the number of craters observed in a small sampled area is representative of the crater density in a much larger area is a function of the number N of such craters counted. The probability that the true value lies within the chosen interval is 0.97 for a Gaussian distribution.

The black triangles in Fig. 15 represent the count made on *Ranger VII* (Mare Cognitum) photographs by R. Hartmann and presented by G. Kuiper (Ref. 4), normalized

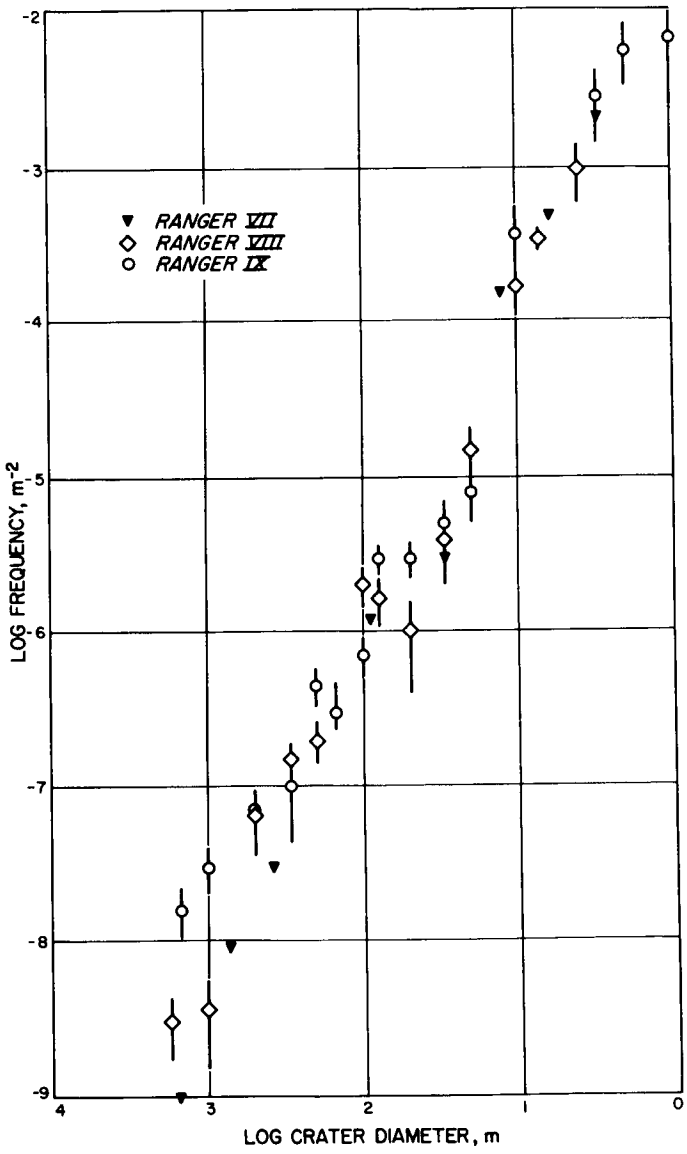


Fig. 15. Crater size-frequency data obtained from *Ranger* photographs

to the same interval ($D \pm 0.1\ D$) as the *Ranger VIII* and *Ranger IX* data.

2. Discussion

Fig. 16 presents curves which summarize the data shown in Fig. 15 for the craters between 1.5 km and 10 m in diameter. The Tranquillitatis and Alphonsus size-frequency distributions are essentially identical below 300 m. Both exhibit marked slope changes at 100 and again at 30 m, with a minimum in the first derivative of the size-frequency curve at 50 m. To put it another way, there are about as many 100-m craters in the sampled areas as there are 50-m craters. The question of interpretation seems to be: does this represent an anomaly in the formation process, or does it imply the loss of $<$ 100-m craters? In the first instance, one could argue that the mass distribution of impacting bodies is such that the $<$ 300-m size-frequency distribution results, or that another, non-impact crater-forming process operates to produce an additional number of 100-m craters. Alternatively, the slope changes could be interpreted as representing the obliteration of $<$ 100-m-diameter craters by the operation of some crater-destruction mechanism.

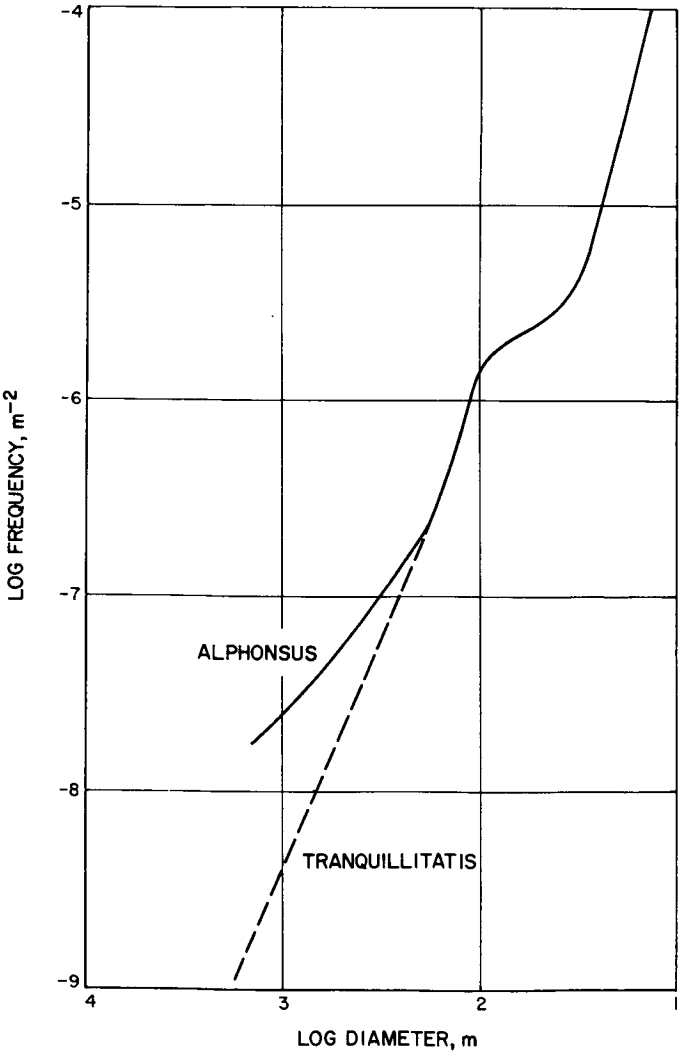


Fig. 16. Crater size-frequency distributions for Alphonsus and Tranquillitatis

One additional bit of observational evidence must be accounted for in the explanation of the size-frequency distribution shown in Fig. 16. This has to do with crater morphology as a function of crater diameter. Examination of the *Ranger VIII* and *IX* photographs reveals that craters of around 100-m diameter exhibit generally more subdued shapes than those of larger ($\sim$ 1-km) and smaller ($\sim$10-m) diameters. That is to say, the proportion of sharply defined, fresh-appearing craters to total craters examined is lower at 100 m than at 1 km and 10 m. One 100-m crater in 20 to 25 appears sharp, while the ratio of 1 km and of 10 m is more like 1 in 5. This "sharpness minimum" at around 100 m was mentioned by O'Keefe and others at the Lunar Surface Conference at the Goddard Space Flight Center, and has been noticed by Hartmann (personal communication) and others at the Lunar and Planetary Laboratory at the University of Arizona. The fact that a morphological effect coincides approximately with the size-frequency slope changes suggests to the writer that these slope changes do not reflect a corresponding anomaly in the mass spectrum of crater-forming bodies. This reduces to two the most probable explanations of the crater size-frequency data. First, the "extra crater" hypothesis can be stated in terms of an independent crater formation process which operates to produce a relatively large number of 100-m size class craters. In this case, the sharpness minimum can be interpreted as representing the initial morphology of the additional craters. The implication of this hypothesis is that the additional craters are formed by a process originating within the Moon; i.e., they are endogenous. The *Ranger IX* photographs provide excellent evidence for postulating the existence of large numbers of endogenous craters on the floor of Alphonsus. In addition to the dark-haloed marrs, a variety of elongate, rimless craters of ½- to 1-km size can be seen coalescing to form long shallow rilles or furrows. The high density of 2- to 300-m rimless craters in and around the dark halos of the rille marrs has been pointed out by Shoemaker (Ref. 5) and others at the Goddard Conference. These craters, which are clearly associated with the internally-generated rille marrs, are also widely distributed in the lighter-colored mantle deposits of the central ridge.

The 100-m craters counted in the accumulation of the data reported here are not believed by the writer to be endogenous. The evidence for this conclusion is not compelling, however, and the resolution of this matter awaits the collection of more conclusive data. The arguments against the extra endogenous crater hypothesis are listed below.

(1) There seems to be no particular reason for a pronounced maximum of endogenous craters at 100 m. That is, it seems to the writer most probable that craters formed by internal mechanisms should, like impact craters, assume a smooth size-frequency distribution without a sharp cutoff at any given diameter.

(2) The 100-m craters counted in the compilation of the size-frequency data exhibit an apparently random distribution in the areas sampled, suggesting that they were formed by an extra-lunar agent (impact) rather than an internal process.

(3) The presence of approximately the same density of endogenous craters in the 100-m size range in both a highland basin and a mare surface would imply a high degree of similarity between these areas for considerable depth, requiring perhaps a common genesis. A common origin for highland basin deposits and mare surface material does not appear probable to the writer, particularly in light of the dissimilarities between these terrains on the km and regional scales.

An alternative interpretation of the crater density data is that the $<$ 300-m curve shown in Fig. 16 represents a size-frequency distribution which is the net product of two opposing processes, i.e., crater formation (impact) and crater obliteration (erosion). The near coincidence of the first (100-m) slope break and the observed minimum in crater sharpness would then be explained as follows. The largest crater obliterated in the history of the sampled areas is around 300 m in diameter. However, around half of the 100-m-diameter craters that have been formed remain. Since the rate of formation of these craters is low relative to smaller craters, and the erosion required to obliterate them small relative to larger craters, they are on the average more nearly obliterated (minimum sharpness) than either larger or smaller craters. The size-frequency distribution of craters smaller than 100 m in diameter represent a dynamic balance between formation and destruction rates, or an equilibrium distribution. Such craters will, of course, present on the average, a sharpness ratio nearly constant for all diameters; i.e., they will exhibit a complete range in crater sharpness from very fresh to nearly obliterated in appearance.

A diagrammatic illustration of the erosional hypothesis is given in Fig. 17 for Alphonsus. The curve marked "all craters formed" presents the size-frequency distribution of all craters formed in post-Alphonsus time. Similarly, the "all craters destroyed" curve depicts the size-frequency distribution of craters obliterated in post-Alphonsus time. The "craters remaining" curve is the present size-frequency distribution for Alphonsus craters, and represents the numerical difference between the other two curves. For small craters ($\sim$ 10 m), the number of craters remaining is almost insignificant compared to the number formed in post-Alphonsus time (at 10 m, 1/40 of those formed remain). In other words, virtually all small craters have been destroyed (98% at 10 m). The "craters remaining" curve is virtually identical for Tranquillitatis and Alphonsus below 300 m, as is shown in Fig. 16. Thus, the destruction and formation curves shown in Fig. 17 could be representative of Tranquillitatis as well as Alphonsus below 300 m. The curves shown in Fig. 17 suggest that the largest crater obliterated in post-

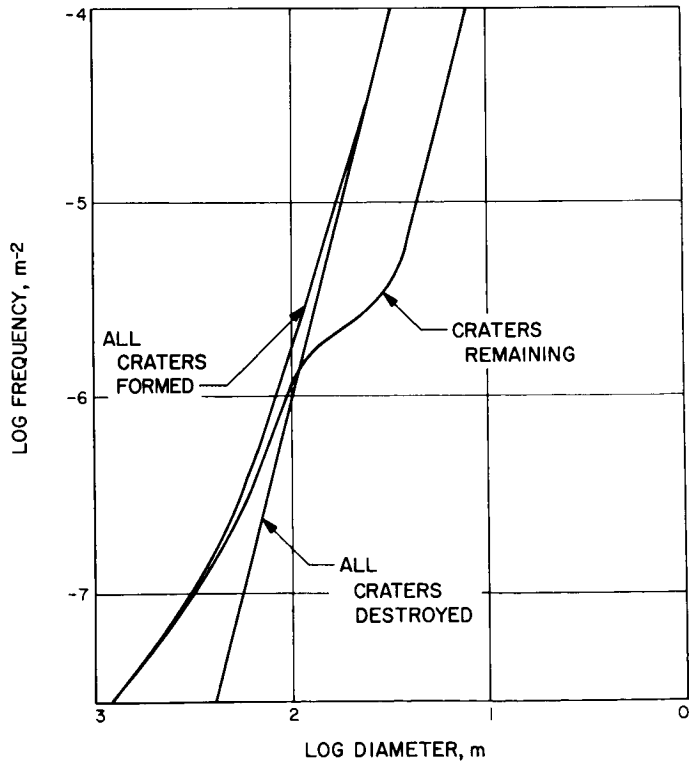


Fig. 17. Hypothetical crater formation and destruction curves, with the resulting crater size-frequency distribution

Alphonsus time is around 300 m in diameter, given an "erosion" rate of around 100 Å yr^{-1}, assuming an age of 4.5 billion years and an initial diameter-to-depth ratio of 5 or 6. The slope of the destruction curve shown in Fig. 17 is -4, and is identical to that of the formation curve for craters of diameter $\leq$ 60 m.

The choice of the latter interpretation of the data, while not requiring a special cratering mechanism which operates at around a 100-m scale, does imply that the maria were formed early in lunar history, since the age difference between Alphonsus and Tranquillitatis must be small compared to their absolute ages in order to explain the size-frequency anomaly as an erosional phenomenon. An additional implication of this hypothesis is that the pre-mare flux must have significantly exceeded post-mare flux only for bodies capable of forming craters in the km and greater diameter range. These two conditions are necessary for the "all craters formed" curve to be identical below 300 m for the terrae and the maria. It is interesting to note that linear extrapolations of the cumulative frequency curves given for Terrae and M. Imbrium by Kreiter (Ref. 6) intersect at about 200 m (Fig. 18).

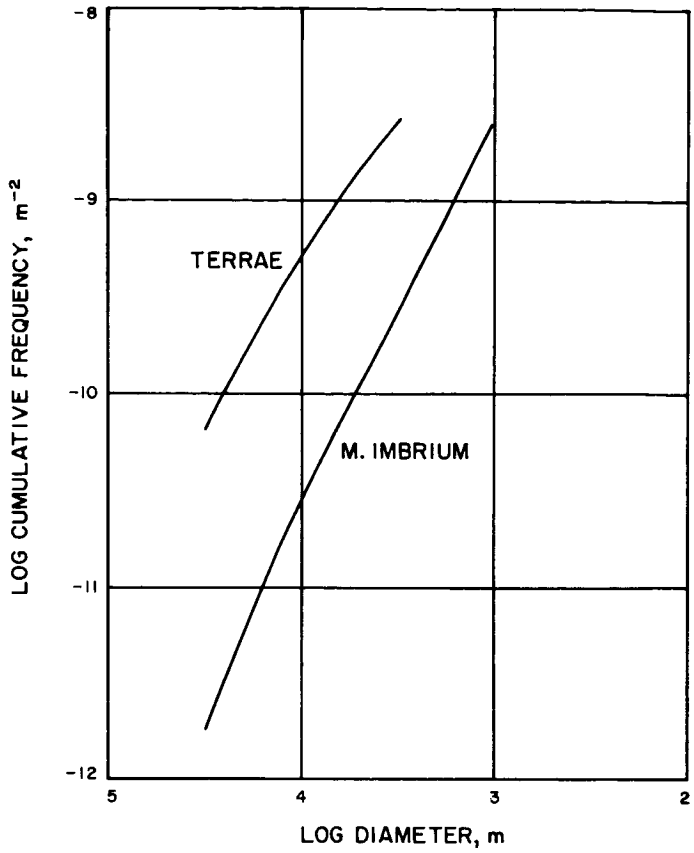


Fig. 18. Cumulative size-frequency curves for craters in M. Imbrium and Lunar Terrae

It is clear from the nature of the data available at this writing that a definitive choice of interpretation cannot be made. The essential question, assuming that the effect discussed (i.e., the discontinuity in crater size-frequency distribution at around 100 m) is a real one and not one of sampling statistics, is whether one is seeing extra endogenous craters or fewer exogenous ones. The writer favors the latter interpretation for two reasons in addition to those given above under the discussion of the first hypothesis. First, this interpretation seems to fit most smoothly with the idea of old maria (Shoemaker, Ref. 7), although still requiring higher pre-mare flux for large bodies. Secondly, the "erosion" (lateral distribution) rate resulting from this interpretation of the data ($\sim$ 100 Å yr^{-1}) is reasonable in light of the upper limit of 15–30 Å yr^{-1} obtained by Whipple (Ref. 8) and Whipple and Fireman (Ref. 9) for etching rates on the basis of the cosmic ray exposure ages of meteorites.

References

1. Sinton, W. M., *Astrophysical Journal*, Vol. 126, p. 231, 1957.
2. Sinton, W. M., *Science*, Vol. 130, p. 1234, 1959.
3. Moroz, V. I., *Astronomicheskii Zhurnal*, Vol. 41, 1964.
4. Kuiper, G. P., *Interpretation of Ranger VII Records*, Technical Report No. 32-700, Part II, Jet Propulsion Laboratory, Pasadena, California, p. 9, February 10, 1965.

References (Cont'd)

5. Shoemaker, E. M., *Preliminary Analysis of the Fine Structure of the Lunar Surface,* Technical Report No. 32-700, Part II, Jet Propulsion Laboratory, Pasadena, California, p. 75, February 10, 1965.

6. Kreiter, T. J., *A Possible Method of Dating the Lunar Surface Features,* Technical Release No. 34-57, Jet Propulsion Laboratory, Pasadena, California, April 11, 1960.

7. Shoemaker, E. M., Hackman, R. J., and Eggleton, R. E., "Interplanetary Correlation of Geologic Time," *Advances in the Astronautical Sciences,* Vol. 8, p. 70, 1961.

8. Whipple, F. L., "Solid Particles in the Solar System," *Journal of Geophysical Research,* Vol. 64, p. 1653, 1959.

9. Whipple, F. L., and Fireman, E. L., "Calculation of Erosion in Space from the Cosmic-Ray Exposure Ages of Meteorites," *Nature,* Vol. 182, p. 1315, 1959.

XVI. Chemistry

A. Further Studies of the Properties of Chelating Agents Coupled to Cellulose

A. J. Bauman, H. H. Weetall, N. Weliky, and L. Taylor

Carboxymethylcellulose sheet to which dithizone is coupled (*SPS 37-31*, Vol. IV, pp. 223–224) is sufficiently translucent to permit ordinary spectrophotometry. Table 1 lists a few visible maxima obtained on 1- × 3-cm sheets of Whatman CM-81 "dithizone cellulose" (DC) by means of a Cary 14 spectrophotometer. The data compares well with that for solution dithizone(ates) (Ref. 1), which suggests that dithizone is coupled through the sulfhydryl group of the dithizonic acid form (Ref. 2). Had coupling occurred in the ring, a marked upward shift in absorption wavelength would have occurred as is the case with azo dyes. Small strips of the DC should be useful in a mixed-color (Ref. 3) field hydrogeochemical method because they are stable when dry and are readily preserved in a notebook. Powdered-form DC is a good trace cation collector used either in batch or column form. We have studied the effects of pH and dilution on cation recovery in batch form and the recovery of cations from sea water by both techniques. In the pH study, samples of DC powder were stirred for 24 hr with solutions excess in zinc, copper, and gold and adjusted to various pHs. Zinc and copper were collected best at pH 5.0 in agreement with solvent extraction data (Ref. 1) but the gold data is ambiguous possibly due to adsorption of AU^{III} complexes at pHs above 7. In the dilution study, Cu^{II} at 6, 3, and 1 mg/liter was equilibrated at pH 5.0 with DC which recovered 30, 25, and 20% of the copper. In each case 0.25 g of DC powder were used but the volumes used were 1.3, 2.7, and 8 liters, respectively. It is possible that some loss of Cu^{II} occurred by adsorption on the walls of the large polyethylene bottles used in the experiment. When DC powder was equilibrated with an excess of Cu^{II} and zinc ion at pH 5.0, the powder picked up cations

Table 1. Comparison between visual wavelength absorption maxima of dithizone(ates) coupled to cellulose and dithizone(ates) in solution

	"Dithizone cellulose" wavelength, Å	Dithizone in CCl_4 wavelength, Å
H^+	4480,6120	4500,6200
Cu^{III}	3750,5250	5480
Zn	5400	5350
Ag	4700	4600
Co	5400	5600
Hg^{II}	5600–5700	5260
Pb	5500	5250

in the simple ratio of 1 meq of metal to 2 meq of carboxymethyl groups in the carboxymethylcellulose starting material. This agrees with Schmuckler's note (Ref. 4) that chelating resins in an excess of cations form complexes of the ratio 1:1 (of the form MR). Dowex-A-1 adsorbs 1 mM of metal ion per mM of nitrogen in the resin irrespective of charge. Complexes of the ratio 1:2 (form MR_2) are formed with trace quantities of cations.

The recovery of trace metals at extreme dilution, as from sea water, has been described by Schmuckler (Ref. 4) in his review of chelating resins. We chose to follow the protocol of Carritt (Ref. 5), who used a dithizone in carbon tetrachloride partition column method to isolate lead, zinc, cadmium, manganese[II], copper[II], and cobalt from sea water. He eluted the first four cations from the column with dilute acid and the last two with concentrated ammonia, which destroyed the column.

Clean sea water was collected directly from off the pier ends at La Jolla and Malibu Beach in large new polyethylene carboys (Ref. 6). We have completed a number of column runs, in each case using 20-g columns of DC, 4.2-cm ID by 5.5 cm high. The water was not filtered

Table 2. Dithizone cellulose column[a] collection of metals from sea water

Metal	Sea water, ppm[b]	% of total solids[b]	Column packing, μg recovered	0.05N HCl eluate, μg recovered	0.5N HCl, μg recovered	3N HCl, μg recovered	1N NH_4OH, μg recovered	Total enrichment ratio
Li	0.1	3×10^{-4}	—	—	—	—	—	—
Na	10561	30.6	—	3000	—	—	—	5.5×10^{-3}
K	380	1.10	—	—	—	—	—	—
Mg	1272	3.68	3200	2.3×10^{4}	700	2400	9.5	4.0×10^{-1}
Ca	400	1.16	9.2×10^{4}	4.4×10^{4}	700	1700	20	2.18
Sr	13	3.77×10^{-2}	—	620	—	—	—	1
Ba	0.05	1.5×10^{-4}	—	—	—	—	—	0
Ti	—	—	trace	—	—	—	—	—
Zr	—	—	—	—	—	—	—	—
Cr	—	—	—	—	trace	trace	—	—
Mo	0.0005	1.5×10^{-6}	—	—	—	—	—	—
Mn	0.001–0.01	3×10^{-5}	—	86	350	—	—	6.5×10^{2}
Fe	—	—	6400	210	1700	1.0×10^{4}	—	6.6×10^{-1}
Co	trace	trace	—	—	—	—	—	—
Ni	0.0001	5×10^{-7}	—	trace	trace	—	—	1
Cu	0.01	3×10^{-5}	740	130	42	980	2	2.2×10^{3}
Ag	0.003	1×10^{-6}	600	97	51	610	—	4.2×10^{4}
Au	0.000006	2×10^{-8}	—	—	1100	8.5×10^{3}	—	2.7×10^{7}
Zn	0.005	1×10^{-5}	—	—	—	—	—	—
Cd	trace	trace	—	—	—	—	—	—
B	4.6	1.3×10^{-2}	3200	22	—	3.4×10^{4}	—	1×10^{2}
Al	0.5	1.5×10^{-3}	1.4×10^{4}	320	—	1.0×10^{4}	—	3.9×10^{2}
Si	0.02–4.0	6×10^{-3}	8.2×10^{4}	350	1000	1.7×10^{5}	120	1.6×10^{3}
Sn	trace	trace	4800	—	2000	1.2×10^{3}	—	—
Pb	0.004	1×10^{-5}	—	420	—	—	—	2.3×10^{3}
Bi	—	—	—	—	—	—	—	—
Other	—	—	—	—	—	—	—	—

[a] 20-g column + 13.90 liter of La Jolla water at pH 7.95.

[b] Adapted from Table V-1 of *Oceanography* by Jerome Williams, Little Brown & Co., Boston, Massachusetts, 1962.

NOTE:

Sea water: total dissolved solids = 3.448×10^{4} mg/kg

Grand total of cations = 1.808×10^{6} μg from [fractions)] + [column—column blank]

= [225 + 315 + 613 + 450 mg] + [206.9 − 2.2]

and therefore contained microplankton. About 10–20 liters of water were passed through the columns (which quickly turned from green to red) at about 10–100 ml/min. The columns were then washed with 3–5 bed volumes of glass-distilled water to a negative chloride test then eluted with HCl and NH_4OH successively. The residual fractions obtained from the acid and base wash were weighed then sent along with the entire column packing to Pacific Spectro Chemical Laboratory for emission spectrographic analysis.

Calculations. Given the spectrographic data on the column and the fractions eluted one may calculate the total recovered weight of each element. One may further total all those weights to give a grand total of all cations recovered from the column. This amounts to "lumping" all recovered metals as one solid sample mixture without regard to anion. Now one can calculate the percentage of each metal recovered in the grand total. If one does a similar calculation for solid sea salt, one may compare the sets of data and obtain an enrichment factor for each element as compared with solid sea salt. This enrichment factor is a direct measure of the efficiency of the column; if no enrichment occurred this value would be 1. Table 2 shows the data for one sea water column run. Briefly, the data show likely quantitative recovery of silver, copper, and gold. Apparently the coupled dithizone chelates magnesium, calcium, and strontium well. We did not confirm Carritt's finding of cobalt. There is no advantage to eluting the columns with base after acid elution because they are essentially clean after acid elution.

B. Soil Studies—Desert Microflora. IX. Measurement of Soil Moisture Suction by Thermal Conductivity Probes

R. E. Cameron, C. W. Odd, and H. E. Britton

1. Introduction

Soil moisture content and soil moisture suction (tension, activity, stress, equivalent negative pressure, or pressure reduction) are important properties of the soil ecosystem. Water availability influences the physical and chemical behavior of the soil, and determines the activities, distribution, and kinds of microorganisms that are present. So far as is known, no organisms can grow and reproduce without the presence of available water at some stage during the life cycle. A test for water "availability" for activities of soil organisms has therefore become one of the most pertinent parameters to be determined in conjunction with extraterrestrial life detection experimentation.

Soil water relationships are influenced by a number of soil properties. Some of these properties include physical factors of soil texture, structure, aggregation, and distribution of soil particles, void relationships, temperature and vapor pressure fluctuation, solar heat flux, evapotranspiration rates, partial pressures of gases and their exchange rates, influences of gravity and topography, depth of soil, its color, and mineralogy, especially clays. Soil chemical and biological factors include the quantity, quality, relative abundances, distribution, and activity of salts and organic matter (including biota), cation-anion exchange capacity, buffer capacity, and oxidation-reduction state, of associated soil micelles. Both pH and redox potentials are properties of soil-water systems. The application of broad thermochemical principles and deductions are therefore necessary for the proper study and evaluation of soil-water phenomena.

Methods for determining and measuring soil moisture are usually applied to some specific range of soil moisture and are designed to indicate soil moisture suction, water content, or both. For biological purposes, it is desirable to know both the water activity as well as the water content of the soil.

Methods for determining soil moisture can be generally categorized as follows: (1) chemical, (2) electrical, (3) thermal, (4) lysimetric, (5) penetrometric, (6) tensiometric, (7) ionizing radiation, and (8) gravimetric (Ref. 7). Depending upon the method used, the measurements are either *in situ* or performed on an isolated sample of soil removed from the soil ecosystem. The gravimetric method has been most generally utilized, but necessitates the disruption of the soil ecosystem by the removal of a sample for analytical purposes. Soil moisture is most commonly indicated as a weight-percent of oven-dry soil (drying to a constant weight at 105 $\pm$ 5°C), although it is also expressed in terms of volume or on the basis of soil water suction. The most frequently applied methods for soil moisture measurement and a brief description of these methods has been given previously[1].

[1] Cameron, R. E., *Soil Moisture Measurement*, Unpublished Mimeograph, Jet Propulsion Laboratory, Pasadena, California, 12 pp., August 8, 1961.

2. Thermal Conductivity Probe Principles

The purpose of this experiment was to study the thermal conductivity (TC) method for measuring soil matrix suction (SMS) and to evaluate it for measuring available soil moisture (0 to 15 bars). No single method of measuring soil moisture is entirely satisfactory, but the TC probe method has some advantages for saline soils (Refs. 8 and 9). SMS measurements are much less affected by the salts in desert soils and can be made *in situ*. Also, temperature measurements of the soil are automatically recorded. Disadvantages are that a relatively long time is required for porous blocks to reach moisture equilibrium in dryer soils and hysteresis effects have been observed when relating these measurements to water content of the soils (i.e., drying curves differ from wetting curves).

The TC probe is essentially a porous medium or block surrounding a thermistor. When the probe is placed in contact with a soil, water flows between the soil and the probe until equilibrium is reached. The amount of water contained in the probe at equilibrium depends on matrix suction of the soil relative to that of the block. The theory of operation of the TC probe is based on the fact that the thermal conductivity of the porous medium is a function of the amount of water which flows into it. If the thermistor is heated by passing an electrical current through it, its temperature will rise. The temperature rise ΔT of the thermistor for a given heating current applied for a given time Δt is a function of the rate at which the heat generated in the thermistor is dissipated in the block. If the block contains water, its thermal conductivity will be greater than for a dry block and heat will be dissipated more rapidly. ΔT will be greatest in a block which has absorbed the least amount of water. Thus, ΔT can be calibrated to read the SMS of the soil or its water content.

3. Experimental

a. Probe construction. Castone block thermistor probes were originally designed, fabricated, and tested. These probes were made by casting glass rod thermistors in Castone (Fig. 1)[2,3]. The self-heating bridge (Fig. 2) was designed to quickly bring the thermistors in the probe to a preset temperature and to maintain that temperature for the duration of the heating period. Bridge current, which was a function of probe heat conductivity, was monitored with an Xy recorder.

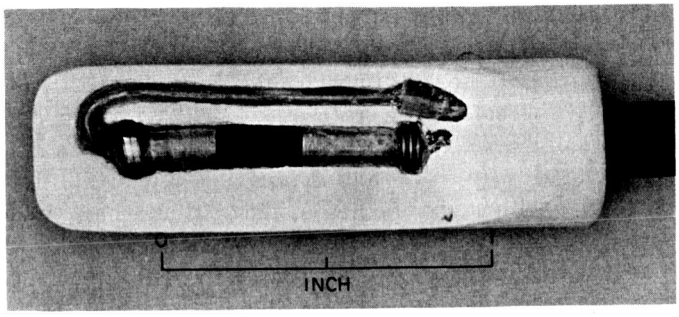


Fig. 1. Castone block thermal conductivity probe

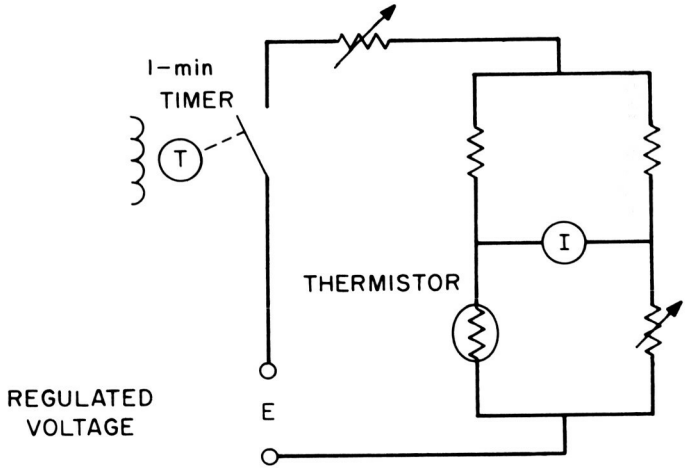


Fig. 2. Self-heating bridge circuit for thermal conductivity probes

Procedures for testing and calibrating the probes were the same as those described by Kelley (Ref. 10) except for some modifications[2]. Following the testing of Castone block thermistor probes and evident failure of all of these probes for reproducibility, non-Castone probes were designed, fabricated, and tested. These new TC probes consisted of bead-type glass ensheathed thermistors embedded in a short, cylindrical plaster absorption block[4] (Figs. 3, 4)[5]. The shaft of each thermistor sheath protruded from the block with an epoxy seal.

[2]Odd, C. W., *Soil Moisture Measurements for Desert Microflora Studies*, Interoffice Memo, Jet Propulsion Laboratory, Pasadena, California, 9 pp., December 10, 1963.

[3]Fenwal thermistor RB 33-L1, 200 Ω @ 25°C.

[4]Odd, C. W., *Thermal Conductivity Probes for Use in Measuring Soil Moisture Suction*, Interoffice Memo, Jet Propulsion Laboratory, Pasadena, California, 8 pp., November 20, 1964.

[5]Fenwal thermistor GB 32-P38, 2000 Ω @ 25°C.

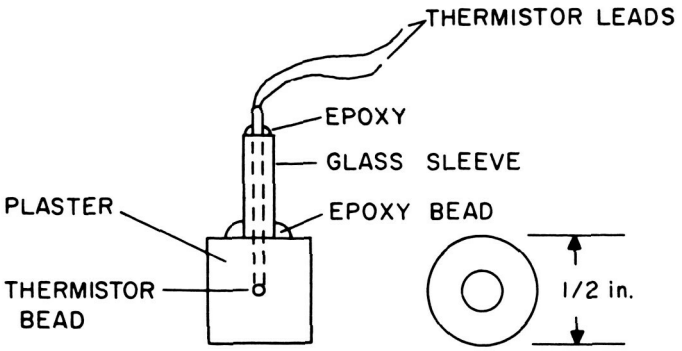


Fig. 3. Plaster block thermal conductivity probe

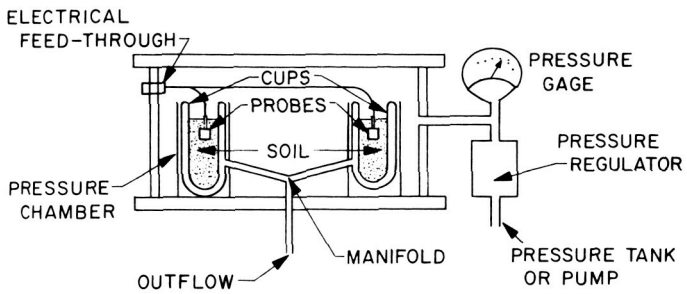


Fig. 5. Soil moisture extraction cups containing thermal conductivity probes in pressure chamber unit

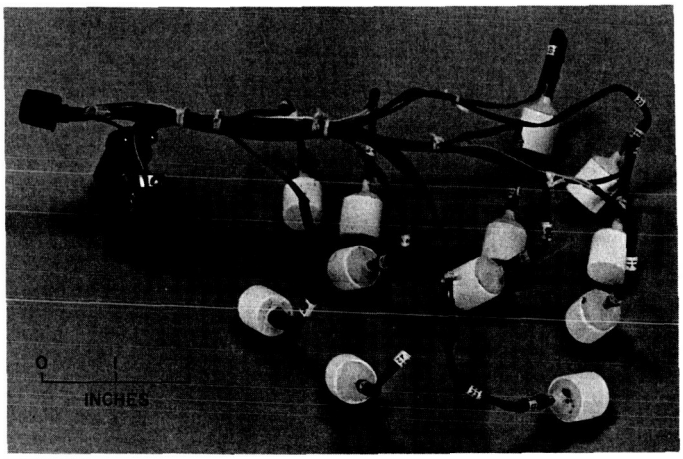


Fig. 4. Multiple plaster thermal conductivity probes arranged for testing in pressure chamber

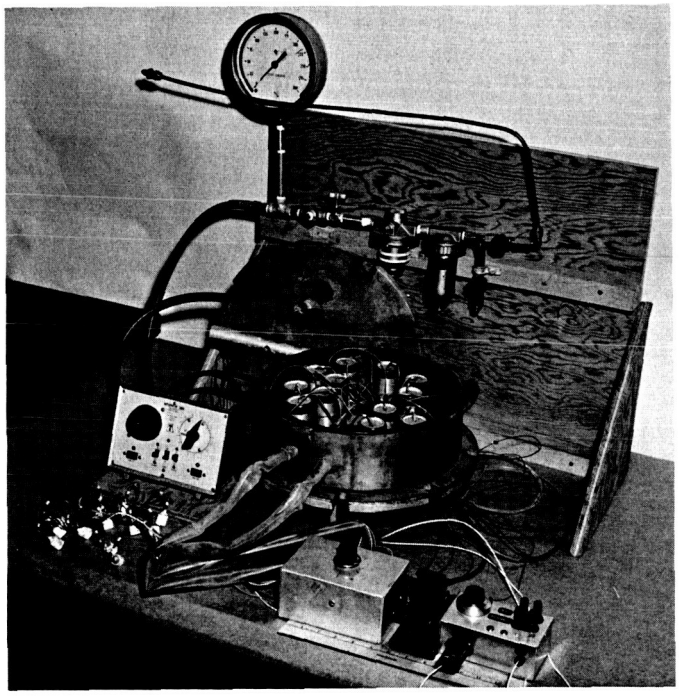


Fig. 6. System for calibrating and testing thermal conductivity probes at various soil moisture suction values

The plaster TC probes were fabricated using a variety of plasters and curing techniques because the sensitivity, stability, and range of probes are affected by the density and permeability of the plaster blocks.

b. Test procedures for plaster TC probes. Following preliminary testing of probes, all measurements were performed in soil moisture extraction cups enclosed in a pressure chamber[6] (Fig. 5) using the entire calibration system as indicated in Figs. 6 and 7. The pressure plate extractor chamber had been modified in order to use several cups simultaneously. Two electrical feedthroughs were inserted through the sidewall to accommodate the thermistor leads (Fig. 5). The membrane in the pressure plate was replaced with a brass plate of similar size and shape, containing a pressure outlet tube in the center. A manifold for 12 cups was attached to the main outlet tube and the cup outlet tubes were connected to the main outlet by a neoprene sleeve (Fig. 6).

A TC probe was placed in each of twelve 15 bar ceramic cups and packed with sieved desert soil, #76-1, described previously (*SPS* 37-32, Vol. IV, pp. 209–212). Soils were brought to the moisture saturation capacity (0.0 bar) and sealed in the pressure chamber. Standard soil moisture extraction procedures were followed to obtain hydraulic equilibrium of the TC probes at 0.0, 5.0, and 15 bars (Ref. 11).

[6]Soil Moisture Equipment Company: 15 bar ceramic cup, Category No. 1550; and pressure plate extractor, Category No. 1000.

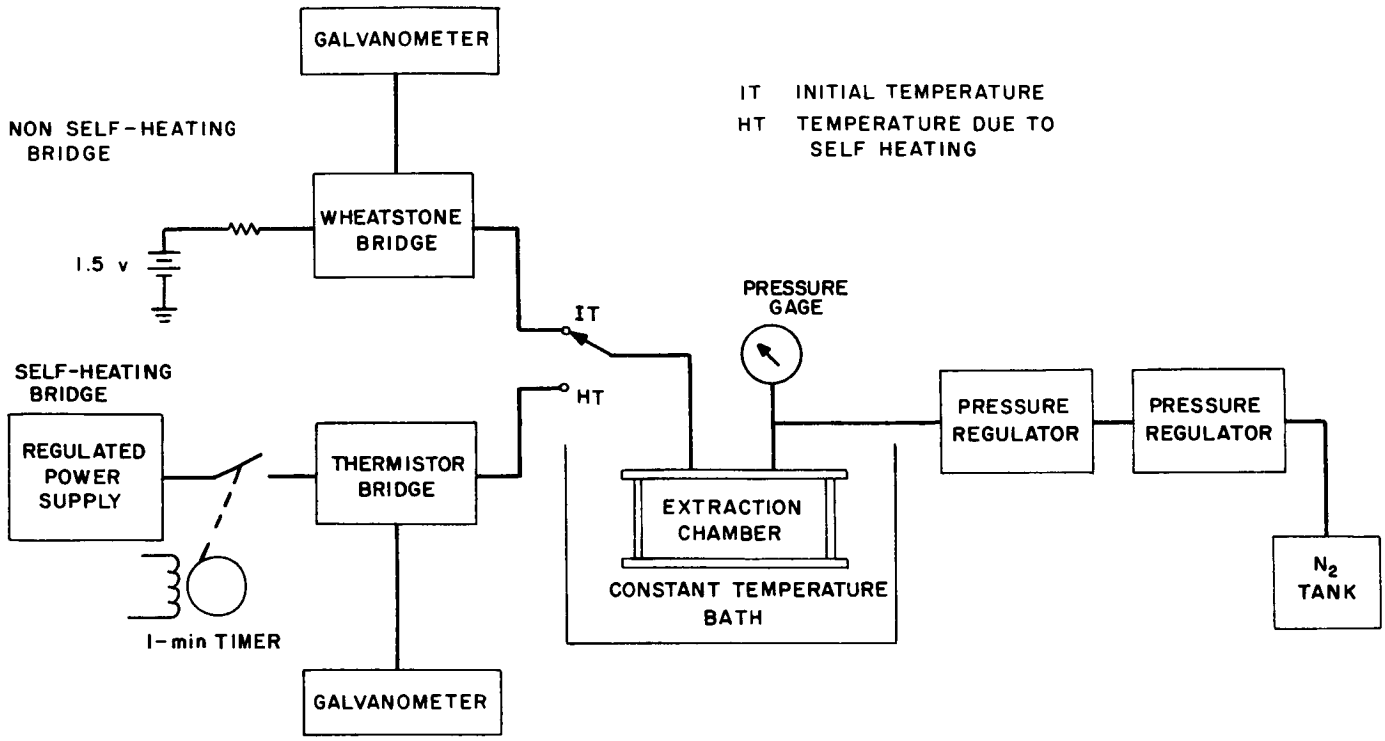


Fig. 7. Diagram of soil moisture suction calibration system for thermal conductivity probes

The temperature evaluation of the probe thermistors was determined by calculating the differences between the resistance of the thermistor before and after applying a self-heating current for one minute. The initial resistance R_i was measured first using a resistance bridge adjusted to apply a negligible self-heating current. R_h was measured with a second bridge which applied a 12-v regulated voltage for self-heating the bridge. Thus, the temperature elevation of the thermistor ΔR may be expressed as $R = R_i + R_h$.

To prevent resistance changes in the thermistors because of ambient temperature changes, the pressure chamber was totally submerged in a temperature bath controlled to 25 $\pm$ 0.03°C. At each pressure increment, new soil moisture and temperature levels were allowed to stabilize for at least overnight before taking resistance measurements.

4. Results and Conclusions

Figs. 8a and 8b show typical drying curves when current of thermistors is constant (Fig. 8a), and where the resistance is constant (Fig. 8b). Slope of drying curves is somewhat similar.

Results for determining soil moisture with plaster TC probes are given in Tables 3 through 7. Table 3 shows the ΔR values obtained for an extraction chamber pressure of 0.0 bar. Tables 4, 5, and 6 list ΔR values for extraction chamber pressure of 5 bar, and 3 equilibration times: (1) overnight, (2) 2 days, and (3) 3 days. Table 6 shows ΔR values obtained at an extraction chamber pressure of 15 bars, with an equilibration time $>$ 3 days.

Maximum deviation of R_h values are shown in Table 7. These values are an indication of the precision or reproducibility of the TC probes.

Although 12 TC probes were used for these tests, 5 of them failed before the tests were concluded. Failure could be attributed either to defective thermistors or else faulty TC probe construction. Probe failures occurred mainly after subjection to increasing chamber pressures, and could indicate that moisture was forced into the glass sheath surrounding the thermistor, thereby lowering the resistance path between the thermistor leads in parallel with the thermistor bead. Proper sealing of the glass sheath at the thermistor lead entrance could probably overcome this problem.

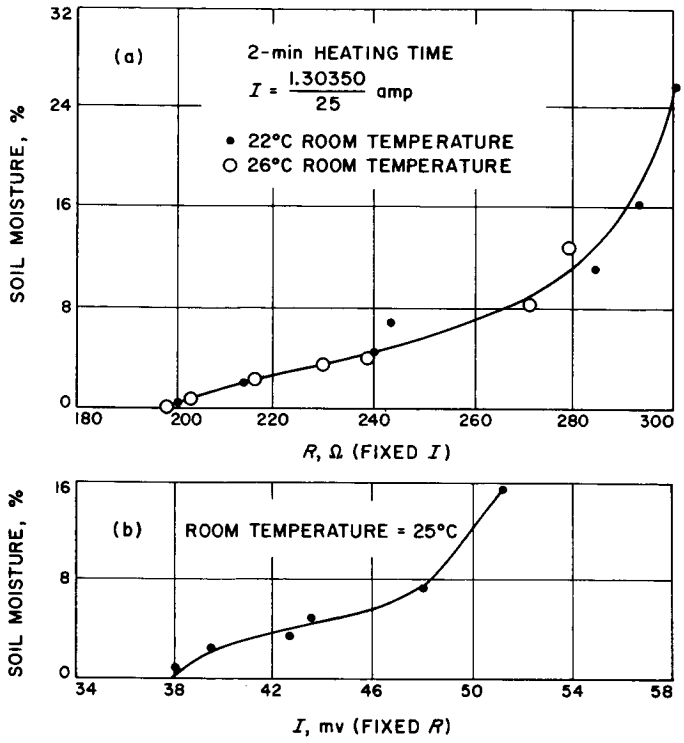


Fig. 8. Thermal conductivity probe sensitivity curves: (a) values of resistance ΔR with constant thermistor current I_t and variable soil moisture content; (b) values of ΔI_t with constant R and variable soil moisture content

According to Table 7, maximum deviation of R_h values was fairly large, approximately 20 to 25% at pressures of 5 and 15 bars. Precision could be improved by obtaining the mean value of a group of similar probes in the soil.

Table 3. Thermal conductivity probe resistance values at 0.0 bar[a]

Probe No.	Average R_i, Ω	Average R_h, Ω	Total deviation, Ω	Average ΔR, Ω
3	2314	266	3	2048
4	1762	216	9	1546
6	1897	259	7	1638
7	2212	252	3	1960
8	2088	227	2	1861
9	1979	221	3	1758
10	2115	321	4	1794

[a] Extraction chamber pressure = 0 bar (ambient at sea level), average of 3 tests.

Table 4. Thermal conductivity probe resistance values at 5 bar with overnight equilibration time[a]

Probe No.	Average R_i, Ω	Average R_h, Ω	Total deviation, Ω	Average ΔR, Ω
3	2314	257	4	2057
4	1762	198	11	1564
6	1897	236	3	1661
7	2212	243	5	1969
8	2088	213	3	1875
9	1979	212	4	1767
10	2115	301	7	1814

[a] Average of 3 tests.

Table 5. Thermal conductivity probe resistance values at 5 bar with equilibration time > 2 days[a]

Probe No.	Average R_i, Ω	Average R_h, Ω	Total deviation, Ω	Average ΔR, Ω
3	2314	255	3	2059
4	1762	191	1	1571
6	1897	223	5	1674
7	2212	232	8	1980
8	2088	208		1880
9	1979	204	2	1775
10	2115	285	2	1830

[a] Extraction chamber pressure = 5 bar, average of 2 tests.

Table 6. Thermal conductivity probe resistance values at 15 bar with equilibration time > 3 days[a]

Probe No.	Average R_i, Ω	Average R_h, Ω	Total deviation, Ω	Average ΔR, Ω
3	2314	248	0	2066
4	1762	188	1	1574
6	1897	215	2	1682
7	2212	218	2	1994
8	2088	202	7	1886
9	1979	196	1	1783
10	2115	275	1	1840

[a] Average of 3 tests.

Table 7. Comparison of normalized ΔR values for thermal conductivity probes varying in composition and curing process[a]

Probe No.	Normalized ΔR values			Probe adsorption block composition	Curing process
	$P = 0$ bar, Ω	$P = 5$ bar, Ω	$P = 15$ bar, Ω		
3	0	11	18	Hydrostonic	Vacuumed
4	0	25	28	Hydrostone & plaster	Vacuumed
6	0	36	44	Hydrostone & plaster	Purged
7	0	20	34	Plaster #1	Air-dried
8	0	19	25	Hydrostone & plaster	Air-dried
9	0	17	29	Plaster #1	Vacuumed
10	0	36	46	Ultracal & plaster	Purged

[a] ΔR values normalized to $P = 0$ bar.

With increasing soil moisture suction, $\Delta R'$ decreased rapidly (Fig. 9). Good sensitivity was obtained for some of the probes up to 5 bar and all were usable up to that point, but beyond 5 bars sensitivity decreased rapidly.

Sensitivity varied according to the particular TC probe that was tested. Two probes, #4 and #8, appeared to be quite insensitive, > 5 bar, while probe #7 showed usable sensitivity between 5 and 15 bars.

This would indicate that probe sensitivity was related to probe fabrication. Dense plasters, or dense plasters without purging appeared to produce TC probes with shorter ranges of soil moisture suction than probes with less dense purged plasters. More probes would need to be tested for results to be conclusive but these tests have indicated that factors affecting permeability of the absorption blocks apparently cannot be optimized to appreciably improve probe performance between 5 and 15 bars.

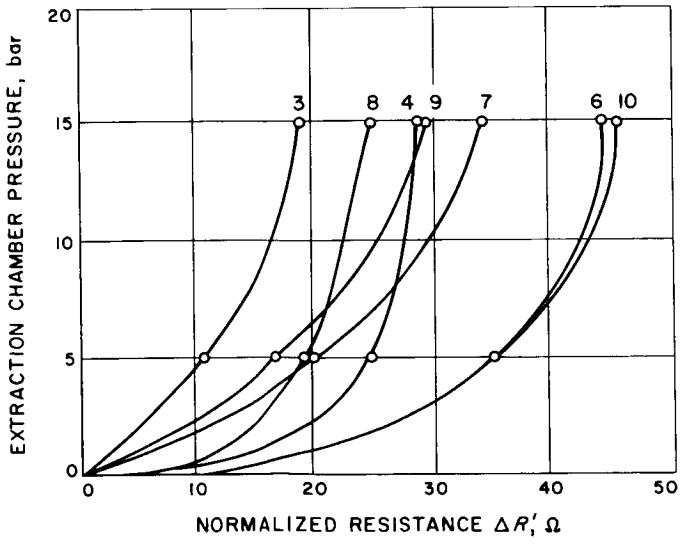


Fig. 9. Normalized resistance of thermistors vs extraction chamber pressure

C. Soil Studies—Desert Microflora. X. Soil Properties of Samples from the Chile Atacama Desert

R. E. Cameron, G. B. Blank, D. R. Gensel, and R. W. Davies

1. Introduction

Preparation for detection and characterization of extraterrestrial microorganisms, especially for Mars, is being supported by studies of terrestrial desert soils. Studies of these soils will provide useful background information for design and instrumentation of life detection experiments by supplying information on desert soil and micrometeorological properties, e.g., abundance and distribution of microflora populations, organic matter content, and moisture status. The purpose of this article is therefore to provide information on properties of soil samples collected from a desert region, the Chile Atacama Desert.

2. Atacama Desert

Among all the desert regions of this planet, portions of the Chile Atacama Desert have the harshest of known environments. This region contains several areas of "absolute" desert (based on extremely low and irregular rainfall and absence of macro life forms). Knowledge of the greater portion of this desert is meager, although it constitutes half the land area of Chile and part of Peru (Loma Desert) and northwest Argentina (Puna de Atacama)

The Atacama-Loma Desert region ranks twelfth in comparative size with other terrestrial deserts. It covers an area of approximately 0.14×10^6 mi^2 and is at an elevation between 1,500 and 13,500 ft with most of the desert between 7,000 and 13,500 ft. The especially arid area of the Atacama Desert extends north from Copiapó in north central Chile, including most of Antofagasta Province (Fig. 10).

In Chile and Argentina, the Andes Mountains form a barrier to westerly winds from the South Pacific. However, the extreme aridity of the Atacama-Loma Deserts is chiefly attributed to the continual subsidence of air in the South Pacific anticyclone belt with concurrent adiabatic influence of the cold Humboldt Current along the coast as a secondary factor.

Under natural conditions, this desert region is entirely barren except for scattered oases and piped water which may moisten the soil and allow for some vegetation. In many places it is also a salt desert. Borax "lakes," saline "dry lakes," and nitrate deposits have provided minerals for export for many years.

3. Area and Soil Site

Soil samples were collected northeast of Antofagasta, Chile, a seaport city which is also a railroad terminus in north Chile (Fig. 10)[7]. Climatological data for Antofagasta indicates that the annual potential evaporation is 66 cm (Ref. 12). The humidity regime is that of an absolute desert, the annual humidity index is 0.02, the normal

[7] Davies, R. W., Manager, Mission Definition, JPL Advanced Studies Section.

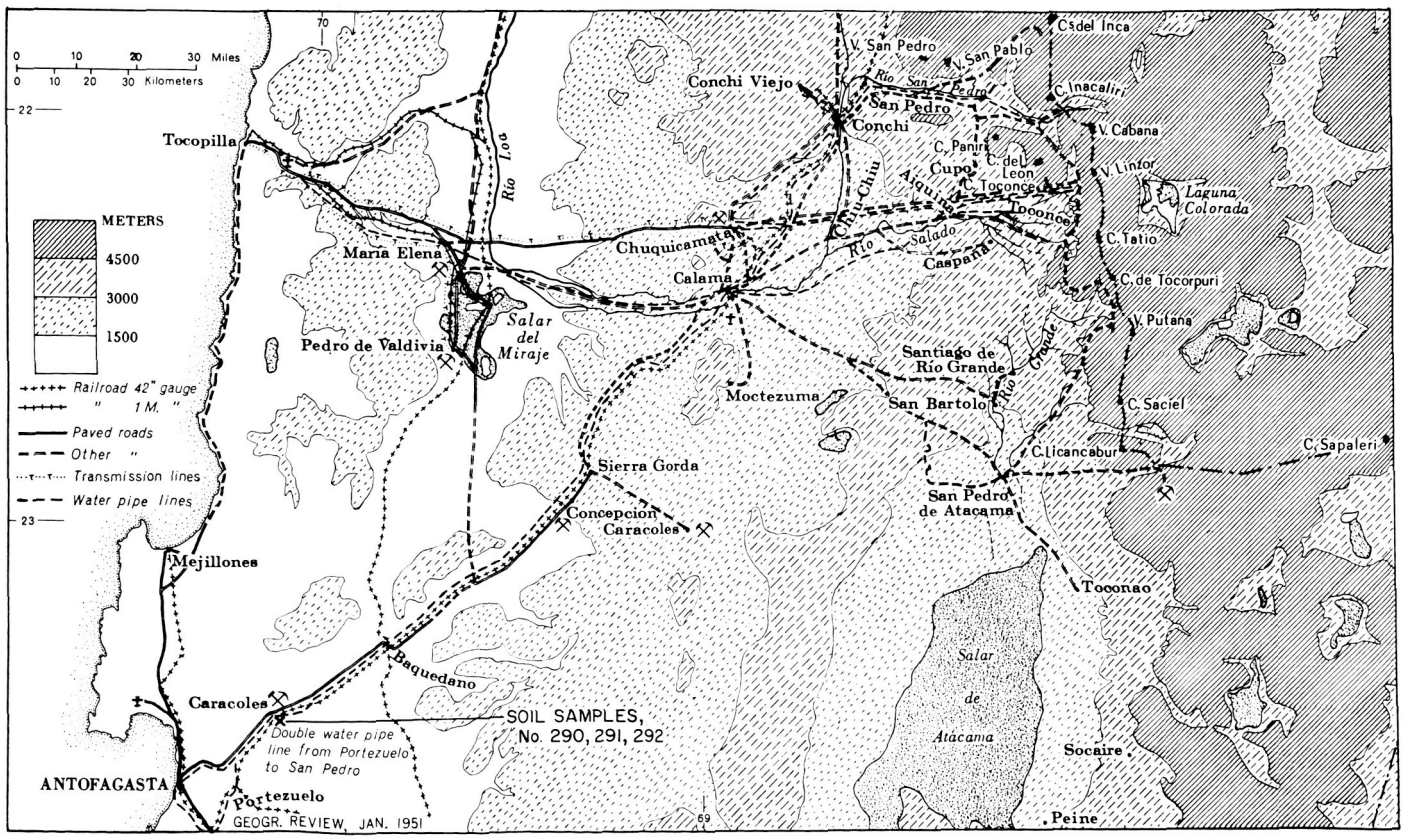


Fig. 10. Location of soil site in Chile Atacama Desert

leaching rainfall is zero (because a humid season is lacking), and all 12 months are categorized as dry. It was reported by faculty members of the Universidad del Norte in Antofagasta that in 1964 precipitation was less than 1 mm.[8] Previous to that precipitation, it rained approximately 12 times four years ago for a total rainfall of 5 mm.

The winter season is entirely frostless, but with a mean annual temperature under 15°C (sufficiently warm for maize but not cotton). A very cool tropical temperature regime is evident and the climate is generally desertic, comparing favorably with the Spanish Sahara in this respect. Soil development in this region is negligible and soils are either desertic entisols or aridisols. Vegetation, if present in the natural environment, would be primarily of scattered xerophytic shrubs or entirely restricted to soil microflora in extremely arid and barren areas.

The soil collection site was near Uriba, elevation 583 m, approximately 25 km northeast of Antofagasta and Baquedano (Fig. 10). Except for a few crustose lichens and cactus at one point, there was no visible vegetation for more than 10 km on the way to Uriba (Fig. 11)[9]. At approximately ⅓ km from the site, a few small rocks were found partially covered by thin desiccated growth of crustose lichens (Fig. 12). The terrain of this area showed no apparent modification or effects due to rainfall or water erosion.

[8] Father German Saa, Orlayer Alcayaga, and Anibal Galvez.

[9] Photograph by Dr. John Opfell, Aeroneutronic Division, Ford Motor Co., Newport Beach, California.

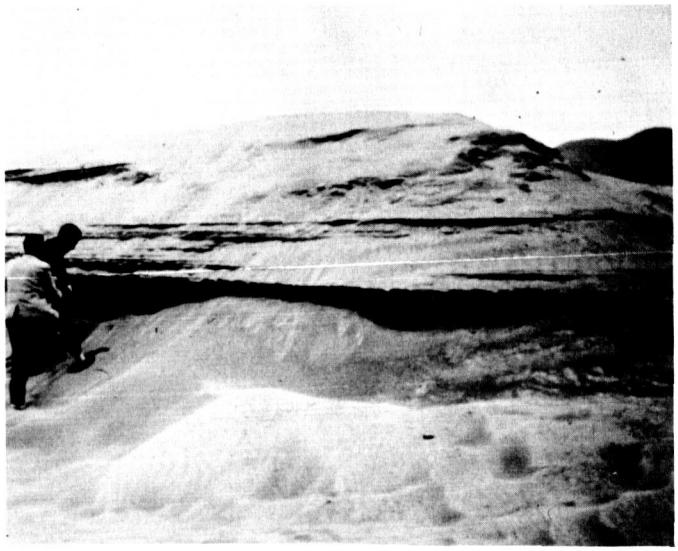


Fig. 11. Sand dunes at Chimba, Chile, approximately 7 mi north of Antofagasta

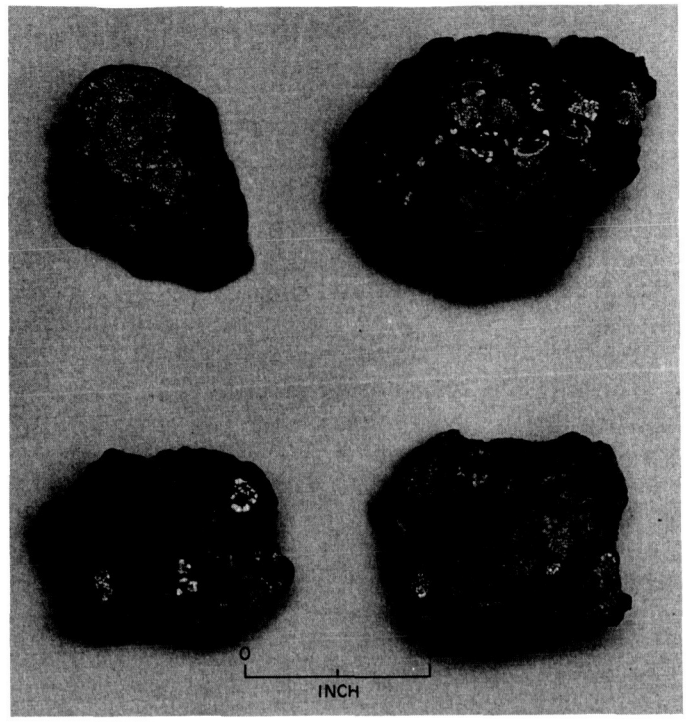


Fig. 12. Crustose ascolichens resembling *Parmelia sp.* on basaltic rocks

4. Materials and Methods

By means of aseptic technique, soil samples were collected from approximately the surface 1-cm (#290), 1- to 4-cm level (#291), and from the 4- to 12-cm depth (#292). At the time of collection it was noted that all of the soils were sandy with some white crusty material evident with an increase in quantity of this material with depth of soil.

Samples were next placed in previously sterilized plastic-lined canvas sacks. Additional samples were taken for soil moisture, porosity, and bulk density measurements. All samples were air-dry at the time of collection (1700 hr). Temperature and humidity measurements were made at that time. The soil temperature was 22°C (at 5 cm), air temperature was 18°C, and the relative humidity was 56%. Samples were later received at the Jet Propulsion Laboratory and analysis begun one month after the collection date. For analytical purposes, samples were homogenized by passing them through a 2-mm

sieve (#10 mesh) and an aliquot powdered by mortar and pestle for appropriate analysis.

a. Physical analyses. Soil color and Munsell notation were determined on both wet and dry soils by reference to Munsell Soil Color Charts. A mechanical analysis for soil particle distribution was performed by the Bouyoucous hydrometer method (Ref. 13) using fresh 5% Calgon for dispersal of salty soils[10]. The soil class was determined by reference to a Soil Textural Class Chart. *In situ* moisture was determined gravimetrically by infrared drying to constant weight at 105°C. Bulk density and porosity were calculated from samples collected in standard soil moisture tins.

b. Physico-chemical analyses. Soil pH was determined on both the saturated paste and 1:5 soil:water extract (Ref. 14)[10]. Soil redox potential was determined potentiometrically on the saturated paste. Electrical conductivity measurements were made on 1:5 soil:water extract and osmotic pressure calculated from these results[10] Combustible materials were determined by ignition in air at 750°C for 30 min. The cation exchange capacity values were obtained by modified barium chloride-triethanolamine procedure for calcareous soils (Ref. 15)[10]. Mineralogical analyses were obtained by X-ray diffraction with a scan speed of ½ deg/min[11]

c. Chemical analyses. Organic matter was determined by comparison of 1:1 soil:water autoclave extract with standard yeast extracts using a Klett-Summerson photometer (Ref. 16). Total carbon, hydrogen, and nitrogen analyses were performed by ignition in Coleman C–H and N-analyzers as described previously (Ref. 16)[12]. Organic carbon and carbonate carbon were obtained by the Allison method, utilizing acid digestion and gravimetric measurement of released CO_2 (Ref. 14). The Kjeldahl method was used to obtain organic nitrogen. Standard colorimetric methods were used to determine available ions on soil-water extracts. These ions included NH_4^+, NO_2^-, NO_3^-, Cl^-, Mn^{++}, $SO_4^=$, and $PO_4^{\equiv}$. Sulfur and P were determined on $HClO_4$ extracts. Sulfur was determined by the methylene blue method and spectrochemical analyses were performed for total elements by igniting soil mixed with graphite using a dc arc[13]. Oxygen values were calculated by difference.

d. Microbiological analyses. One-gram quantities of air-dried soil samples were suspended and dispersed by a wrist-action shaker in 9 cm^3 of distilled water. This suspension was then diluted serially by 1/10 and appropriate aliquots pipetted into selective media. For the various groups of microflora, this included the following media: (1) aerobic bacteria (including actinomycetes), trypticase soy agar, (2) microaerophilic bacteria, fluid thioglycollate medium, (3) anaerobes, trypticase soy agar in an atmosphere of CO_2, (4) fungi (molds), rose bengal agar, and (5) algae, 8 parts Pochon inorganic medium plus 1 part 1:1 desert soil extract (#76 soil from Mohave, California). All bacteria and fungi cultures were incubated for 2 to 5 weeks; algae were incubated in warm fluorescent light (450 to 550 ft-c) for at least 12 weeks. All incubations were conducted at room temperature of 22°C, except that the psychrophilic bacteria were incubated at 5°C. Details of media and culture conditions, as well as details of other analyses, will be presented in a future report[14]

5. Results and Discussion

A summary of the characteristics of the three soil samples is given in Tables 8–14. Physical determinations (Table 8) showed a narrow range of brownish-colored sandy soils with no material > 2 mm in diameter. None of the samples contained sufficient available *in situ* moisture. *In situ* moisture content increased with depth of soil, with increase in silt and salt content as well porosity, and varied inversely with the bulk density. Bulk density and porosity measurements are comparable to those of finer textured soils rather than sands and therefore indicate the influence of the crusty white salt (primarily nonpearly gypsum) on the formation of weak soil aggregates and the noticeable ease of soil compaction when walked upon at the collection site. Because of this specific salt influence, bulk density and porosity values simulated those obtained for clay or pumice soils, which are generally uncommon in desert regions.

Mineralogically, these soils are composed primarily of quartz, and secondarily of oligoclase (a plagioclase mineral, AB 90-70), and amphibole (possibly hornblende) (Fig. 13). All three of these are common in rock formations.

Physico-chemical determinations are indicated in Table 9. The pH values are typical for saline soils, which ordinarily have pH values less than 8.5. These soils are also

[10] Analyses by Edward S. Babcock and Sons, Riverside, California.

[11] Analyses by Neil L. Nickle, Engineer, JPL Lunar and Planetary Sciences Section.

[12] Analysis by Elek Microanalytical Laboratories, Torrance, California.

[13] Analyses by Pacific Spectro-Chemical Laboratory, Los Angeles, California.

[14] Cameron, R. E., and Blank, G. B., *Desert Soil Characteristics. I. Preliminary Studies*, Technical Report, Jet Propulsion Laboratory, Pasadena, California (to be published).

Table 8. Chile Atacama Desert soils, physical determinations

Soil No.	Depth of sample	Color and Munsell notation		Mechanical analysis				Textural class	*In situ* moisture, %	Bulk density, g/cm³	Porosity, %
		Wet	Dry	Sand (2 mm–50 μ), %	Silt (50–5 μ), %	Clay (5μ or <), %	Material (> 2 mm), %				
290	Surface 1-cm	7.5YR 4/2–5/4[a], dark brown-brown	7.5YR 4/4–5/4, dark brown-brown	98	1	1	0	Sand	1.18	0.99	62.6
291	1–4 cm	7.5YR 4/2–4/4[b], dark brown	7.5YR 5/2, brown	96	3	1	0	Sand	2.50	0.93	64.9
292	4–12 cm	7.5YR 4/4[c], dark brown	7.5YR 6/2, pinkish gray	93	6	1	0	Sand	2.60	0.88	66.8

[a] Considerably mottled.

[b] Slightly mottled.

[c] Gypsum mottles present.

Table 9. Chile Atacama Desert soils, physico-chemical determinations

Soil No.	Depth of sample	pH		Saturated paste Eh, +mv	1:5 soil: H_2O extract electrical conductivity × 10^6, mhos/cm at 25°C	1:5 soil: H_2O extract osmotic pressure, atm	Loss on ignition, 30 min at 750°C	Cation exchange capacity, meq/100 g soil[a]
		Saturated paste	1:5, soil: H_2O extract					
290	Surface 1-cm	7.6	6.7	566	325	0.117	1.19	2.5
291	1–4 cm	7.7	7.4	445	760	0.344	1.67	2.5
292	4–12 cm	7.7	7.6	463	2225	0.801	2.62	5.0

[a] Primarily Ca^{++}.

called "white alkali" soils or "Solonchaks" and are often recognized by the presence of white crusts, and a flocculated soil condition. The electrical conductivity, osmotic pressure, and cation exchange capacity values also indicate a saline soil. As indicated in Table 9, there is an increase in all of these values, including weight loss upon ignition, with increase in depth of soil and concurrent increase in soluble salts. Electrical conductivity and corresponding osmotic pressure are noticeable for Soil #292, and growth of some plants would be restricted. Weight

Table 10. Chile Atacama Desert soils, chemical determinations % of dry weight

Soil No.	Depth of sample	Total organic matter	Carbon-nitrogen-hydrogen values								
			Total carbon	Total nitrogen	Organic carbon	Organic nitrogen	Carbonate carbon	Total hydrogen	C:N total	N:H total	H:C total
290	Surface 1-cm	0.06	0.10	0.01	0.03	0.01	0.05	0.20	8.7	0.05	2.1
291	1–4 cm	0.06	0.08	<0.01	0.02	<0.01	0.09	0.18	11.4	0.04	2.3
292	4–12 cm	0.07	0.12	<0.01	0.04	<0.01	0.10	0.43	20.7	0.01	3.5

Table 11. Chile Atacama Desert soils, abundance of chemical elements % of dry weight

Element	Soil number		
	290	291	292
O	48.2	47.7	44.8
Si	29.5	25.5	22.0
Al	8.5	11.5	8.9
K	4.95	4.80	5.90
Fe	3.05	3.85	3.55
Na	2.45	2.60	2.45
Ca	1.10	1.45	7.0
Mg	1.02	1.25	1.25
Ti	0.505	0.545	0.780
H	0.201	0.183	0.434
S	0.058	0.191	2.517
P	0.123	0.129	0.128
C	0.096	0.080	0.124
Mn	0.091	0.081	0.057
Ba	0.05[a]	0.05[a]	0.05[a]
Sr	0.031	0.018	0.018
Pb	0.0225	0.0225	0.010[b]
Zr	0.0225	0.0325	0.0155
N	0.011	0.007	0.006
Cr	0.0155	0.0185	0.0125
V	0.0110	0.0135	0.0120
Ga	0.0049	0.0058	0.0046
Cu	0.00615	0.00585	0.00465
Co	0.00285	0.00315	0.00250
Ni	0.00260	0.00275	0.00180
B	0.00195	0.00190	0.00165

[a] 0.05% or less.
[b] 0.01% or less.

Table 12. Chile Atacama Desert soils, available cations, ppm by weight

Cation	Soil: H_2O extract	Soil number		
		290	291	292
Ca^{++}	1:5	168	625	2700
Na^{+}	1:5	50	150	240
Mg^{++}	1:5	53	55	53
K^{+}	1:5	25	50	30
Mn^{++}	1:10	<0.5	<0.5	<0.5
NH_4^{+}	1:5	1.2	1.0	0.5

Table 13. Chile Atacama Desert soils, available anions, ppm by weight

Anions	Soil: H_2O extract	Soil number		
		290	291	292
$SO_4^{=}$	1:10	225	700	450
Cl^{-}	1:10	<20	50	50
NO_3^{-}	1:5	1	3	4
$PO_4^{\equiv}$	1:10	1	1	1
NO_2^{-}	1:5	0	0	0

loss of soil upon ignition is due primarily to salt and "bound" water, and not organic matter or clay. Eh determinations show a well-oxidized soil.

Chemical determinations for organic matter, carbon, nitrogen, and hydrogen are indicated in Table 10. Organic matter content, as obtained by the autoclave method

Table 14. Chile Atacama Desert soils, microbiological determinations

Soil No.	Depth of sample	Aerobes: bacteria + actinomycetes/g soil		Microaerophiles growth in dilutions/g soil[b] 10^5	10^6	10^7	Anaerobes/g soil	Fungi/g soil	Algae/g soil
290	Surface 1-cm	Mesophiles 8.0	Psychrophiles 3.5	+	+	+	0.5	1.7	0.0
291	1–4 cm	4.0	80	+	+	+	0.0	0.0	0.0
292	4–12 cm	38[c]	14	+	+	−	0.0	1.0	0.0
		Trypticase soy agar		Fluid thioglycollate medium			Trypticase soy agar; incubation in CO_2	Rose bengal agar	Pochon salt solution + soil extract

[a] Mesophiles: incubation at 20°C for 2 weeks; psychrophiles: incubation at 5°C for 5 weeks; all other incubations at room temperature.
[b] + = growth; − = no growth.
[c] Predominately actinomycetes.

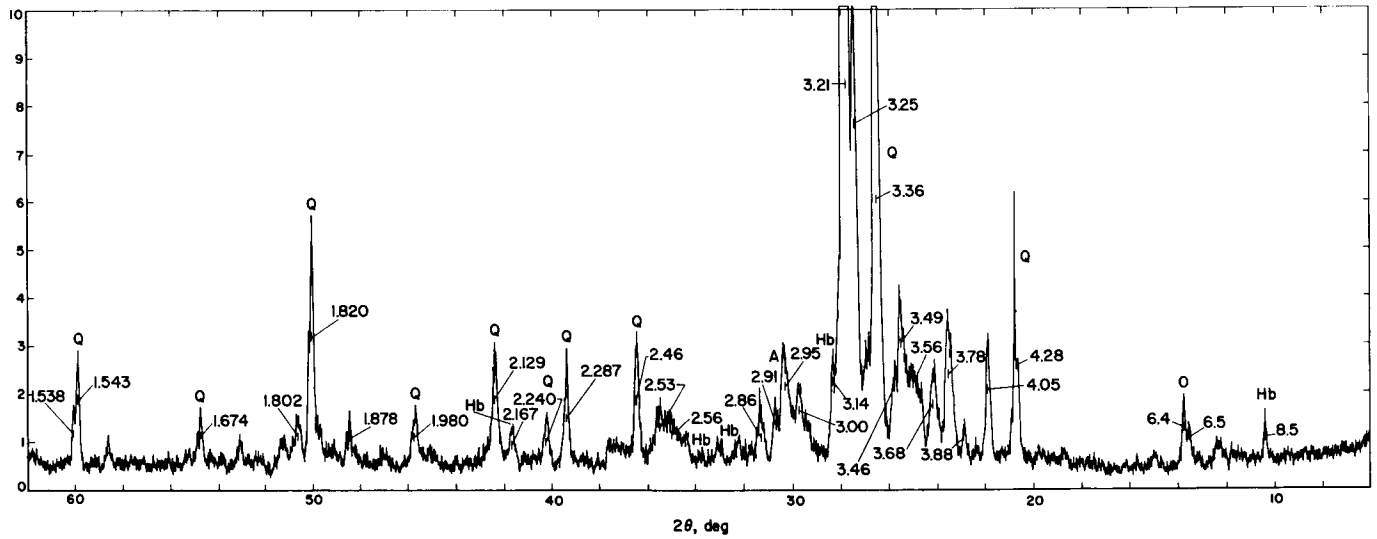


Fig. 13. X-ray diffraction diagram for Soil #291 [Q = quartz O = oligoclase, A = amphibole (hornblende ?), Hb = hornblende]

(Ref. 16), is quite low, but also corresponds to calculated values based upon the organic carbon in these soils. Organic matter content corresponds to the lowest values found at depths of 3 to 4 ft in California desert soils[14]. Total carbon is primarily inorganic, in the carbonate form, and increases with depth of the soil "profile." Organic carbon is evenly distributed with depth, and is less than that found in typical nonabsolute deserts or volcanic deserts. Total nitrogen is entirely in the organic form, except for a few ppm of ammonia and nitrate (Tables 12 and 13). Organic nitrogen figures compare favorably with low organic nitrogen content of subsurface soils in California deserts.

C:N values, as based on total carbon and nitrogen, are comparable with those of other desert soils, but organic C:N values are 3.0, 2.0, and 4.0, respectively. with depth of soil and are very narrow. This would indicate that the organic matter is ir the form of protein and of a microbial nature.

Total hydrogen increased with Sample #292, with corresponding decreases in the total N:H value and inversely with the H:C value. Hydrogen was present mainly in salt-silt-water complexes, and was not derived from organic matter, which provides a substantial quantity of hydrogen in agricultural soils.

The availability of prominent cations and anions on the exchange complex is given in Tables 12 and 13. As is indicative of saline soils, the chief anions are sulfate and chloride. The chief cations are calcium, sodium, magnesium, and potassium. A comparison of concentration of ions in Tables 12 and 13 with total abundance of elements in Table 11 indicates the relative availability status of these elements.

Abundance of chemical elements shows that the oxygen and aluminum content of these soils decreases with depth of soil primarily because the content of calcium and sulfur increase. There was also a higher concentration of potassium, titanium, and hydrogen in Soil #292. Two minor or trace elements, boron and lead, were of higher concentration than normally found in our previously investigated desert soils, and boron toxicity has been noted in arid areas where sodium and calcium borates are not removed by leaching. A soluble boron content > 2 ppm is undesirable for plant growth. The lead content is 10 to more than 20 times that found in soils and could also present a toxicity problem if in soluble form. The concentration of lead found in these soils is of greater concentration than any other desert soils we have investigated. Other elements of above the average content for agricultural soils include cobalt, copper, gallium, magnesium, and potassium, as well as sodium, calcium, and sulfur (Ref. 17).

Abundance of common groups of microflora found in soils is shown in Table 14. The only abundant group found in these soils was the microaerophilic organisms. The numbers of these organisms are comparable to their abundance in California desert soils. The abundance of other common groups of microflora, aerobes, anaerobes, molds, and algae, found in these desert soils was the lowest we have ever found, comparing only with a few atypical soils previously collected from volcanic deserts. No algae were detected and the anaerobes and fungi were either not detected or present in concentrations of about one organism per gram of soil. The aerobes were present at slightly higher concentrations, but predominately in soil collected from below the surface 1-cm. A significant proportion of the aerobes were psychrophiles, and predominately an actinomycete (Fig. 14). This

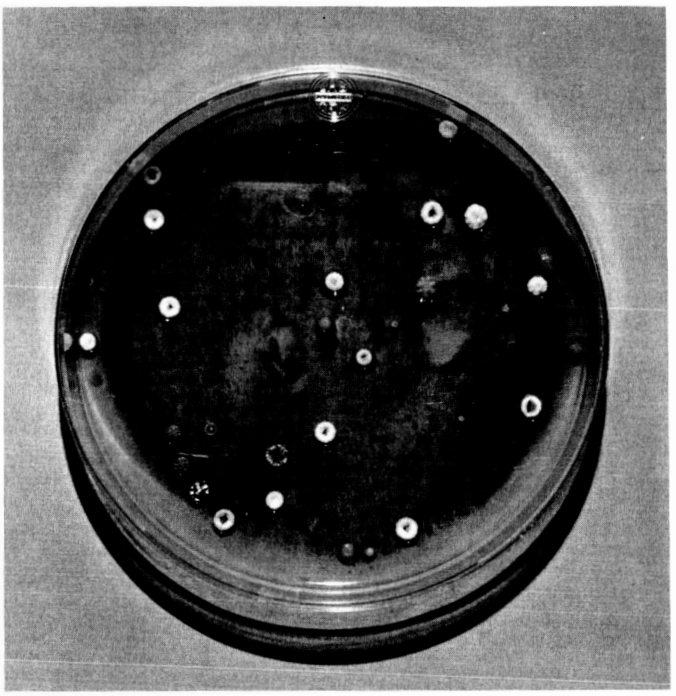


Fig. 14. *Streptomyces sp.* from one gram of soil growing on trypticase soy agar

actinomycete most closely resembled a *Streptomyces sp.* The molds were represented by three species which grew in comparable abundance on both rose bengal agar and trypticase soy agar. *Penicillium sp.*, which is commonly found in other desert soils, was the only mold present in Soil #292 (Fig. 15).

Although no algae were found in these soil samples, the presence of a few thin, crustose, desiccated, depauperate lichens on rocks approximately ⅓ km from the collection site (Fig. 12) indicates that algae, as well as other microflora, could be present in other soil samples from the same general area. An examination of this lichen showed a number of simple black ascospores and coccoid green algae. These algae were *Protococcus grevillei* and represent the most abundant species of green alga found in southwestern United States desert soils. Filamentous blue-green soil algae, including several new species, were collected previously near Antofagasta, and Caldera, Chile (Ref. 18).

6. Concluding Remarks

Three soil samples from the Chile Atacama Desert were studied and their properties compared with values obtained for other desert soils. Although all three samples were taken from the same site, there were apparent

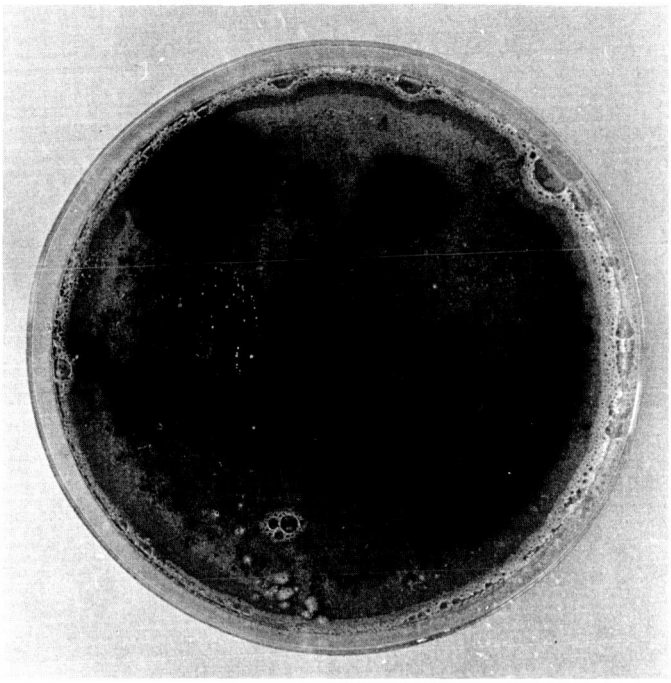


Fig. 15. *Penicillium sp.* from one gram of soil growing on rose bengal agar

differences within a 12-cm depth of soil, such as in textural composition, moisture content, bulk density, porosity, pH, electrical conductivity and osmotic pressure, weight loss upon ignition, cation exchange capacity, availability and abundance of some ions and elements, and abundance of microflora.

Compared with other desert soils, the most noticeable difference was in regard to the very low abundance of common groups of microflora. These values compared with some of the values on aerobic microorganisms grown out in trypticase soy agar from soil samples collected in nearby areas of the Antofagasta-Atacama desert region.[15] It is especially important to note the abundance of the microaerophiles, which is similar to that in other desert soils having much higher counts of the other groups of microflora. This would indicate that specialized groups of microorganisms are present and selective culture media and conditions are necessary to show their presence. The very low values obtained for organic matter, carbon, and nitrogen are indicative of the harsh environmental nature of this absolute desert, and would also show that autotrophy, rather than heterotrophy is the primary mode of existence. Those microorganisms which did grow out, grew slowly and macroscopic appearance of colonies took more than 48 to 72 hr or as much as 30 days. This was contrary to the usual rapid growth of most microorganisms which we have cultured from desert soils. It should also be noted that a comparatively large amount of soil sample is needed in order to detect and grow out some of the common groups of these desert soil microflora.

More soil samples are obviously needed from widely spread sites and depths of soil. Only the periphery of the Atacama Desert has now been examined, and this is the area most likely to be contaminated by nonindigenous microorganisms. A transect through the higher elevations of the Atacama, beginning at the seacoast, but especially from 7,000 to 13,500 ft and higher to the Puna-Atacama and Andes Mountains is desirable. Detailed environmental measurements should be made, e.g., time, duration, and amount of dew formed, and these measurements correlated with abundance and activity of the desert soil microflora. This desert area should present an excellent field test situation for extraterrestrial life detection experimentation prior to actual tests in an even more harsh extraterrestrial environment.

[15] Opfell, John, *Effect of Soil Extracts in Medium in Detecting Aerobic Microbial Populations in Desert Soils from the State of Antofagasta, Chile,* Aeroneutronic Division, Ford Motor Co., Newport Beach, California, 13 pp., July 26, 1965 (unpublished report).

D. The Reaction of $O(^1D)$ with Methane

W. DeMore and O. Raper

1. Introduction

The reactions of atomic oxygen with saturated hydrocarbons are among the least understood of all atom reactions. Rates of reaction of $O(^3P)$ with several of the lower alkanes have been reported (Refs. 19–24), but the reaction mechanisms are not known with certainty and in some cases the reported rates include secondary reactions to an unknown extent. In the case of $O(^1D)$, Yamazaki and Cvetanovic (Ref. 25) have shown that insertion into the CH bonds of propane gives an excited propanol which is either deactivated or undergoes radical fission. This latter result is consistent with the work of Basco and Norrish (Ref. 26) on the flash photolysis of O_3–CH_4 mixtures in which the formation of OH radicals was demonstrated spectroscopically.

We have investigated the reaction of $O(^1D)$ with CH_4 by photolyzing O_3–CH_4 mixtures dissolved in liquid Ar at

87°K. The results show that the insertion reaction has an activation energy $\leq$ 0.5 kcal/mole and occurs at a rate comparable to the reaction of $O(^1D)$ with O_3. The insertion product CH_3OH is formed in substantial yield but in addition there is strong evidence of a radical fission process giving CH_3 + OH and a molecular fission process giving $CH_2O + H_2$. The experiments have also led to new information regarding the oxidation of the radicals, CH_3, CH_3O, and OH.

2. Experiments

The photolysis cell used in this work has previously been described in detail (Ref. 27). High purity tank grade Ar for use as the solvent was found to contain not more than 1 ppm O_2 by the lamp test (Ref. 28). The methane was Matheson Research Grade, which was used without further purification except in some cases where it was passed over Ca metal at 400–500°C to remove traces of O_2. Tank grade CF_4 (Matheson) was similarly freed of O_2. The light source was a low-pressure mercury lamp with a methanol filter which limited the absorbed radiation almost entirely to 2537 Å.

Samples were prepared by first distilling the required amount of O_3 into the evacuated cell at 77°K, followed by condensation of sufficient argon to fill the cell. During solvent condensation the cell was allowed to warm to 87°K by replacing the N_2 coolant with liquid argon. Methane was introduced by means of a doser which was arranged so that the methane could be swept into the cell by the argon flow.

For analysis of the product H_2 the cell was opened to an evacuated gas chromatography sampling tube. The column was molecular sieve at 35°C and the carrier gas was Ar. The procedure was repeated until analysis of the gas sample showed no further H_2 content, with three or four samplings usually being sufficient. The H_2 peak heights from each sample were summed and were related to the original H_2 concentration by comparison with similar measurements on synthetic mixtures of H_2 in liquid argon. The calibration curve is shown in Fig. 16.

Products not volatile at 77°K were isolated by pumping off the Ar and CH_4 and were then collected by vacuum transfer to an ampoule at 77°K. The sample was then flash-evaporated into a mass spectrometer inlet system. Analysis of synthetic samples of CH_3OH and CH_2O by this method showed a collection efficiency of 85 to 90%, and therefore all reported yields of these products are corrected on the basis of 90% recovery.

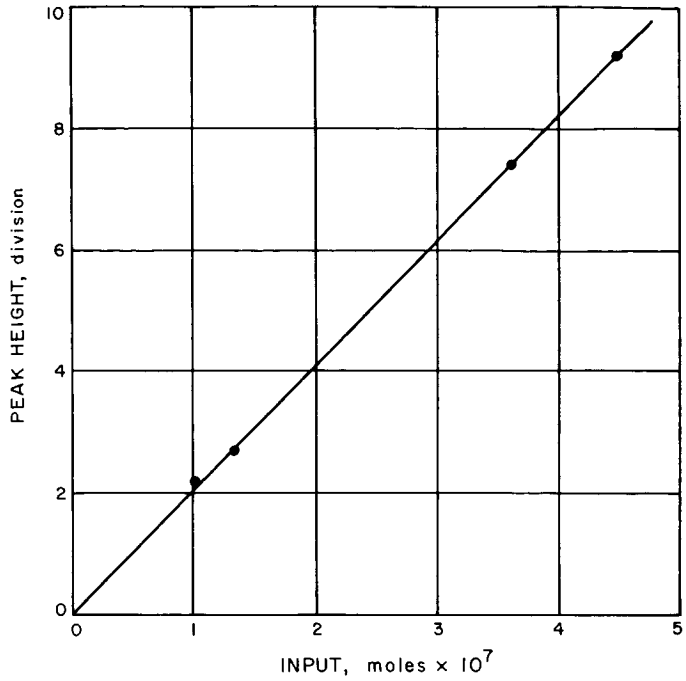


Fig. 16. Calibration for hydrogen analysis

3. Results

a. Products without added oxygen. The products found in the photolysis of 0.4 mole/liter solutions of methane at varying O_3 concentrations were CH_3OH, CH_2O, and H_2. The mass spectrum of the non-volatile sample fraction showed no trace of any peak not due to CH_3OH or CH_2O, with the possible exceptions of H_2O and CO_2, which are always present as background impurities. In particular there was no mass spectral evidence for the presence of CH_3OOCH, H_2O_2, or CH_3OOH. However, these latter compounds are unstable and are subject to decomposition in the mass spectrometer. The product yields of CH_3OH, CH_2O, and H_2, based on O_3 decomposed, are shown in Table 15. The O_3 was photolyzed to completion in all experiments except for the one case indicated in the table. The product yields appear to be independent of the initial O_3 concentration, and partial decomposition of O_3 gave product yields which did not differ significantly from those obtained for complete O_3 decomposition. The latter result shows that product photolysis is not a source of error in these experiments, since O_3 absorbs virtually all of the 2537 Å radiation except in the last stages of those experiments in which O_3 is photolyzed to completion.

Analysis of the reaction mixtures before irradiation confirmed the absence of H_2 as an initial impurity. It was also found that further irradiation of solutions which

Table 15. Product yields

O_3 initial, mole/liter × 10^3	ΔO_3, moles × 10^6	Product yields based on ΔO_3, %		
		CH_3OH	CH_2O	H_2
1.00	3.00	55	6.7	
0.95	2.84	51	8.4	
0.87	2.63	49	5.4	
0.65	1.97	52	9.2	6.0
0.62	1.86	49	7.3	6.2
0.61	1.84	51	3.9	
0.58	1.08[a]	51	8.8	5.1
0.56	1.67	52	6.5	
0.50	1.50	53	7.4	6.0
0.44	1.33	51	5.9	

[a] O_3 62% photolyzed.

had been photolyzed to completion and stripped of H_2 did not produce any more H_2. Thus H_2 is formed only during the stage of O_3 photolysis.

b. Products with added oxygen. The product yields with O_2 added to the reaction mixture are shown in Table 16. With $O_2 \gg O_3$, the yields are all increased and reach limiting values of 68% (CH_3OH), 11% (CH_2O), and 7% (H_2).

Table 16. Product yields with added oxygen

O_3 initial, mole/liter × 10^3	O_2 initial, mole/liter × 10^3	Product yields based on ΔO_3, %		
		CH_3OH	CH_2O	H_2
0.50	3.9	67	10.7	6.5
0.67	2.8	68	11.3	6.7
0.86	2.2	67	11.9	6.6
0.69	1.2	69	11.0	6.4
0.80	0.07	58	9.1	6.8

c. Effect of added CO. The effects of O_2 on the product yields and the quantum yields of O_3 disappearance (discussed later) suggested the presence of scavengable radicals, of which one possibility would be CH_2. The presence of the latter species can be detected by its reaction with CO to form ketene (Ref. 29). For this purpose an experiment was carried out with CO added at a concentration equal to that of methane (0.4 mole/liter). No trace of ketene could be found among the products, thus arguing against the presence of CH_2. The product composition also supports this view, as discussed later.

d. Quantum yields of O_3 photodecomposition. The initial quantum yield of O_3 photodecomposition at CH_4 concentrations of 0.4 mole/liter was found to be 0.12, and a decrease from this limiting value was observed with increasing extent of photolysis. These results suggest a scavenging effect of O_2, which is a by-product of the O_3 photolysis. The latter assumption was confirmed by the observation that initial quantum yields were decreased by addition of O_2 to the mixture, and also is consistent with the effect of added O_2 in raising the product yields. Fig. 17 shows the observed dependence of Φ_{O_3} on the ratio O_3/O_2. The lower limiting value of Φ_{O_3} at $O_2 \gg O_3$ is 0.065.

Experiments were also carried out on the photolysis of O_3 dissolved in nearly pure CH_4 at 90°K, using liquid O_2 as the coolant. A small proportion (about 5–7% by volume) of argon was added to lower the freezing point of the mixture, since CH_4 freezes near 90°K. The chemical effects of this amount of argon can be neglected in view of the rapid reaction of $O(^1D)$ with CH_4. Under conditions of complete O_2 scavenging ($O_2 \gg O_3$), Φ_{O_3} was found to be 0.7. In the absence of O_2 scavenging, Φ_{O_3} was 1.3. The product yields of CH_3OH and CH_2O in the scavenged and unscavenged cases were quite similar to those in the corresponding experiments in argon solvent. The analytical method for H_2 did not permit determination of H_2 in the methane solvent.

To further examine the effect of solvent on the reaction, a few experiments were carried out in a CF_4 solvent. As

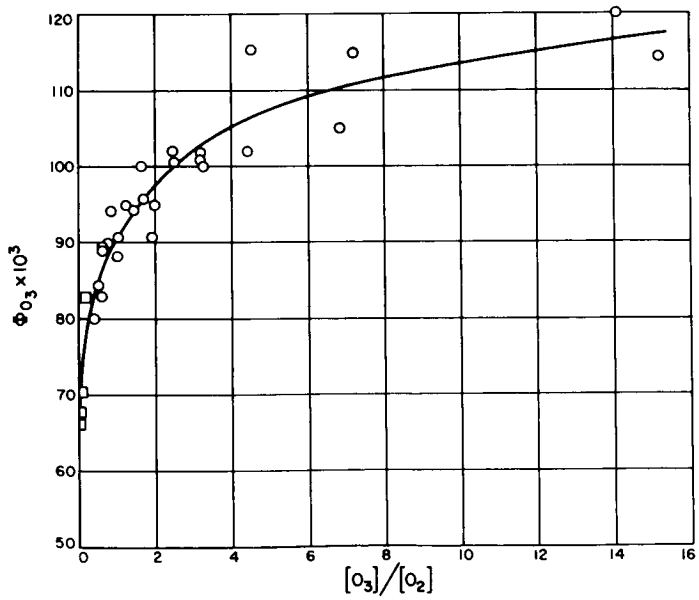


Fig. 17. The quantum yield for O_3 disappearance as a function of the O_3/O_2 ratio

in the cases of Ar and CH_4 solvents, the initial Φ_{O_3} was approximately halved by the presence of excess O_2 in the reaction mixture.

e. Experiments with labeled oxygen. Thermal isotopic exchange between O_3 and O_2^{36} does not occur in low-temperature solvents (Ref. 30), unlike the gas phase where the exchange is quite rapid. It was therefore possible to use O_2^{36} in order to determine the origin of oxygen in the products. The experiment was particularly applicable to CH_3OH since that product was formed in high yield and $CH_3O^{18}H$ can be readily detected in the presence of CH_3OH by mass spectroscopy. Appearance of labeled oxygen in the products could arise by two paths: first, by direct participation of O_2^{36} in product formation, and, secondly, by a photo-exchange between O_3 and O_2^{36}. Photo-exchange would occur whenever $O(^3P)$ appeared in the system, since this latter species has no fate other than recombination with O_2^{36}. On the other hand $O(^1D)$ reacts essentially entirely with CH_4, since $[CH_4] \gg [O_2^{36}]$, and therefore $O(^1D)$ does not lead to exchange except insofar as the interaction with CH_4 is itself a deactivation process. The quantum yield of O_3 disappearance under these conditions (Part 3d, above) is a measure of the extent to which such deactivation occurs, since the initial quantum yield of $O(^1D)$ formation in the O_3 photolysis step is known from previous work (Ref. 30) to be unity. Therefore, by combining the quantum yield results with the results pertaining to labeling in the products, the extent to which the products derive oxygen from O_3 and O_2 can be determined.

First the extent of product labeling arising from photo-exchange must be calculated. It can be shown that the average ratio of labeled ozone to total ozone during a photolysis time t is given by:

$$< O_3^L/O_3^t > = \left(\frac{f}{f'}\right)^2 \left[1 + \frac{f O_3^0 \left[(O_3^t/O_3^0)^{1/f} - 1\right]}{(O_3^0 - O_3^t)}\right] \quad (1)$$

where f is the fraction of $O(^1D)$ which is deactivated and f' is the fraction which reacts to form products. The symbol O_3^0 represents the initial O_3 concentration and O_3^t represents the total O_3 remaining after time t. Assuming that photodissociation of O^{18}–O–O gives O^{18} 50% of the time, then the expected fraction of $CH_3O^{18}H$ resulting from photo-exchange is $\frac{1}{2} < O_3^L/O_3^t >$.

Results from the exchange experiments are shown below and are compared with the calculated results based on Eq. (1), taking $f = 0.3$ and $f' = 0.7$ as indicated by the quantum yield data from Part 3d.

O_3^0, mole/liter $\times 10^3$	O_2^{36}, mole/liter $\times 10^3$	Decomposition of O_3, %	Labeling of CH_3OH, %	
			Observed	Calculated
0.48	25	42	3.4	3.6
0.39	25	67	5.5	5.2

The agreement between observation and calculation is very good and allows two important conclusions. First, the correctness of the deactivation fraction $f = 0.3$ is strongly supported since the fraction of product labeling is very sensitive to f. Second, the results show that CH_3OH derives essentially all of its oxygen from O_3 and not from O_2. The experimental error, however, allows for the possible presence of a few percent of labeled CH_3OH in excess of that arising from photo-exchange.

4. Discussion

a. The initial reaction of $O(^1D)$ with CH_4. The observed products and the scavenging effect of added O_2 indicate that the initial reaction follows three paths as shown below, where we denote the fractional occurrence of each by α, β, and γ.

$$O(^1D) + CH_4 \xrightarrow{\alpha} CH_3OH \qquad \Delta H = -135 \text{ kcal/mole} \quad (2)$$

$$O(^1D) + CH_4 \xrightarrow{\beta} CH_2O + H_2 \qquad \Delta H = -115 \text{ kcal/mole} \quad (3)$$

$$O(^1D) + CH_4 \xrightarrow{\gamma} CH_3 + OH \qquad \Delta H = -37 \text{ kcal/mole} \quad (4)$$

Estimation of α, β, and γ is deferred until the question of radical oxidation has been discussed.

The reaction

$$O(^1D) + CH_4 \rightarrow CH_2 + H_2O \qquad \Delta H \simeq -60 \text{ kcal/mole} \quad (5)$$

has sometimes been postulated as a primary step in the reaction of O with CH_4 but is excluded on the grounds that no CH_2CO could be found as a product with added CO. Also, the products show no significant excess of CH_2O which would be expected to arise from the reaction $CH_2 + O_3 \rightarrow CH_2O + O_2$. Another possible reaction,

$$O(^1D) + CH_4 \rightarrow CH_2O + 2H \qquad \Delta H = -11 \text{ kcal/mole} \quad (6)$$

is also rejected on the grounds that there is no evidence that it occurs.

Reactions (2)–(4) probably involve the common intermediate CH_3OH^* formed by insertion of $O(^1D)$ into a C–H bond of CH_4:

$$O(^1D) + CH_4 \rightarrow CH_3OH^* \text{ (excitation energy = 135 kcal/mole)} \quad (7)$$

Excited CH_3OH^* would either be collisionally deactivated or otherwise undergo thermal decomposition according to Reactions (3) and (4). The latter two reactions are the predicted modes for decomposition of CH_3OH^* on the basis of the Rice-Ramsperger-Kassel theory of unimolecular reactions. Unfortunately, the various third bodies (Ar, CH_4, CF_4) used in the present experiments show no differing efficiencies which could be taken as proof of the role of CH_3OH^* as a reaction intermediate.

A fourth reaction which must be included is deactivation of $O(^1D)$:

$$O(^1D) + CH_4 \rightarrow O(^3P) + CH_4 \quad (8)$$

Both the quantum yield data and the product labeling experiments indicate that 30% of the $O(^1D)$ produced undergoes deactivation by CH_4. It is not clear at the present time whether deactivation occurs by a chemical mechanism involving CH_3OH^* or whether the deactivation is physical in nature, arising from solvent effects.

Occurrence of the reaction of $O(^1D)$ with CH_4 at 87°K is a strong indication that the activation energy is zero. This suggestion is supported by comparison with the reaction of $O(^1D)$ with O_3, which was studied under very similar conditions (Ref. 30). The ratio of rate constants k_{CH_4}/k_{O_3} is given by

$$k_{CH_4}/k_{O_3} = \frac{\Phi_{CH_4}(n_{O_3} - \Phi_{O_3})[O_3]}{\Phi_{O_3}(n_{CH_4} - \Phi_{CH_4})[CH_4]} \quad (9)$$

where n_{O_3} is the net number of O_3 molecules decomposed each time the reaction occurs, and Φ_{O_3} is the observed quantum yield of O_3 photolysis. Similar definitions hold for n_{CH_4} and Φ_{CH_4}. For the $O(^1D) - CH_4$ reaction we take $n = 1$ although the actual value may be 0.7 if deactivation does in fact occur by a chemical mechanism, as discussed above. In the case of the reaction $O(^1D) + O_3 \rightarrow 2O_2$ the correct value of n is 2. For $[O_3] = [CH_4] = 0.4$ moles/liter, Φ_{O_3} was found to be 0.4 in previous work (Ref. 30), and from the present work $\Phi_{CH_4} = 0.065$. The latter value is the quantum yield with complete O_2 scavenging since that quantity represents the O_3 decomposed in the initial step, as shown in Part 4b. From this data we find

$$k_{CH_4}/k_{O_3} = 0.3 \quad (10)$$

The reaction of $O(^1D)$ with CH_4 therefore occurs with a rate comparable to the reaction of $O(^1D)$ with O_3. Since the latter reaction was found to have $E_a < 0.5$ kcal/mole, it follows that the methane reaction almost certainly has a similarly low E_a, since rates are extremely sensitive to activation energy at the temperature of these experiments.

b. Radical oxidation. The effect of added O_2 on the O_3 quantum yield and the product yields, as shown in Fig. 18 and Tables 15 and 16, is attributed to scavenging of CH_3, which otherwise reacts with O_3:

$$CH_3 + O_3 \rightarrow CH_3O + O_2 \quad \Delta H = -37 \text{ kcal/mole} \quad (11)$$

$$CH_3 + O_2 \rightarrow CH_3O_2 \quad \Delta H = -26 \text{ kcal/mole} \quad (12)$$

Other possible radical oxidation reactions are

$$CH_3O + O_3 \rightarrow CH_3O_2 + O_2 \quad \Delta H = -28 \text{ kcal/mole} \quad (13)$$

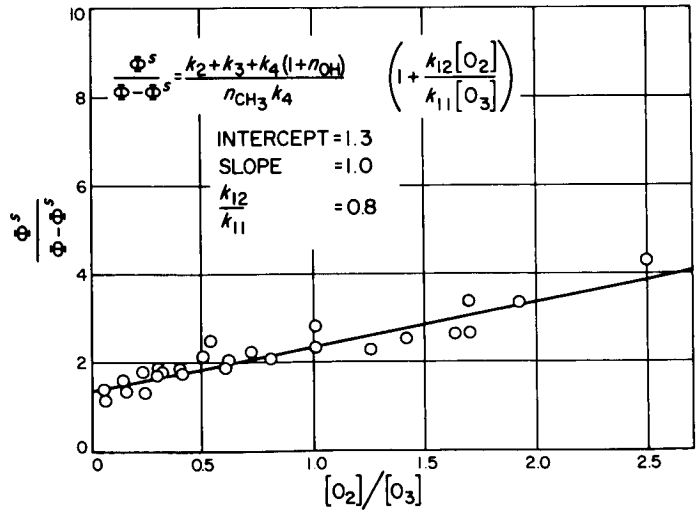


Fig. 18. The effect of added O_2 on the quantum yield for O_3 disappearance

$$OH + O_3 \rightarrow HO_2 + O_2 \qquad \Delta H = -37 \text{ kcal/mole} \tag{14}$$

The scavenging efficiency of O_2 as expressed by the rate ratio k_{12}/k_{11} can be correlated with the scavenging data by means of the following expression:

$$\frac{\Phi^s}{\Phi - \Phi^s} = \frac{k_2 + k_3 + k_4(1 + n_{OH})}{n_{CH_3} k_4}\left(1 + \frac{k_{12}[O_2]}{k_{11}[O_3]}\right) \tag{15}$$

where Φ^s represents the O_3 quantum yield in the limit of complete scavenging, and n_{OH} and n_{CH_3} represent the number of O_3 molecules destroyed by OH and CH_3, respectively. The scavenging data in Fig. 18 are plotted according to Eq. (15). From the slope we find $k_{12}/k_{11} = 0.8$ so that CH_3 reacts with O_3 and with O_2 at essentially equal rates.

The intercept in Fig. 18 is related to the quantity γ, which is defined by Reaction (4), but is also a function of the ratio n_{OH}/n_{CH_3}:

$$\frac{k_2 + k_3 + k_4(1 + n_{OH})}{n_{CH_3} k_4} = \frac{1}{\gamma} + \frac{n_{OH}}{n_{CH_3}} \tag{16}$$

It is of interest to determine n_{OH} and n_{CH_3} since these quantities indicate whether or not Reactions (13) and (14) occur. The quantity n_{OH} equals 1 if Reaction (14) occurs and zero if it does not; the quantity n_{CH_3} has a value of at least 1 because of the known Reaction (11) and equals 2 if a second O_3 is consumed by Reaction (13). The primary experimental observations which must be explained by the radical oxidation mechanism are that Φ (unscavenged)/Φ (scavenged) = 1.77 and that at the same time the CH_3OH yield (scavenged) is as high as 0.68. The keys to the quantities α, β, and γ are the CH_3OH yield, the H_2 yield, and the scavenging effect, respectively. In addition, the relation of $\alpha + \beta + \gamma = 1$ must hold. The values for n_{OH} and n_{CH_3} which best fit the experimental results can be determined by a case-by-case comparison of prediction with observation, as shown in the table below.

Inspection of this table shows that Cases 1 and 2 are completely inadequate to explain the CH_3OH yields. Possible routes for CH_3OH formation through radical-radical reactions such as the following can be ruled out:

$$CH_3O_2 + CH_3O_2 \rightarrow CH_3OH + CH_2O + O_2 \tag{17}$$

$$CH_3O + CH_3O \rightarrow CH_3OH + CH_2O \tag{18}$$

In each case the corresponding yield of CH_2O is not found, and in addition Reaction (17) would produce labeled CH_3OH in the O_2^{36} experiments, contrary to observation.

Case 3 accounts for the O_3 quantum yield ratios and also predicts a CH_3OH yield of 0.55 which is not too much lower than the observed 0.68. Furthermore, additional CH_3OH formation is allowed by Reaction (19),

$$CH_3O_2 + OH \rightarrow CH_3OH + O_2 \tag{19}$$

since in this reaction CH_3OH would not be labeled in the presence of O_2^{36}. Case 3 also has the advantage that the same radical species are present in both the scavenged and unscavenged experiments, which is in agreement with the observation that added O_2 has little effect on the qualitative product distribution, but changes product yields by an amount which would be predicted on the basis of the quantum yield ratios and the assumption of identical sources of CH_3OH in each experiment.

The quantum yield (scavenged) of O_3 photolysis in CH_4 solvent provides independent evidence that $n_{OH} = 0$. If Reaction (14) occurred, the observed quantum yield of 0.7 would include O_3 loss due to that reaction. This would imply that $f > 0.3$, which is contrary to the results of the product labeling experiments.

The reaction $OH + O_3 \rightarrow HO_2 + O_2$ [Reaction (14)] occurs quite rapidly in the gas phase and is responsible for high quantum yields in the photolysis of wet ozone (Ref. 31). Since the present work indicates that the reaction is slow at 87°K, we propose that there is an activation energy of 1–2 kcal/mole for Reaction (14). It is of interest to note that Reaction (13), which is somewhat analogous to Reaction (14), apparently has a zero activation energy even though it is less exothermic than Reaction (14).

Case	Assumed value		Calculated value			Scavenged CH_3OH yield		
	n_{OH}	n_{CH_3}	α	β	γ	Equal to	Calculated	Observed
1	1	2	0.27	0.11	0.62	$\alpha/1+\gamma$	0.17	0.68
2	0	1	0.16	0.07	0.77	α	0.16	0.68
3	0	2	0.55	0.07	0.38	α	0.55	0.68

E. Atomic Integrals with the Magnetic Dipole–Dipole Operator

M. Geller

In a recent note (Ref. 32), Roberts has obtained the atomic integrals with the magnetic dipole–dipole operator

$$r_{12}^{-3} P_2(\cos\theta_{12}) = r_{12}^{-5} (3z_{12}^2 - r_{12}^2)/2 \tag{1}$$

in the form of a finite sum of integrals over hypergeometric functions (which are themselves infinite series). These integrals correspond to the one-center integrals needed in the evaluation of zero-field splitting. Geller and Griffith (Ref. 33) have previously obtained the two-center Coulomb-type, zero-field splitting integrals by the Fourier convolution method (Refs. 34 and 35) in the form of a finite sum of integrals for which analytic closed form expressions were given. In the simpler one-center (atomic) case, equivalent to that treated by Roberts, the method of Ref. 33 can easily be applied and yields the following results.

For the one-center, two-electron integral[16]

$$I = \int [NLM]_1 O_1 [N'L'M']_2 \, d\tau_1 \, d\tau_2 \tag{2}$$

where the $[NLM]$ are the Roothaan charge distributions and O_1 is the operator defined above, the integral is given by[17]

$$I = (-1)^{L+M} \delta_{M,M'} \sum_{s=0}^{[\frac{1}{2}(N-L)]} M_{NL}(s,p_a) \sum_{t=0}^{[\frac{1}{2}(N-L)]} M_{N'L'}(t,p_b) \times G_{NL}^{N'L'}(s,t;p_a,p_b) (C_1 \delta_{L,L'-2} + C_2 \delta_{L,L'}) \tag{3}$$

where

$$M_{NL}(s,p_a) = (-1)^s 2^{2L} (s+1)_L [(2L)!]^{-1} p_a^{2N-L-2S+2} \times \left(\frac{N+L+1}{2s+2L+1}\right) \left[\frac{N+L+1}{2L+1}\right]^{-1}$$

$$C_1 = (L+M+2)(L+M+1)(L-M+2)(L-M+1)^{1/2}/D$$

$$D = (L+L'+3)(L+L'+1)(L+L'-1),$$

$$C_2 = -\frac{2}{3}(L(L+1) - 3M^2)/D \tag{4}$$

and

$$G_{NL}^{N'L'}(s,t;p_a p_b) = \frac{1}{\pi} \int_0^\infty \frac{k^{2s+2t+L+L'+2} dk}{(k^2+p_a^2)^{N+1}(k^2+p_b^2)^{N+1}} \tag{5}$$

This latter integral G has been evaluated in closed, analytical form by Lounsbury (Ref. 36) and is readily amenable to automatic machine coding as it involves a sum of algebraic terms from 0 to N plus a second summation ranging from 0 to N'.

[16] For details of the rotation and the application of the method see Ref. 35(a).

[17] The integral is arranged so that $L \leq L'$.

F. A Micro Gas Chromatograph for Descent Analysis of the Martian Atmosphere

W. F. Wilhite

1. Introduction

As part of the proposed Mars 1966 capsule mission, rapid analysis of the Martian atmosphere during the descent of the capsule to the surface is desired.

Since the primary use of this descent analysis information would be more accurate design of later capsule missions, it would be necessary to determine only the major components of the Martian atmosphere. However, the allotted time for atmospheric analysis during descent could be as short as 1 min; this constraint, and a requirement for several analyses, would limit the analysis time to less than 10 sec.

This limitation, plus the fact that the scientific payload could be very limited in weight, dictated a very lightweight gas analysis instrument capable of rapid analysis of the major components of the Martian atmosphere.

A gas chromatographic system has been developed at JPL to meet these limitations. The goal of this development was a gas chromatograph weighing less than 3 lb, capable of performing the analysis in a few seconds.

Rapid analysis with gas chromatography has been achieved many times (Refs. 37 and 38) in the brief history of this technique. However, rapid analysis heretofore has usually been accomplished with very high flow rates of

carrier gas. The requirement for carrying a large quantity of carrier gas to achieve high flow rates would make design of a lightweight chromatograph unfeasible. The goal of this development was rapid analysis capability using a total carrier gas flow rate of less than 1 cm^3/min. This requirement has necessitated evolving new techniques in gas chromatography component development. Although these new techniques were aimed specifically at rapid analysis of the Martian atmosphere, they are equally applicable to other space-borne gas chromatographs.

2. Component Development

a. Columns. The two requirements of low flow rate of carrier gas and rapid analysis dictated the use of small-bore columns. Packed columns, rather than classical capillaries, were selected for this particular mission since the atmospheric pressure of Mars at present is not accurately known, and therefore the amount of sample loading on the column could vary over a wide range. Stainless steel tubing of 0.25-mm inner diameter was used for the columns. This tubing diameter limited the usable particle size of the packing to a range of 10 to 80 μ. Since narrow cuts were not commercially available in this particle size range, a sieving system was assembled (Fig. 19) to provide such material.

This sieving system uses electroformed nickel screens with opening tolerances of $\pm$ 1 μ, manufactured by Buckbee Mears (Ref. 39), St. Paul, Minnesota. The ranges of opening sizes used are 25 to 40 μ in 3-μ steps, 40 to 60 μ in 5-μ steps, and 60 to 80 μ in 10-μ steps. The screens are inserted in adapters that are placed on an Alpine air jet sieve, manufactured by Alpine of Augsburg, West Germany. Five to 10 cm^3 of the material to be sieved is placed on the screen, and the screen is then covered with an airtight plastic lid. Directly under the screen, an arm with a radial slot on top rotates at 40 rpm. An industrial vacuum cleaner is used to create a vacuum in the chamber in which the rotator turns. The slot in the rotating arm is open to outside air through its hollow axle, and, because of the vacuum, air is pulled up through this slot, lifting the particles off the top of the screen. As the rising jet of air moves from under the particles, they are pulled down through the screen by the vacuum. This action removes agglomerates that may have formed and continuously prevents the sieve from being clogged. The fine particles that pass through the screen are collected in a large filter mounted under the sieve.

This system provides very uniform particles of small size using a dry sieving technique (Refs. 40 and 41). One added advantage is that the system removes fines, which usually adhere to the particles of commercially sieved material. These fines are undesirable, as they cause increased pressure drop across a packed column. All the material sieved with this system was provided by Analabs (Hamden, Conn.) from their waste material, which is smaller than 170 mesh.

Finding a plug to hold the packing material in the micro packed columns presented another problem. The plugs had to be porous and able to withstand a column activation temperature of 400°C. Success was achieved with 0.25-mm-outer-diameter plugs electronically machined from 400-series magnetic sintered stainless steel 0.8 mm thick, with a pore size of 10 $\pm$ 5 μ. The plugs are cone-shaped (Fig. 20) because of the nature of the electronic machining process. Through the use of a stereomicroscope, these plugs are inserted into the end of the column with tweezers and are wedged into place with a soft brass tool.

Proper preparation of the tubing is an important step in making a micro packed column. First of all, the proper length of 1.6-mm-outer-diameter, 0.25-mm-inner-diameter tubing is cut. The ends of the tubing are machined in a jeweler's lathe to provide a perfectly flat surface perpendicular to the axis of the column, so that no dead space occurs between the end of the column and the tubing fitting. Next, each end of the tubing is made up using a special Swagelok gas chromatography low dead-volume fitting. The procedure for packing these micro packed columns primarily involves performing the normal column packing operations under a stereomicroscope. The lower end of the column is attached to a vacuum line and supported by a micromanipulator under the stereomicroscope. The upper end of the column is then fitted with a short length of 1.6-mm-inner-diameter Teflon tubing for use as a funnel. The column is then carefully packed using a standard hand vibrator and a vibrating spatula with a small "v" notch in the tip to feed the packing material. Both of these devices are mounted rather than held by hand, and are activated or deactivated by a foot switch. After the column has been packed, the open end is plugged with the stainless steel plug. The column is then turned upside down, the vacuum line is removed, and the other end is plugged.

b. Detector. A dual-cell micro thermal conductivity detector has also been developed for use with the micro packed columns. The thermal conductivity method of detection was chosen because of the relative simplicity with which a hot-wire filament detector of small internal

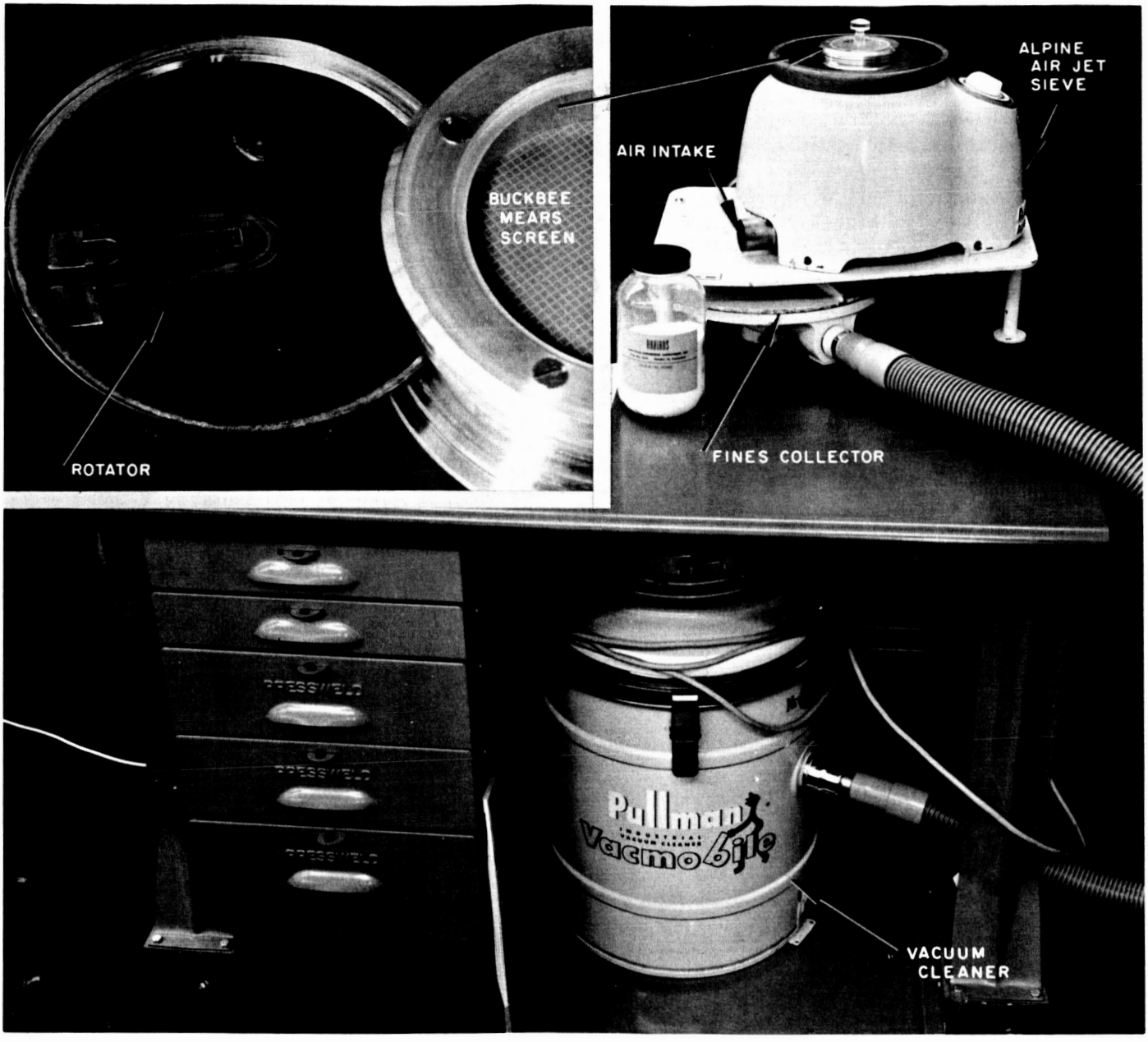


Fig. 19. Alpine air jet sieving system

volume could be built. A cross-sectional drawing of the micro thermal conductivity detector is shown in Fig. 21. The detector is made of 303 stainless steel and has an internal volume of 0.1 μl. One of the special Swagelok gas chromatography fittings was machined and silver-soldered into the detector block to provide a minimum dead volume between the end of the column and the detector.

The hot-wire filament is soft-soldered between two 0.13-mm-diameter platinum wires that protrude out of the detector probe. The length of the filament is 0.3 mm and the diameter is 0.5 μ. The filament wire is 10% rhodium-platinum and is manufactured by the Sigmund Cohn Corporation, Mt. Vernon, N.Y., using the Wollaston process. (This process involves inserting the wire to be drawn in a tube of silver and then drawing the incased wire down to a calculated diameter.) When the rhodium-platinum wire is to be used, the silver coating is dissolved by nitric acid. The small length of bared wire is placed across the two platinum needles and soft-soldered into

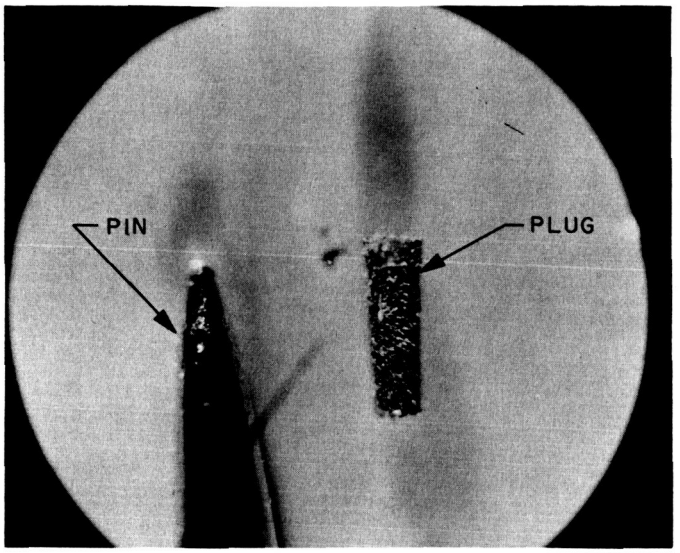


Fig. 20. Photomicrograph of sintered stainless steel plug compared with straight pin for size

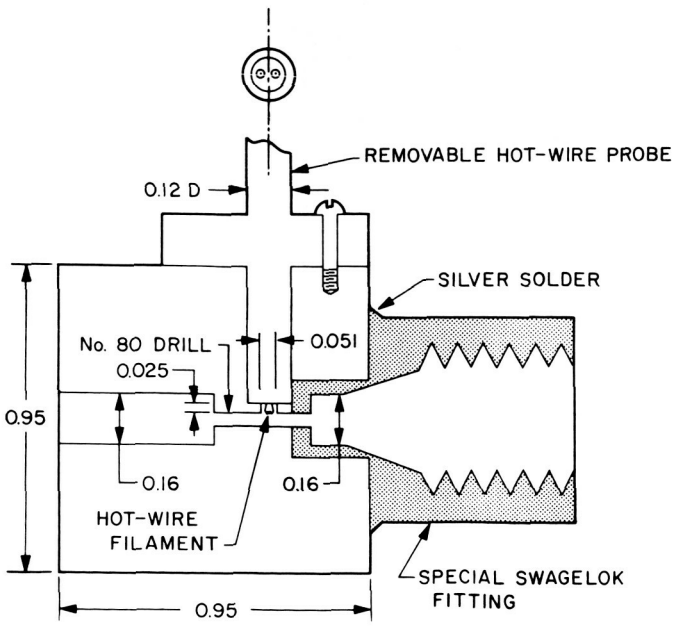


Fig. 21. Cross-sectional drawing of micro thermal conductivity detector

place. The hot-wire filament is then annealed to a dull red glow.

During operation, the filament current is usually about 3.5 ma. The resistance of the wire at this current averages 400 Ω to give a filament power dissipation of 5 mw. This power consumption is several orders of magnitude below that of commercial filaments and offers a distinct advantage for spacecraft missions that require low power consumption.

c. Sampling valve. A sampling valve that requires very little force to actuate and is capable of injecting microliter quantities of gas was also developed. The valve, as shown in Fig. 22, is a sliding valve using the shear-seal O-ring sealing principle. The O-ring seal was used so that a minimum amount of force would be required to actuate the valve. For laboratory use, the valve is actuated by opposing pneumatic bellows that are pressure-pulsed by a four-way solenoid valve. The time required for movement of the valve in either direction is 25 msec. The valve is provided with several interchangeable sliders of the same overall dimensions but with various-sized O-rings and passages through the slider to provide sampling volumes from 0.5 to 10 μl. The valve was tested for leaks and showed no leakage of helium up to a pressure of 100 psi.

These developed microcomponents were assembled into a system to perform rapid analysis of components believed to compose the Martian atmosphere. It has been determined by astronomical observation that carbon dioxide is a major component in the Martian atmosphere. Two other major components are thought to be argon and nitrogen.

Because argon, nitrogen, and carbon dioxide cannot be separated easily on one column in gas chromatography, two columns and detectors were assembled in series as follows:

(1) A silica gel column for separating carbon dioxide from the composite (argon-nitrogen peak).

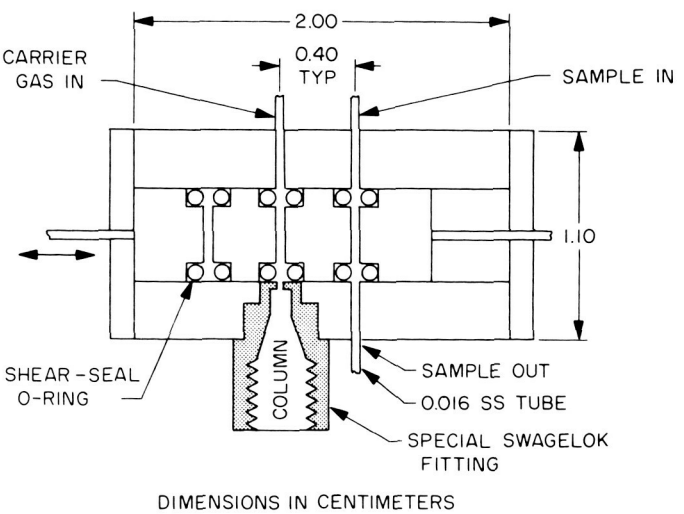


Fig. 22. Cross-sectional drawing of sampling valve

(2) A detector for detecting the composite and carbon dioxide peaks.

(3) A molecular sieve 5A column for separating argon and nitrogen.

(4) A detector for detecting argon and nitrogen.

Separation of carbon dioxide from air (composite) is rather fundamental in gas chromatography. However, to perform the separation rapidly required resolution with very sharp, narrow peaks. These necessary characteristics are not usually achieved by using silica gel as a separation material, since the carbon dioxide peak becomes extremely broad, even though easily separated.

To overcome this difficulty we tried an unactivated silica gel column to achieve a narrow carbon dioxide peak without losing the desired resolution. This proved to be initially successful, but over a period of time the silica gel column would slowly activate as it was purged with the dry helium. The carbon dioxide peak would then broaden and increase in retention time, slowing the overall analysis. Success was finally accomplished by activating 25- to 28-μ silica gel for 1 hr at 250°C and then deactivating the silica gel with a low loading of a polar substrate. The best results were obtained with 1% diglycerol. Using this preparation, a silica gel micro packed column 2.5 cm long was required for the resolution of the composite and carbon dioxide peaks in 1.5 sec.

To prepare the molecular sieve 5A column for separating argon and nitrogen, the 25- to 28-μ molecular sieve was activated in the column for 3 hr at 400°C. A column 3 2 cm long was used for resolution of these two components.

The two column-detector combinations were then assembled into a series system (Fig. 23). A delay line 16.5 cm long was needed between the first detector and the molecular sieve column to delay detection of argon until after the carbon dioxide peak was detected. Dummy columns paralleling the actual columns were connected to the reference cells of the detectors to ensure identical operating conditions in both cells of each detector. An electronic system combining two detector bridges and a low-noise amplifier into one package was designed and built to process the detector signals. The output from this system was fed to an oscilloscope for readout.

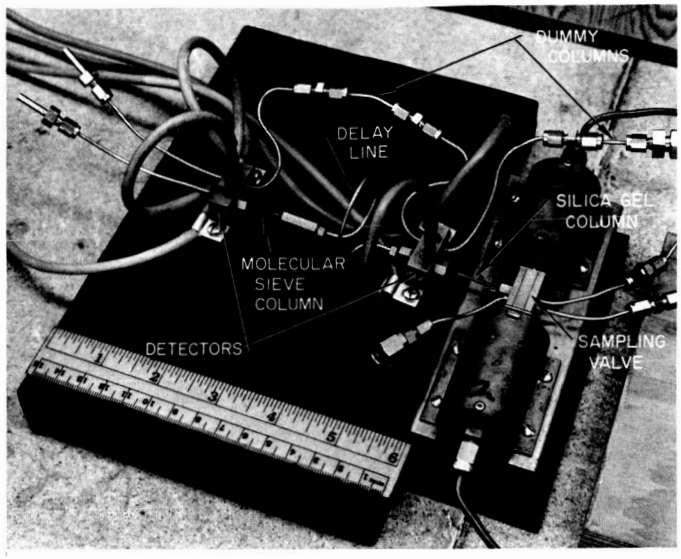


Fig. 23. Series column-detector system

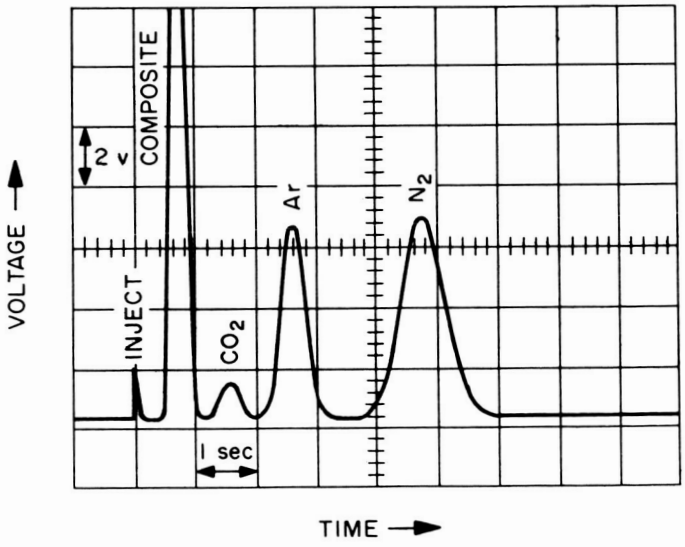


Fig. 24. Chromatogram reproduced from oscilloscope trace: Column 1: 2.5 cm, 0.25-mm ID, 1% diglycerol on 25 to 28-μ silica gel; Column 2: 3.2 cm, 0.25-mm ID, 25 to 28-μ molecular sieve 5A; helium carrier gas

3. *Discussion of Results*

a. System performance. Fig. 24 shows the series column-detector system chromatogram as displayed on the oscilloscope. Each division represents 1 sec on the horizontal coordinate and 2 v on the vertical coordinate. As shown, carbon dioxide, argon, and nitrogen were resolved in 6 sec. The analysis could be performed in even less time if shorter columns, packed with smaller particles, were used. The analysis was made at room temperature with an inlet pressure of 33 psi. The flow rate of the system was 0.5 cm^3/min, and the sample size was 1 μl.

b. Column performance. To check the performance of a micro packed column, a series of runs was made using a column 3.8 cm in length, 0.25-mm inner diameter, packed with 25- to 28-μ molecular sieve 5A. The column was activated at 400°C for 3 hr. Several runs at room temperature and at different flow rates of helium carrier were made using a 1-μl sample of 20% argon and 80% nitrogen mixture. The height of equivalent theoretical plate (HETP) and average velocity were calculated for each run. A plot of these data (Fig. 25) shows a continuing gradual improvement in efficiency as velocity increases. The data demonstrate that the micro packed column is capable of high efficiency at low velocities and correspondingly low flow rates. This is one of the most desirable characteristics of the micro packed column since the primary purpose of this work was to achieve an efficient gas chromatographic system at very low flow rates.

It is relevant to compare the performance of micro packed columns with the work reported by Halász and Heine in which a classical packed column and packed capillary columns are compared. Halász and Heine (Ref. 42) used alumina as the column packing material to compare a classical packed column with 2- and 10-m packed capillary columns under similar operating conditions. (It is felt that a valid comparison can be made between their work and the results obtained with the micro packed columns, even though the packing material and operating conditions were somewhat different in the two investigations.) The efficiency of the micro packed column compares favorably with the shorter packed capillary column and, at higher velocities, could approach the efficiency of the longer packed capillary, as shown by the trend of the curve of the HETP vs velocity (Fig. 25).

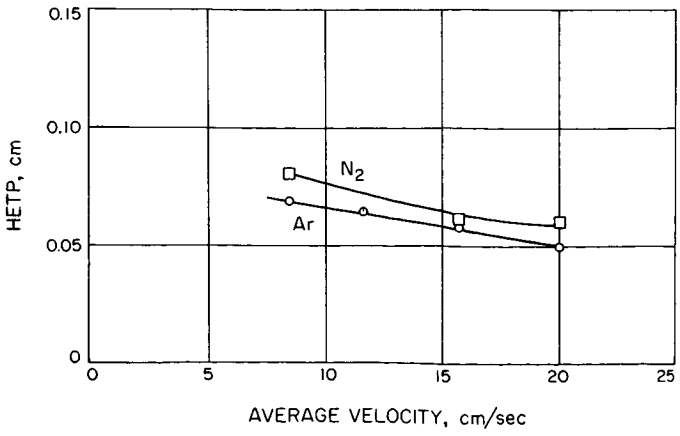


Fig. 25. HETP vs average velocity curve for micro packed column

The micro packed columns have the added advantage of being able to handle much larger samples than the packed capillary column without overloading.

c. Detector performance. The performance of the micro thermal conductivity detector was tested using the mixing-chamber exponential decay method described by Lovelock (Ref. 43). In the present study, the mixing chamber was a Teflon vessel with an internal volume of 89 ml, containing a magnetically-driven stirrer. The carrier gas was passed through a three-way valve into the mixing chamber. Into the third port of the three-way valve, argon was fed so that the input to the mixing chamber could be switched between the helium gas and argon. Part of the mixing chamber effluent was passed through the detector, with the rest bypassing the detector. The carrier gas flow rate was set at approximately 0.5 cm³/min through the detector and 9.5 cm³/min through the bypass. By switching the three-way valve, the carrier gas was shut off, and argon, at a much higher flow rate, was introduced into the mixing chamber. When the detector output, as read on a strip-chart recorder, leveled off, showing that the gas flowing through the detector was 100% argon, the three-way valve was switched back to the helium carrier at the preset flow rates. Time zero was indicated on the recorder chart, and the detected decay of argon concentration in helium was recorded. Fig. 26 shows a plot of this concentration decay vs the detector output in millivolts. A rough indication of the sensitivity of this

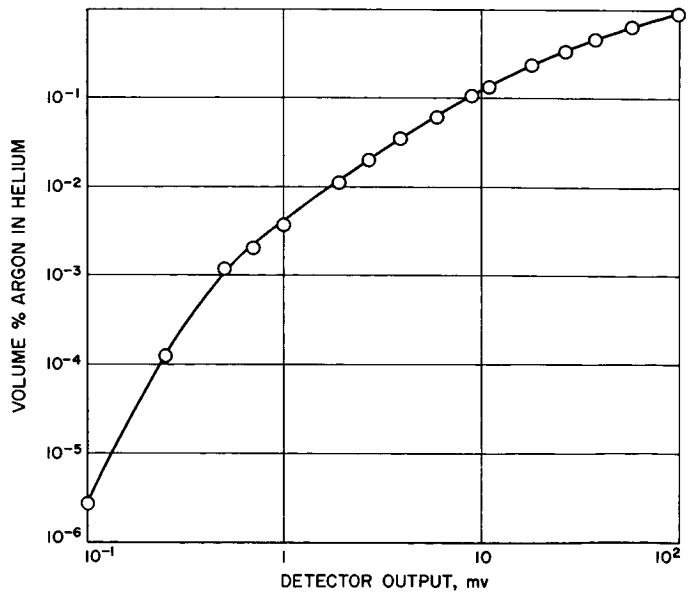


Fig. 26. Response curve for micro thermal conductivity detector

detector at a 2-to-1 signal-to-noise ratio was obtained at less than 5 parts per million argon in helium.

4. Concluding Remarks

Future work in this area of micro gas chromatography will involve two new aspects. One of these will be testing the developed micro components for capability of surviving a shock of 10,000 *g*. This development will aim at designing a gas chromatograph capable of surviving a hard impact landing on the surface of Mars.

The other area of development will involve applying the techniques described in this article to analysis of highly polar organic compounds. This work will aim at developing a gas chromatograph for inclusion in a pyrolysis/gas chromatograph/mass spectrometer system for detection of life-related compounds. The low flow rate of the micro components is ideally suited for the combination gas chromatograph–mass spectrometer because a space-borne mass spectrometer must rely on chemical and ion pumping for vacuum. Since these two pumping techniques are relatively low in capacity, they are limited to handling small quantities of carrier gas.

Acknowledgment. The author expresses his sincere appreciation to J. J. Norris, for laboratory assistance, and to G. E. Tennant, who built the micro detector and the sampling valve.

References

1. Kodama, K., *Methods of Quantitative Inorganic Analysis,* Table 19, p. 64, Interscience Publishers, New York, N.Y., 1963.
2. Sanders, K. H., *The Aromatic Diazo-Compounds and Their Technical Applications,* pp. 209 et seq, Edward Arnold & Co., London, England, Second Ed., 1949.
3. Sandell, E. B., *Colorimetric Determination of Traces of Metals,* pp. 283–289, Interscience Publishers, Inc., New York, N.Y., 1944.
4. Schmuckler, G., "Chelating Resins—Their Analytical Properties and Applications," *Talanta,* Vol. 12, Part 3, pp. 281–290, 1965.
5. Carritt, D. E., "*Separation and Concentration of Trace Metals from Natural Waters,*" *Analytical Chemistry,* Vol. 25, pp. 1927–1928, 1953.
6. Goya, H., et al, *A Method for the Collecton, Storage and Monitoring Analysis of Natural Sea Water,* USNEDL-TR-637, U. S. Naval Rediological Defense Laboratory, San Francisco, California, 1963.
7. Shaw, M. D., and Arble, W. C., *Bibliography on Methods for Determining Soil Moisture,* Engineering Research Bulletin B-78, Pennsylvania State University, University Park, Pa., 152 pp., 1959.
8. Bloodworth, M. E., and Page, J. B., "Use of Thermistors for the Measurement of Soil Moisture and Temperature," *Soil Science Society of America Proceedings,* Vol. 21, pp. 11–15, 1957.
9. Shaw, B. T., and Bauer, L. D., "An Electrothermal Method for Following Moisture Changes of Soil *In Situ,*" *Soil Science Society of America Proceedings,* Vol. 4, pp. 78–83, 1940.
10. Kelley, O. J., "A Rapid Method of Calibrating Various Instruments for Measuring Soil Moisture *In Situ*" *Soil Science,* Vol. 58, pp. 433–439, 1944.

References (Cont'd)

11. Richards, L. A., Editor, *Diagnosis and Improvement of Saline and Alkali Soils,* Agriculture Handbook No. 60, pp. 109–110, 1954.

12. Papadakis, J., "Climatic Tables for the World," J. Papadakis Publishers, Buenos Aires, Argentina, 175 pp., 1961.

13. Bouyoucos, G. J., "Hydrometer Method for Making Particle Size Analyses of Soils," *Agronomy Journal,* Vol. 54, pp. 464–465, 1962.

14. Jackson, J. L., "Soil Chemical Analysis," Prentice-Hall, Inc., Englewood Cliffs, N.J., 498 pp., 1958.

15. Chapman, H. D., and Pratt, P. F., "Methods of Analysis for Soils, Plants, and Waters," University of California, Berkeley, California, 309 pp., 1961.

16. Cameron, R. E., and Blank, G. B., *Soil Organic Matter,* Technical Report No. 32-443, Jet Propulsion Laboratory, Pasadena, California, May 23, 1963.

17. Vinogradov, A. P., "The Geochemistry of Rare and Dispersed Chemical Elements in Soils, Second Edition," Translation by Consultants Bureau, Inc., New York, N.Y., 209 pp., 1959.

18. Schwabe, G. H., "Zur autotrophen Vegetation in ariden Böden, Blaualgen und Lebensraum. IV," *Osterreichischen Botanischen Zeitschrift,* Vol. 107, pp. 281–309, 1960.

19. Wong, E. L., and Potter, Jr., A. E., *The Journal of Chemical Physics,* Vol. 39, p. 2211, 1963.

20. Cadle, R. D., and Allen, E. R., *Journal of Physical Chemistry,* Vol. 69, p. 1611, 1965.

21. Avramenko, L. I., Kolesnikova, R. V., and Kuznetsova, N. W., *Izvestiya Akademii Nauk* SSSR, *Otdelenie Khimicheskikh Nauk,* Vol. 620, 1963.

22. Azatyan, V. V., Nalbandyan, A. B., and Tsui, M. Y., *Kinetika i Kataliz,* Vol. 201, 1964; *Doklady Akademii Nauk* SSSR, Vol. 147, p. 361, 1962; *ibid.,* Vol. 36, p. 23, 1963.

23. Elias, L., and Schiff, H. I., *Canadian Journal of Chemistry,* Vol. 38, p. 1657, 1960.

24. Avramenko, L. I., Kolesnikova, R. V., and Savinova, G. I., *Izvestiya Akademii Nauk* SSSR, *Otdelenie Khimicheskikh Nauk,* Vol. 976, 1963.

25. Yamazaki, H., and Cvetanovic, R. J., *The Journal of Chemical Physics,* Vol. 41, p. 3703, 1964.

26. Basco, N., and Norrish, R. G. W., *Canadian Journal of Chemistry,* Vol. 38, 1960.

27. Raper, O. F., and DeMore, W. B., *The Journal of Chemical Physics,* Vol. 40, p. 1047, 1964.

28. *Argon, Helium and Rare Gases,* Vol. II, edited by G. A. Cook, Interscience Publishers, New York, N.Y., p. 525, 1961.

29. DeMore, W. B., and Benson, S. W., *Advances in Photochemistry,* Vol. 2, Interscience Publishers, New York, N.Y., p. 219, 1964.

References (Cont'd)

30. DeMore, W. B., and Raper, O. F., "Primary Processes in Ozone Photolysis," *Space Programs Summary No. 37-31*, Vol. IV, Jet Propulsion Laboratory, Pasadena, California, pp. 225–229, February 28, 1965.

31. McGrath, W. D., and Norrish, R. G. W., *Proceedings of the Royal Society (London)*, Vol. A 242, p. 265, 1957.

32. Roberts, P. J., *The Journal of Chemical Physics*, Vol. 43, p. 765, 1965.

33. Geller, M., and Griffith, R. W., *The Journal of Chemical Physics*, Vol. 40, p. 2309, 1964; and Erratum to this article, *The Journal of Chemical Physics*, Vol. 41, p. 2946, 1964.

34. Prosser, F. P., and Blanchard, C. H., *The Journal of Chemical Physics*, Vol. 36, p. 1112, 1962.

35. (a) Geller, M., *The Journal of Chemical Physics*, Vol. 39, p. 84, 1963; (b) Geller, M., *The Journal of Chemical Physics*, Vol. 39, p. 853, 1963; (c) Geller, M., *The Journal of Chemical Physics*, Vol. 41, p. 4006, 1964.

36. Lounsbury, J. B., *The Journal of Chemical Physics*, Vol. 42, p. 1549, 1965.

37. Farré-Rius, F., and Guiochon, G., *Journal of Gas Chromatography*, Vol. 1, No. 11.

38. Halász, I., and Horvath, C., *Analytical Chemistry*, Vol. 35, pp. 499–505, 1963.

39. Daeschner, H. W., Seilert, E. E., and Peters, E. D., "Application of Electroformed Precision Micromesh Sieves to the Determination of Particle Size Distribution," *American Society for Testing Materials*, STP No. 234, 1958.

40. Lauer, O., *Staub*, Vol. 20, pp. 69–71, 1960.

41. Lauer, O., *Staub*, Vol. 18, pp. 306–309, 1958.

42. Halász, I., and Heine, E., *Analytical Chemistry*, Vol. 37, pp. 495–500, 1965.

43. Lovelock, J. E., *Analytical Chemistry*, Vol. 33, pp. 162–178, 1961.

XVII. Fluid Physics

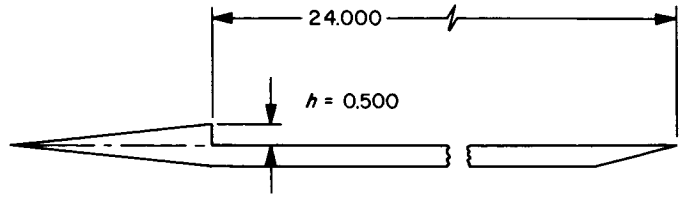
A. Experimental Investigations of the Base-Flow Problem. II. Wedge

F. R. Hama

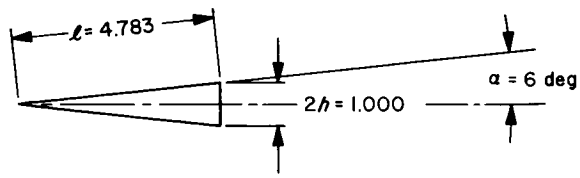
Reported in Ref. 1 is a brief account of the optical observations and pressure measurements of the flow behind a backward-facing step formed by a wedge and a splitter plate as shown in Fig. 1a. In the present test, not only the base pressure but also the surface-pressure distribution, particularly near the separation edge of a wedge without a splitter plate (Fig. 1b), were measured. There have been some speculations without experimental verifications that the base pressure of a wedge might be smaller than that with the splitter plate due to the difference in the reattachment condition. Since, on the other hand, the strength of the lip shock was found to be large (Ref. 2), the measurement of the pressure distribution near the separation edge was thought to be of importance for examining the separation condition.

Fig. 2 shows the base pressure p_b normalized by the free-stream static pressure p_1 for the free-stream Mach numbers $M_1 = 2.01$ through 4.54 and in the range of the Reynolds number based upon wedge-surface condition and its length $Re_2 = 0.2 \sim 2.0 \times 10^6$. The base pressure of the wedge is indeed lower than that for the corresponding backward-facing step, although no substantial difference exists at Mach 2.61 and 3.02. Horizontal lines at the right edge of the figure indicate the base pressure with a boundary-layer trip on the wedge surface for the step configuration.

All the data shown in Fig. 2 were obtained with fences as shown in Fig. 3. The fences are expected to isolate the



(a) WEDGE WITH SPLITTER PLATE



(b) WEDGE

Fig. 1. Experimental configurations

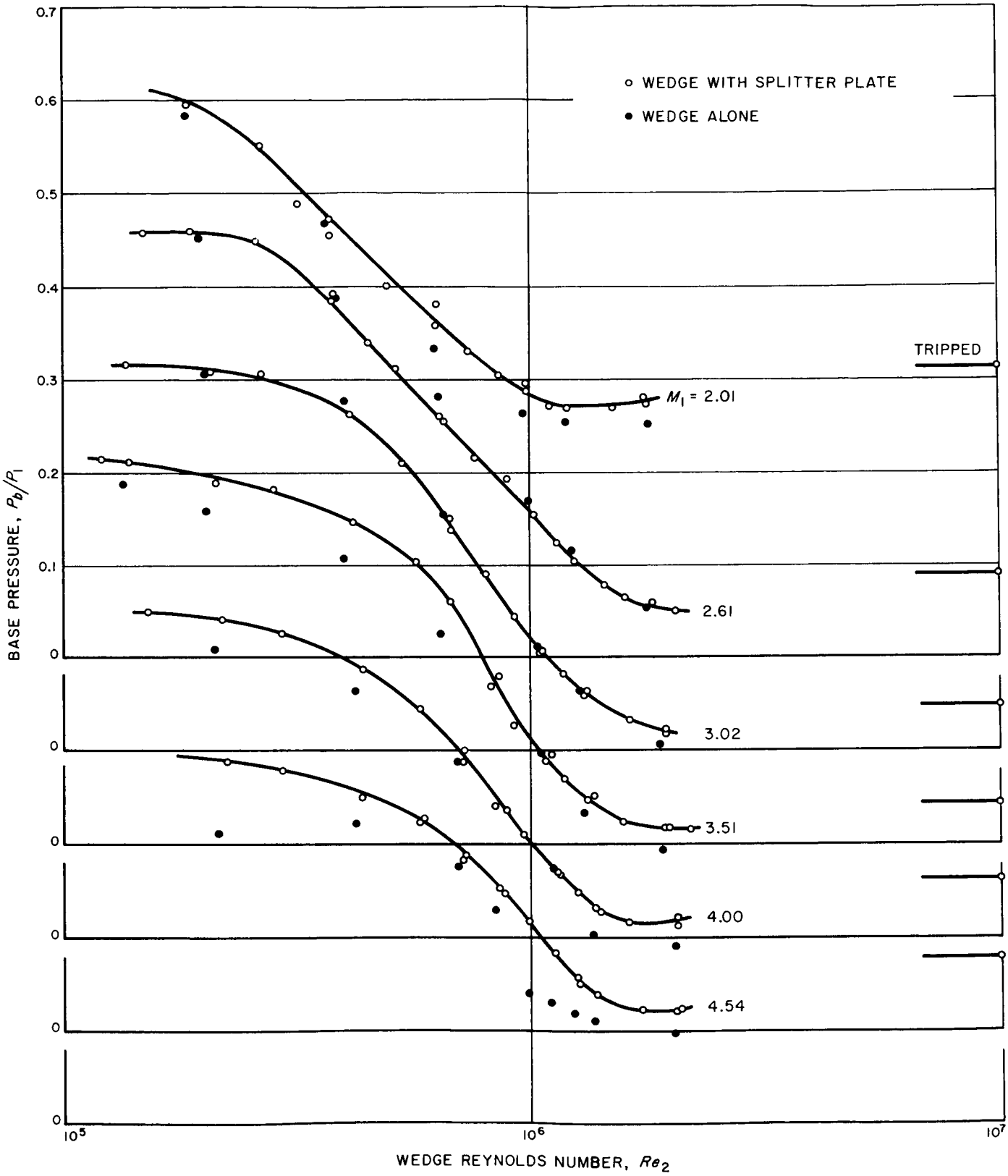


Fig. 2. Base pressure with and without splitter plate

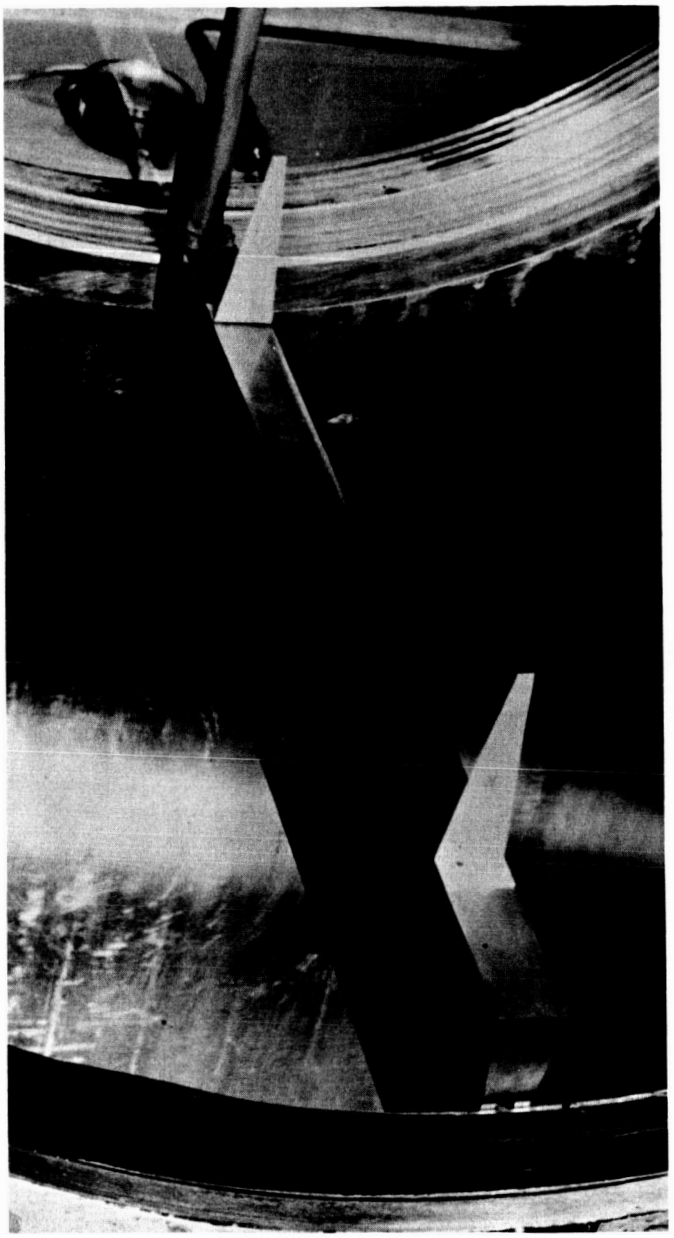


Fig. 3. Wedge model with fences

side-wall effect coming into the central portion of the wake when they are mounted outside of the wall boundary layer (Refs. 3 and 4). The base pressure was also measured without the fences at Mach 2.61, 3.51, and 4.54, and was found to be almost always higher than with the fences, except at lower Reynolds numbers and at Mach 2.61, which apparently corresponds to the case reported in Ref. 3. Base-pressure distributions in the spanwise direction are shown in Fig. 4 for two extreme flow conditions. In all cases, the wall boundary layer was turbulent. The

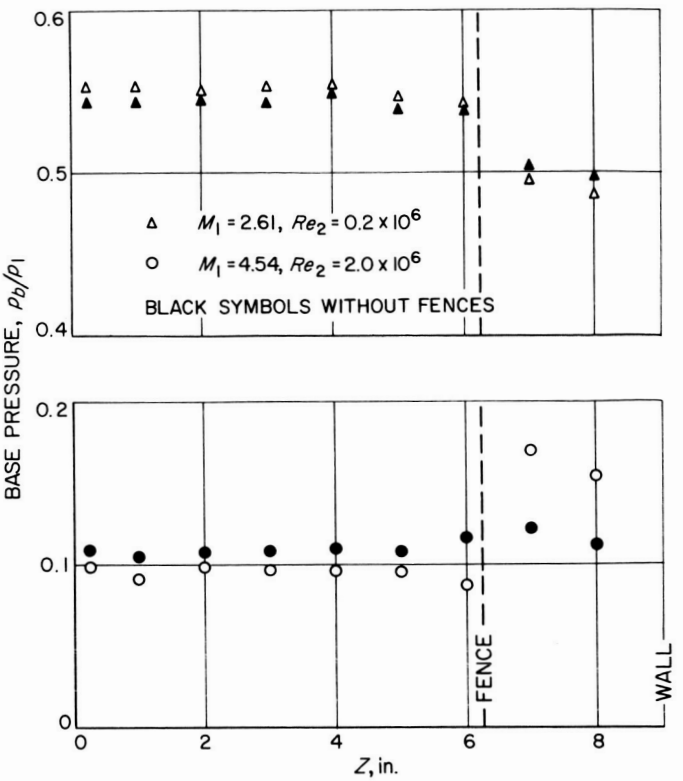


Fig. 4. Spanwise distribution of base pressure

reattachment was laminar at the lowest Reynolds number, whereas it was turbulent at the highest Reynolds number.

The surface-pressure distribution at Mach 4.54 is shown in Fig. 5 in which X is the distance along the windward surface and Y the vertical distance along the base, both measured from the separation edge. The pressure along the windward surface agrees with the inviscid bow-shock solution with an indication of higher values at lower Reynolds numbers due to the viscous interaction, and decreases as the flow approaches the expansion corner. A surprising result is the base-face pressure distribution which revealed a sharp drop immediately behind the separation edge. In addition, there seems to be a slight pressure peak near the edge at lower Reynolds numbers. The sharp pressure drop was consistently observed at Mach 3.51 and higher, whereas the pressure was essentially constant on the base face at lower Mach numbers. This experimental evidence might be interpreted as that the flow did not separate immediately at the edge but in fact turned sharply around the edge and then separated, although the presence of a small separation bubble is not ruled out. If this is indeed the case, the cause of the lip shock might be different than that recently proposed by Weinbaum (Refs. 5 and 6).

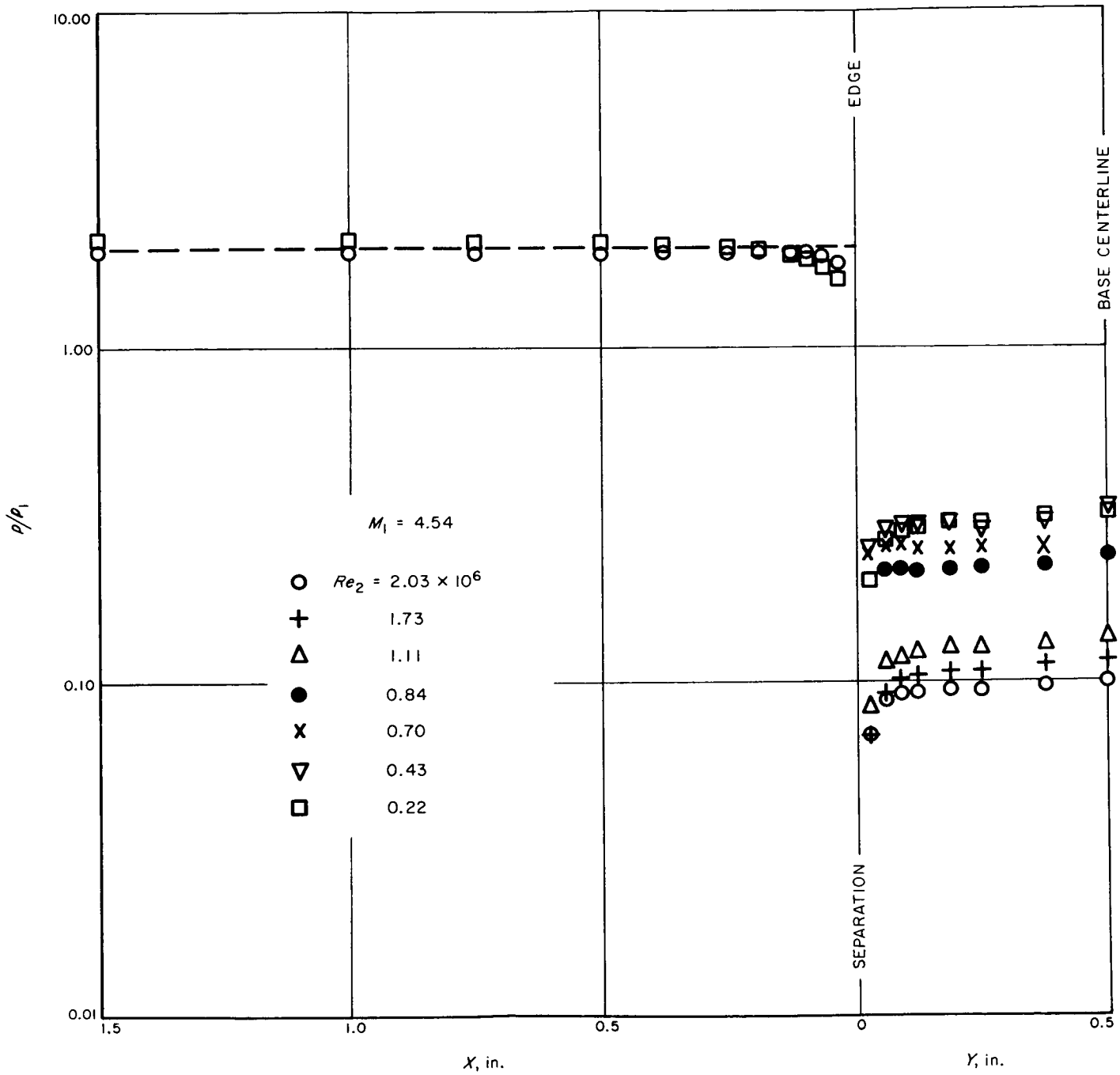


Fig. 5. Surface-pressure distribution

B. The Inviscid Stability of the Cooled Laminar Boundary Layer. Part II

L. M. Mack

In Ref. 7 the effect of wall cooling on the stability of the laminar boundary layer was studied by means of the inviscid theory. The chief result was the computation at $M_1 = 5.8$ of the maximum dimensionless time rate of amplification, $(\alpha c_i)_{max}$, for the first three modes as a function of T_w/T_r, the ratio of wall temperature to recovery temperature. Cooling was found to stabilize the first mode, but to destabilize the second and third modes.

In the present contribution, this calculation is extended to higher Mach numbers. The inviscid eigenvalues of neutral and amplified disturbances have been computed at $M_1 = 8$ and 10 for the insulated-wall boundary layer and the cooled-wall boundary layer with $T_w/T_r = 0.05$. The $M_1 = 8$ computation is for a free-stream temperature, T_1, of 80°K, with an additional computation for the insulated wall with $T_1 = 50$°K. At $M_1 = 10$, $T_1 = 50$°K.

The wave numbers of the first-mode amplified disturbances lie between zero and α_{s1}, the first-mode subsonic neutral solution. The phase velocity of α_{s1} is c_s, the mean boundary-layer velocity at η_s, the point in the boundary layer where $(u'/T)' = 0$. The latter point is called the generalized inflection point because it plays the same role in the stability of the compressible boundary layer as does the inflection point for the incompressible boundary layer. The insulated-wall boundary layer has a single η_s in the interior of the boundary layer. When the wall is cooled, there are two values of η_s, but the second one is always located where $u < 1 - 1/M_1$, and, therefore, does not lead to a subsonic neutral disturbance. As the wall temperature is decreased, the outer η_s moves inward. As long as M_1 is not too large, sufficient cooling will move η_s below the point where $u = 1 - 1/M_1$ and completely stabilize the first mode. Fig. 6 gives the behavior of the mean velocity at η_s as a function of T_w/T_r for a few Mach numbers. It is seen that at about $M_1 = 8$, depending on the free-stream temperature, η_s remains subsonic relative to the free stream for all wall temperatures and the first mode can no longer be completely stabilized by cooling.

The results of the eigenvalue calculations at $M_1 = 8$ and 10 are shown in Figs. 7 and 8. In Fig. 7, for the insulated wall, the amplification rates of the first three modes are given for the two free-stream temperatures. Increasing T_1 from 50 to 80°K increases by a small amount the most unstable wave numbers of all three modes, and decreases the maximum amplification rate 19% for the first mode and 32% for the third mode. The second-mode maximum amplification rate is almost unaffected by T_1. In addition, the decrease in the neutral second-mode wave number, α_{s2}, and the increase in the singular neutral third-mode wave number, α_{i3}, are sufficient so that at $T_1 = 80$°K the second and third modes no longer join as at $T_1 = 50$°K.

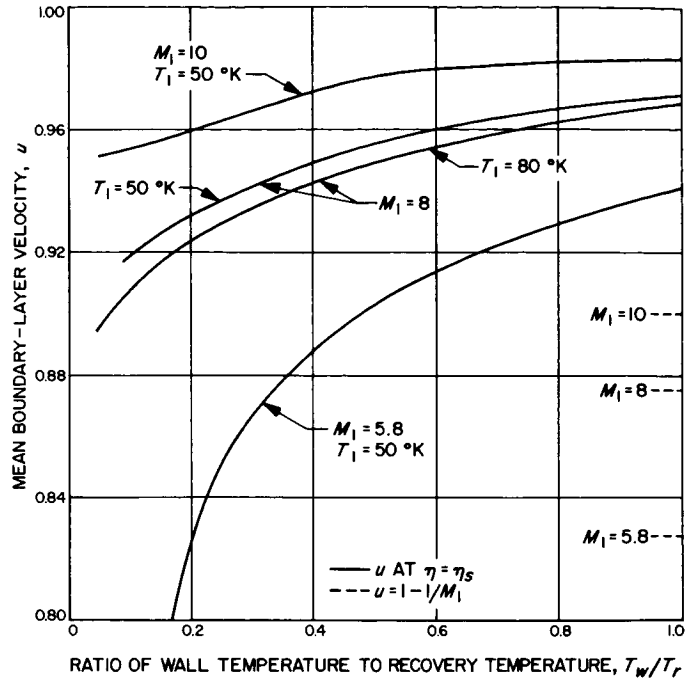


Fig. 6. Mean boundary-layer velocity at generalized inflection point as function of ratio of wall temperature to recovery temperature

The effect of cooling at $M_1 = 8$ is the same as was found at $M_1 = 5.8$. The first mode is stabilized, and the second and third modes are destabilized and shifted to higher wave numbers. In accordance with Fig. 6, the first mode is not completely stabilized by cooling, but the amplification rates at $T_w/T_r = 0.05$ are so small that a first-mode curve cannot be drawn in Fig. 7.

Fig. 8 gives the results for the first four modes at $M_1 = 10$ with a free-stream temperature of 50°K. For the insulated wall, all four modes are joined together and the peaks at the most unstable wave numbers do not

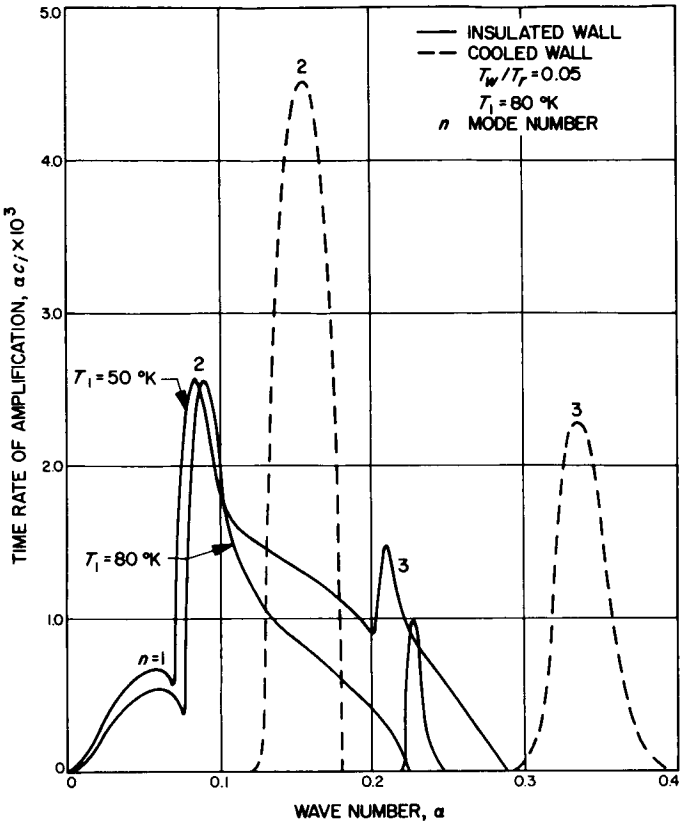


Fig. 7. Time rate of amplification vs wave number for first three modes for insulated wall and cooled wall ($M_1 = 8$)

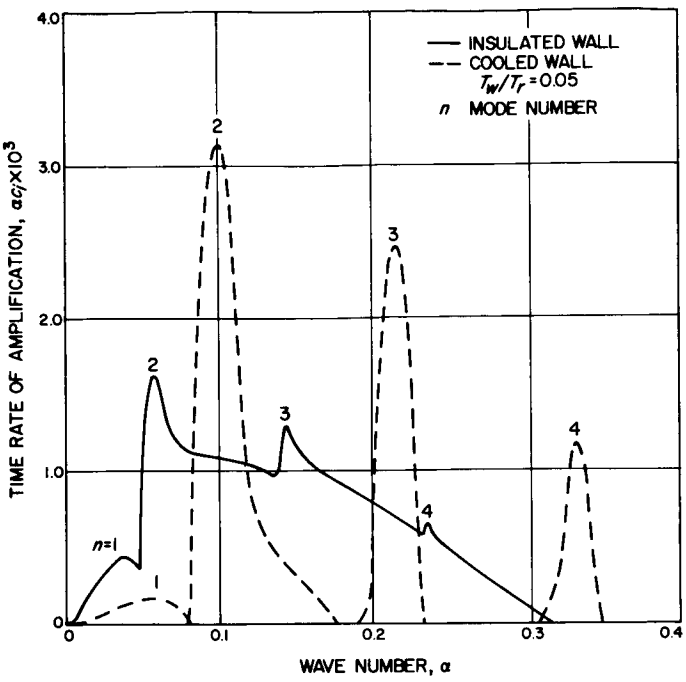


Fig. 8. Time rate of amplification vs wave number for first four modes for insulated wall and cooled wall ($M_1 = 10$)

stand out as sharply from the surrounding wave numbers as at low Mach numbers. For the cooled wall, the first mode is still unstable with an amplification rate of about ⅓ of its value for the insulated wall. The most unstable first-mode wave number is larger for the cooled wall than for the insulated wall. At lower Mach numbers, the most unstable wave number decreases with decreasing wall temperature. The higher modes are destabilized and shifted to higher wave numbers in the same manner as at lower Mach numbers.

At $M_1 = 5.8$, when the over-all amplification of fixed frequency disturbances at finite Reynolds numbers was computed (Ref. 8), it was found that the increase in amplification rate with cooling was almost evenly balanced by the narrowing of the unstable region. The result was that cooling has little effect on the total amplification. In view of the marked decrease with cooling of the width of the unstable region found in the present calculation, it is entirely possible that a similar result holds at both $M_1 = 8$ and 10.

C. The Effective Diameter of an Orifice in Low Density Flow

H. Ashkenas

In Ref. 9 detailed analytical and experimental data regarding the structure of the low density free jet are presented. The variation of gas properties within the jet geometry is generally described in terms of the orifice diameter. At low densities, however, due to viscous and/or transonic effects, the geometric orifice diameter no longer correlates measured jet structural parameters with those obtained by calculation. This has led to the concept of an "effective orifice diameter," a modified diameter which does correlate experimental with analytical data. This modified diameter may be determined experimentally as follows:

(1) Using a Pitot tube, the impact pressure is measured on the axis of the jet, at some axial distance, x, from the orifice.

(2) Applying previously determined Pitot tube corrections (Ref. 10), the local Mach number is measured.

(3) The measured local Mach number is used, in conjunction with the characteristics calculation of Ref. 11 to determine the location, x/D, where that Mach number occurs.

(4) The calculated x/D is compared with the station at which the impact pressure was measured; thus $D_{eff} = x_{meas}/(x/D)_{calc}$.

The procedure described above for the determination of the effective orifice diameter is somewhat cumbersome to use, especially if a large number of different orifices are to be used, since the effective orifice diameter must be defined over the entire operational range of Reynolds number. In order to examine the behavior of the effective orifice diameter, data reported previously in Ref. 9 was re-examined with a view toward correlation in terms of the effusive flow discharge coefficient, Γ, defined in Ref. 12, i.e.,

$$\Gamma = \Gamma_{Re\to\infty}\left(1 - \frac{1}{2\sqrt{Re}} \pm \cdots\cdots\right)$$

where

$$\Gamma_{Re\to\infty} = 0.582$$

$$Re = \frac{D}{\nu_0}\left(\frac{p_0 - p_1}{\rho_0}\right)^{1/2}$$

where ν_0 = stagnation kinematic viscosity, p_0/p_1 = pressure ratio across the orifice, and ρ_0 = stagnation density

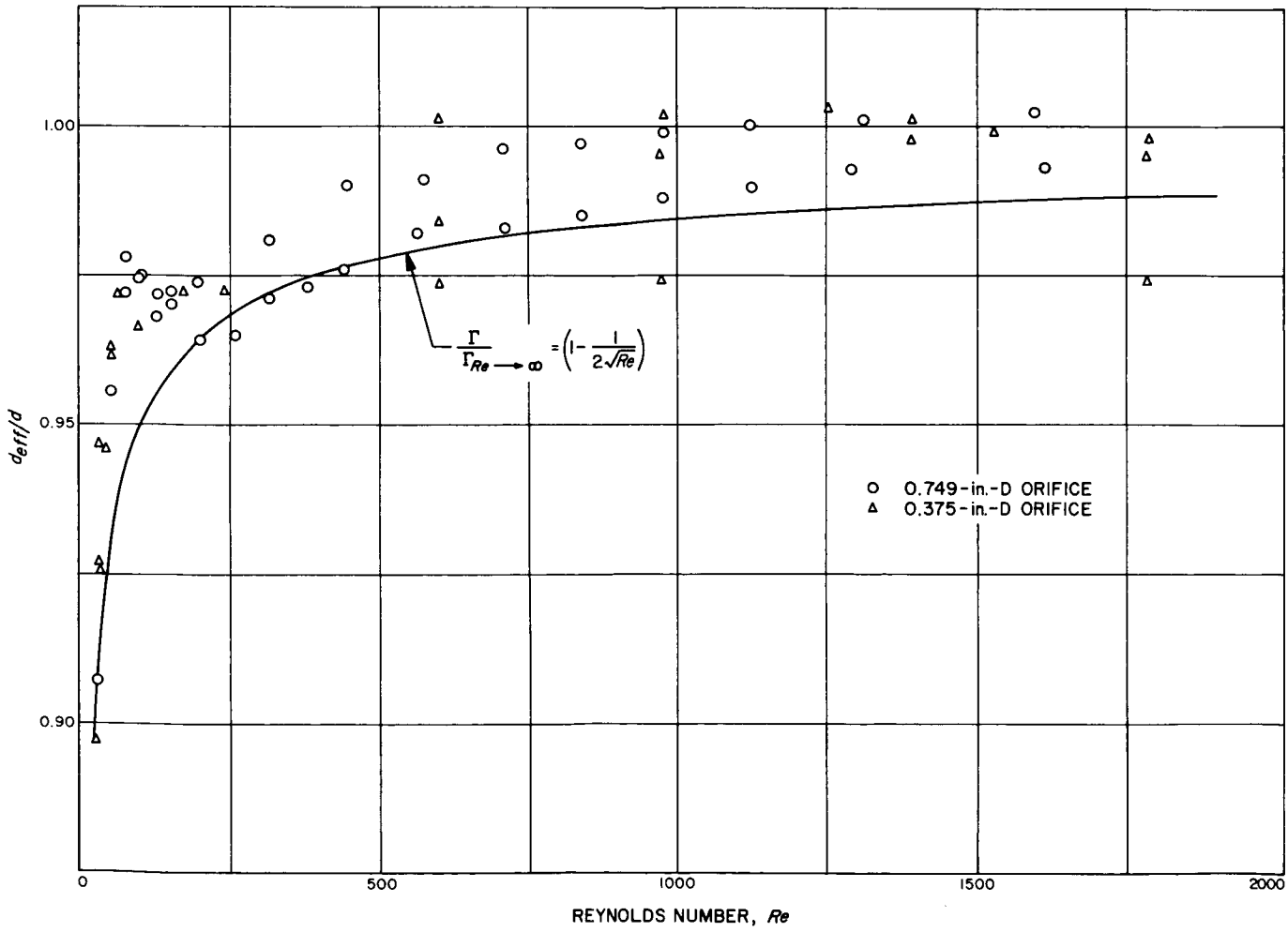


Fig. 9. Correlation between effective diameter and discharge coefficient for low Reynolds numbers

The discharge coefficient Γ relates the mass flow through the orifice to a characteristic velocity, W, and the stagnation density, viz.,

$$\Gamma = \dot{m}/\rho_0 WA$$

(see Ref. 12 for details).

Fig. 9 shows the correlation obtained with two orifices ⅜-in. diameter and ¾-in. diameter and the discharge coefficient; plotted here, as a function of the Reynolds number defined above, are d_{eff}/d and Γ/Γ_∞. It is evident that the data indicate a reasonable correlation between d_{eff}/d and Γ/Γ_∞; i.e., for an unknown orifice, if d_{eff}/d is taken to be numerically equal to Γ/Γ_∞, the error in d_{eff}/d should be no greater than 3 to 5%.

No claim is made here that the correlation shown by Fig. 9 adds to our understanding of the physical processes whereby the effective orifice diameter arises; the sole purpose of this article is to describe a simple means of effective orifice determination for wind tunnel workers.

References

1. Hama, F. R., "Experimental Investigations of the Base-Flow Problem, I. Backward-Facing Step," *Space Programs Summary No. 37-34*, Vol. IV, Jet Propulsion Laboratory, Pasadena, California, July 31, 1965.
2. Hama, F. R., "Estimation of the Strength of Lip Shock," *AIAA Journal* (to be published).
3. Ginoux, J. J., *On the Existence of Cross Flows in Separated Supersonic Streams*, Technical Note 6, Training Center for Experimental Aerodynamics, Rhode-St. Genèse, Belgium, February 1962.
4. Hama, F. R., "Interaction of Tunnel-Wall Boundary Layer with Two-Dimensional Models," *Space Programs Summary No. 37-31*, Vol. IV, Jet Propulsion Laboratory, Pasadena, California, February 28, 1965.
5. Weinbaum, S., *The Rapid Expansion of a Supersonic Shear Flow*, Research Report 204, Avco-Everett Research Laboratory, Everett, Massachusetts, January 1965.
6. Weiss, R. F., and Weinbaum, S., *Hypersonic Boundary Layer Separation and the Base Flow Problem*, Research Report 221, Avco-Everett Research Laboratory, Everett, Massachusetts, July 1965.
7. Mack, L. M., "The Inviscid Stability of the Cooled Laminar Boundary Layer," *Space Programs Summary No. 37-29*, Vol. IV, pp. 171–174, Jet Propulsion Laboratory, Pasadena, California, October 31, 1964.
8. Mack, L. M., "The Finite-Reynolds Number Stability of the Cooled Laminar Boundary Layer at $M_1 = 5.8$," *Space Programs Summary No. 37-31*, Vol. IV, pp. 241–244, Jet Propulsion Laboratory, Pasadena, California, February 28, 1965.
9. Ashkenas, H., and Sherman, F. S., "The Structure and Utilization of Supersonic Free Jets in Low Density Wind Tunnels," *Proceedings of the Fourth International Symposium on Rarefied Gas Dynamics*, Academic Press, Inc., New York, N.Y., 1965.

References (Cont'd)

10. Ashkenas, H., "Pitot Tube Corrections in Low Density Flows," *Space Programs Summary No. 37-15,* Vol. IV, p. 123, Jet Propulsion Laboratory, Pasadena, California, June 30, 1962.

11. Owen, P. L., and Thornhill, C. K., "The Flow in an Axially Symmetric Supersonic Jet from a nearly Sonic Orifice into a Vacuum," ARC R&M 2616, 1948.

12. Liepmann, H. W., "Gas Kinetics and Gas Dynamics of Orifice Flow," *Journal of Fluid Mechanics,* Vol. 10, Part 1, 1961.

XVIII. Physics

A. Construction of Invariant Scattering Amplitudes

J. S. Zmuidzinas

In this article we give a prescription for constructing a set of invariant amplitudes for an arbitrary scattering process involving two initial and two final massive particles. The invariant amplitudes are believed to be analytic functions of the Lorentz scalars, which can be formed from the various particle momenta, and are independent of the spin of the particles participating in the scattering reaction. As such, they are the appropriate entities for dynamical calculations.

The construction, or isolation, of invariant amplitudes has been done in the literature for a large number of physical processes involving two initial and two final particles. The general case of arbitrary spins of participating particles has been treated by Barut et al (Ref. 1). They employ the spinor formalism together with the very complicated and cumbersome Clebsch-Gordan machinery which makes the isolation of invariant amplitudes a rather tedious process. Here we offer what we believe to be a simpler procedure based solely on the Poincaré-group properties of relativistic states (Ref. 2). We use the helicity scheme since it is appropriate to both massive and massless particles (although we do not consider the latter) We consistently ignore isospin, strangeness, etc. Although we restrict ourselves to processes involving only four particles, the method is quite general and remains simple in the case of a larger number of particles.

Consider a two-particle system whose state vector may be written as (H denotes the helicity convention)

$$| p_1 s_1 \lambda_1; p_2 s_2 \lambda_2 >_H , \tag{1}$$

where p_i, s_i, and λ_i are the momenta, spins, and helicities, respectively. Under a restricted Poincaré transformation, we have

$$u(a, \ell) \, | \, p_1 s_1 \lambda_1; p_2 s_2 \lambda_2 >_H$$
$$= e^{-ia \cdot \ell \, (p_1 + p_2)} \sum_{\lambda_1' \lambda_2'} D^{s_1}_{\lambda_1' \lambda_2'} [R(\ell, p_1)]$$
$$\cdot D^{s_2}_{\lambda_2' \lambda_2} [R(\ell, p_2)] \, | \, \ell p_1 s_1 \lambda_1'; \ell p_2 s_2 \lambda_2' >_H$$

in the standard notation (Ref. 2). Since a physical scattering amplitude is just the matrix element of the operator $T = (S - 1)/i$ between state vectors of the form Eq. (1), the D-functions are seen to be the essential source of complication in the transformation properties of the scattering amplitude. To show how to avoid this complication, we must first discuss how to construct what we call "invariant state vectors."

The state vector [Eq. (1)] is the simultaneous eigenvector of the twelve commuting operators

$$P_\mu(1), \underset{\sim}{S}(1)^2, H(1), P_\mu(2), \underset{\sim}{S}(2)^2, H(2),$$

where

$$\underset{\sim}{S}(i)^2 = -W(i)^2/P(i)^2,$$

$$W_\mu(i) = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma} M^{\nu\rho}(i) P^\sigma(i),$$

$$H(i) = W_0(i)/[\underset{\sim}{P}(i)^2]^{1/2},$$

and $P_\mu(i)$ and $M_{\mu\nu}(i)$ are the usual momentum operators of the ith particle (see Ref. 3). The helicity operators $H(1)$ and $H(2)$ are clearly not invariant under an overall Lorentz transformation of the two-particle system. However, it is possible to define two operators, $I(1)$ and $I(2)$, which are Lorentz-invariant and which coincide with $H(1)$ and $H(2)$, respectively, in a particular Lorentz frame of the two-particle system, namely, in the center-of-mass (c.m.) frame They are (Ref. 2)

$$I(1) = 2P(2)\cdot W(1)\,[\Delta(m_1^2, m_2^2, s)]^{-1/2},$$

$$I(2) = 2P(1)\cdot W(2)\,[\Delta(m_1^2, m_2^2, s)]^{-1/2},$$

where

$$\Delta(a, b, c) \equiv a^2 + b^2 + c^2 - 2(ab + bc + ca),$$

$$s = (p_1 + p_2)^2.$$

We denote the eigenvalues of $I(i)$ by σ_i and introduce new two-particle state vectors,

$$|\, p_1 s_1 \sigma_1;\, p_2 s_2 \sigma_2 >_I,$$

diagonalizing $P_\mu(i)$, $\underset{\sim}{S}(i)^2$, and $I(i)$, i = 1, 2. It is immediately verified that

$$U(a, \ell)\,|\, p_1 s_1 \sigma_1;\, p_2 s_2 \sigma_2 >_I$$

$$= e^{-ia\cdot\ell(p_1+p_2)}\,|\, \ell p_1 s_1 \sigma_1;\, \ell p_2 s_2 \sigma_2 >_I;$$

i.e., the new state vectors transform under the Poincaré group like the state vectors of two spinless particles! Considering the simple example of a two particle → two particle scattering, we have

$$\tau \equiv {}_I\!<p_1 s_1 \sigma_1;\, p_2 s_2 \sigma_2 \,|\, T \,|\, p_3 s_3 \sigma_3;\, p_4 s_4 \sigma_4>_I$$

$$= {}_I\!<\ell p_1 s_1 \sigma_1;\, \ell p_2 s_2 \sigma_2 \,|\, T \,|\, \ell p_3 s_3 \sigma_3;\, \ell p_4 s_4 \sigma_4>_I$$

so that τ is an invariant amplitude depending on s_i, σ_i, $m_i^2 = p_i^2$, and the Lorentz scalars (Mandelstam variables) $s = (p_1 + p_2)^2$, $t = (p_1 - p_3)^2$, and $u = (p_1 - p_4)^2$, $s + t + u = \Sigma m_i^2$. For a fixed physical process, different sets of values of the σ_i, $i = 1, 2, 3, 4$, label the different independent invariant amplitudes. For example, if $s_1 = s_2 = s_3 = s_4 = \frac{1}{2}$, then $\sigma_i = \pm\frac{1}{2}$ and there are $2^4 = 16$ invariant amplitudes; their number may of course be smaller if T has some special invariance properties under parity, time reversal, and charge conjugation, or under their products.

It remains to find the transformation coefficients between the new invariant and the old helicity state vectors. This will provide us with the relation between τ and the matrix elements of T in the helicity representation. Clearly,

$${}_I\!<p_1 s_1 \sigma_1;\, p_2 s_2 \sigma_2 \,|\, p_1' s_1' \lambda_1;\, p_2' s_2' \lambda_2>_H$$

$$= \delta(p_1 - p_1')\,\delta_{s_1 s_1'}\,\delta(p_2 - p_2')\delta_{s_2 s_2'}\,\tilde{C}(\sigma_1\lambda_1;\, \sigma_2\lambda_2), \qquad (2)$$

where $\tilde{C}$ may depend on spins and momenta. In the c.m. frame *only* $\tilde{C}$ has the simple form

$$\tilde{C}(\sigma_1\lambda_1;\, \sigma_2\lambda_2) = \delta_{\sigma_1\lambda_1}\,\delta_{\sigma_2\lambda_2} \quad (\underset{\sim}{p}_1 + \underset{\sim}{p}_2 = 0). \qquad (3)$$

To find it in any other frame, we use the Lorentz transformation properties of the state vectors in the following fashion:

$${}_I\!<p_1 s_1 \sigma_1;\, p_2 s_2 \sigma_2 \,|\, U(\ell)^{-1}\, U(\ell) \,|\, p_1' s_1' \lambda_1;\, p_2' s_2' \lambda_2>_H$$

$$= \sum_{\lambda_1'\lambda_2'} D^{s_1}_{\lambda_1'\lambda_1}[R(\ell, p_1')]\, D^{s_2}_{\lambda_2'\lambda_2}[R(\ell, p_2')]$$

$$\cdot\, {}_I\!<\ell p_1 s_1 \sigma_1;\, \ell p_2 s_2 \sigma_2 \,|\, \ell p_1' s_1' \lambda_1';\, \ell p_2' s_2' \lambda_2'>_H. \qquad (4)$$

Taking $p_1 = p_1'$ and $p_2 = p_2'$ as arbitrary and considering ℓ as *that particular Lorentz transformation which takes p_1 and p_2 to their common c.m. frame*, we find from Eqs. (2), (3), and (4) that

$$\tilde{C}(\sigma_1\lambda_1;\, \sigma_2\lambda_2) \equiv C(p_1 s_1 \sigma_1 \lambda_1;\, p_2 s_2 \sigma_2 \lambda_2)$$

$$= D^{s_1}_{\sigma_1\lambda_1}[R(\ell, p_1)]\, D^{s_2}_{\sigma_2\lambda_2}[R(\ell, p_2)].$$

It should be noted that C is not unique due to the fact that the c.m. frame of two particles is arbitrary within a spatial rotation or, stated differently, due to the arbitrariness in the *direction* of $(p_1)_{cm} = -(p_2)_{cm}$. This nonuniqueness of C is removed if we agree to take ℓ to be a *straight rotationless* Lorentz transformation. Explicit formulas for ℓ (and other quantities appearing in the above) may be found in the literature, e.g., in Ref. 4.

Generalization of our results to a larger number of particles and application to specific physical processes is left for a future communication.

B. The JPL Spark Chamber Program

W. S. McDonald and L. L. Lewyn

1. Objectives

The objective of this program is to conduct measurements in space of the distributions in energy and direction of energetic particles, using an omnidirectional spark chamber. The primary effort is directed towards studies of protons associated with solar flares; however, we also intend to adapt the apparatus to neutron measurements to be correlated with the charged particle results.

A JPL development program (Phase I) has been underway for two years and has led to a demonstration of the feasibility of the key components of the system. These components will be described in the paragraphs to follow. We plan to describe the more important developments in a series of reports to appear in coming months.

The objective for this fiscal year is to conduct a low-latitude balloon experiment (Phase II) which will constitute a flight feasibility study of an elementary spark chamber and to prepare an advanced instrument for a high-latitude flight (Phase III) during the summer of 1966.

2. Scientific Background

a. Charged-particle experiment. The propagation of charged particles from the Sun to the Earth is controlled by the large-scale structure of the interplanetary magnetic field. The trajectories of these particles depend on the direction and strength of that field, and on the scale of its irregularities. If the directional distribution is measured as a function of gyro radius (which depends on energy) outside the magnetosphere immediately after a solar flare, then information concerning both the initial field and the flare-produced fields can be derived. Such an experiment, therefore, will not only give information on the particles themselves, but will provide an indirect experimental method for studying the large-scale interplanetary field over much of the inner solar system. The experiment will also measure the directional properties of the galactic cosmic rays.

McCracken (Ref. 5) has determined the gross propagation features of solar protons using neutron monitor data. The major result of the analysis is a verification of the spiral structure of the field in a quiet solar wind. The angular coverage of available neutron monitors, however, is quite limited, and the monitors are not sensitive to protons of energy less than about 450 Mev. Little additional information might therefore be gained from further development of the analysis. The present experiment will significantly extend the low-energy cutoff and the angular coverage over that available from the neutron monitors. Also, performing the experiment outside the magnetosphere will avoid the uncertainties introduced by the Earth's field.

Bryant et al (Ref. 6) have obtained data on the anisotropy of solar protons in space by comparing the results of orthogonal scintillation and Geiger counter telescopes. The angular coverage and resolution of that experiment will be greatly extended by the present experiment.

Spark chamber experiments are in various stages of development at several laboratories. These are generally directed toward other goals than the present program. For example, gamma ray chambers are being developed by groups at MIT, Case, Rochester, and Goddard Space Flight Center, and electron chambers are being developed at Caltech and Chicago. Active information exchange between various groups has led to general dissemination of solutions to common problems of design or technique. The much more ambitious spark chamber efforts associated with accelerators have sufficiently different goals that their inputs to this program are minor.

b. Neutron experiment. It is often suggested that neutrons are produced in the Sun during solar flares. For example, protons accelerated to about 30 Mev can produce neutrons from helium. Since neutrons would not be influenced by magnetic fields, time correlation of the proposed neutron measurement with the charged-particle measurement would give information on the acceleration and modulation of the charged particles and on the production of neutrons. The directional discrimination feature of the neutron detector (which represents a modest elaboration of the basic spark chamber) would have the capability of showing whether or not the neutrons were coming from the Sun or whether they were simply background-produced by the interaction of charged particles with the spacecraft or with the detector. This feature is particularly important, since the flux of solar neutrons is expected to be small compared to the solar charged particle flux.

Experiments are now being designed to examine the feasibility of the spark chamber as a neutron detector.

3. Instrumentation

The spark chamber system was chosen in preference to conventional counter telescopes for the charged particle

experiment in order to have both good angular resolution and wide angular coverage. A spark chamber is basically a device for tracing the trajectory of a charged particle. It is thus capable of giving information on the direction of travel of the particle. A schematic view of a conventional four-gap spark chamber is shown in Fig. 1. The gaps are simply spaces between flat metal plates enclosed in a noble gas. When a charged particle passes between the coincidence counters, a high voltage is very rapidly (typically within 1 μsec) switched onto alternate plates. A spark is then drawn between the plates more or less along the ionization wake left by the particle as it traverses the gas. The positions of the sparks therefore trace the path of the particle. In conventional spark chambers the spark positions are recorded photographically. The present system (described below), however, presents the spark data in digital form, which is better suited for a space experiment.

The high-latitude balloon-borne spark chamber is designed to measure the energy spectrum and angular distribution of protons above 50 Mev and alpha particles above 200 Mev.

A schematic drawing of this chamber is shown in Fig. 2. The ultimate spacecraft instrument which is our long term goal will be a lightened and ruggedized version of this chamber. The instrument consists of 6 double-gap chambers placed in a 6-in. cubical array. Inside this spark chamber is a smaller 3-in. chamber of similar configuration. The 3-in. chamber provides the neutron detection feature of the instrument and will be discussed in a later article on the neutron experiment.

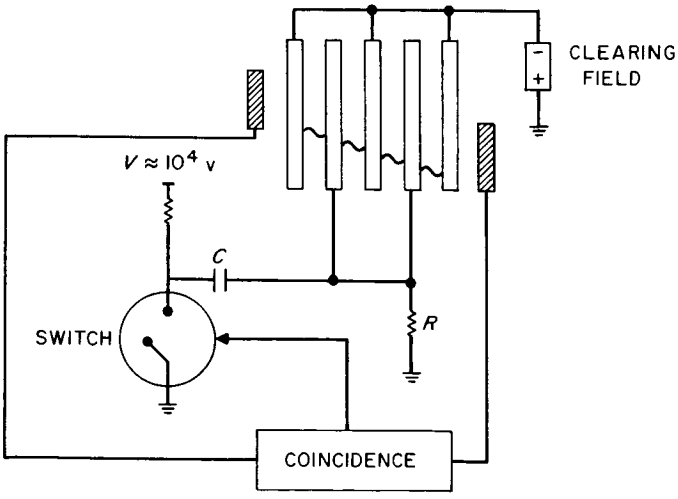


Fig. 1. Schematic of a 4-gap spark chamber

All of the spark chamber plates consist of fine wires strung 16 to the inch on a glass-bonded ceramic frame, so that most of the area of the plate is open. Each wire passes through a magnetic core. When the chamber is triggered, the core associated with the particular wire that carries the spark current is flipped, and that wire can be identified by examining, electronically, the state of the core system associated with a given plate. Adjacent spark gaps have plates with wires strung in mutually orthogonal directions. Therefore, identification of the current-carrying wires in adjacent gaps is equivalent to identification of the position in the chamber at which the spark occurred.

On each outside face of the larger spark chamber are two planar plastic scintillators, optically isolated from each other. A charged particle passing through any two inner scintillators will trigger the spark chamber high voltage. The direction of the particle can be inferred, with an angular accuracy of about 2 deg, from the location of the sparks in the outer spark chamber. Particle energy information is obtained by observing the number of scintillators penetrated and by performing a pulse-height analysis in each scintillator. The chamber is triggered by protons above 50 Mev. For protons from 50 to 250 Mev an energy spectrum and the path and sense of travel of the particle are measured. For protons from 250 to 600 Mev the energy and path of the particle are obtained, but the sense of travel of the particle along its path becomes ambiguous.

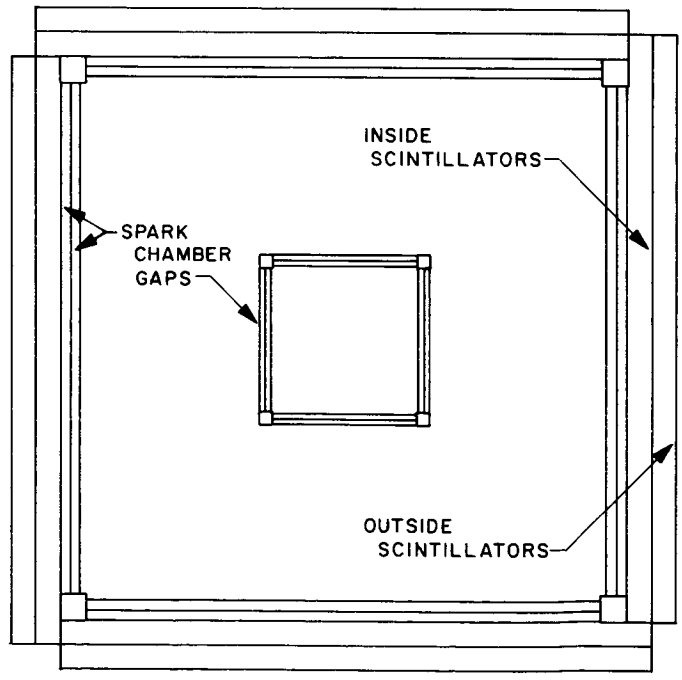


Fig. 2. Schematic of an omnidirectional spark chamber

This is because protons in this energy range penetrate through the entire system with energy losses in each scintillator which are too nearly equal to be discriminated from each other. An energy spectrum for protons above 600 Mev cannot be measured because these high-energy protons exhibit statistical fluctuations in energy loss which are of the same order as the differences in energy loss.

The low-latitude-balloon-flight instrument is a simplified version of the instrument described above. It consists of only two faces of the cubical chamber shown in Fig. 2. Fabrication of this instrument is well underway. It is scheduled for a balloon flight in March 1966.

The low-latitude-balloon-instrument block diagram is shown in Fig. 3. The signals resulting from the passage of a charged particle through the scintillators are amplified and held in separate stretching circuits if the following conditions are met:

(1) The signals from the inner scintillators exceed the threshold setting.

(2) The signals from the inner scintillators arrive within a time separation of less than 0.5 μsec.

(3) The system is not busy processing other signals.

After the passage of the peak of the input signal, the stretcher input is electronically blocked so that the noise resulting from the application of the high-voltage modulator will not alter the value of the signal held in the stretcher. Immediately after the application of the high-voltage modulator pulse, the height-to-time analog-to-digital converters begin to process the signal. When the pulse-height information has been processed, the readout

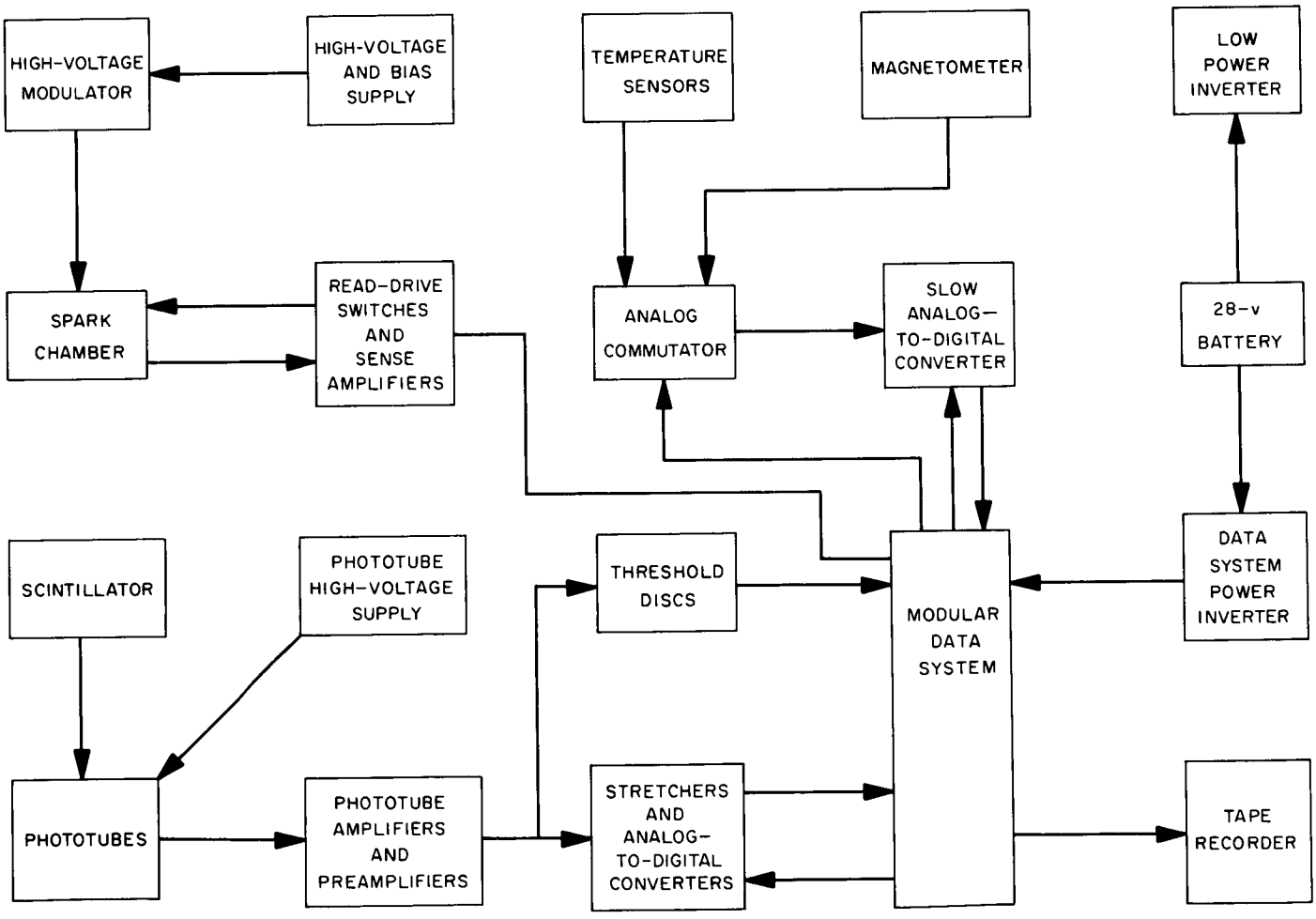


Fig. 3. Low-latitude-balloon-experiment block diagram

sequence begins. The pulse-height and spark location information is read out serially. Each data subframe contains 32 bits. The first 16 bits contain identification and pulse height information and the last 16 bits contain identification and spark location information. Since only one pulse-height value may be read out on each subframe, a minimum of 128 bits are required to read out a single event. While pulse-height information is being read out in the first half of the subframe, the data system scans the cores associated with all four planes in parallel. The scan progresses at a rate high compared to the bit rate until the sense amplifiers indicate that one or more disturbed cores have been interrogated. The location of these cores will then be read out on the last half of the subframe. Every sixteenth subframe is devoted to the readout of supplementary data. The first 16 bits of the supplementary data subframe are devoted to the readout of identification and rate register information. The rate register counts all coincident signals from the inner scintillators. The last 16 bits are devoted to readout of the slow analog-to-digital converter. The slow analog-to-digital converter scans one analog input each master frame. Seven inputs contain temperature information and one input is from a single-axis magnetometer which yields directional information. All data and associated synchronization signals are recorded on magnetic tape.

C. Low-Noise Amplifier for Solid-State Detectors

R. Lockhart

There has been an ever-increasing number of radiation experiments proposed that employ the use of solid-state particle detectors. This is primarily because of their high energy resolution, small size, rugged construction, and negligible power consumption. Solid-state detectors are becoming available with very good noise resolutions, typically 3 to 10 Kev full width at half maximum (FWHM). With detector noise resolution of this order the amplifier will become the limiting factor in system noise resolution. It is therefore necessary to amplify the output signals from these devices with a very low noise amplifier. Because of wide variations in the parameters of these devices (leakage current, capacity, etc.) it is also important that this amplifier be as versatile as possible consistent with small size, low power consumption, reliability, and good design practices. The amplifier must also be able to perform under the wide range of environmental conditions that are normally encountered in space flight. It was with these goals in mind that the present amplifier system was developed.

Since the solid-state detector is in reality a back-biased diode, it appears electrically to the amplifier as a capacitor, and the signals it presents are in the form of charge. The charge produced by the detector is proportional to the energy loss of a charged particle within the detector's depletion layer and is given by:

$$Q_D = q\frac{E}{e_0}$$

where q = charge of an electron = 1.6×10^{-19} coulombs, e_0 = energy loss per electron-hole pair = 3.5 ev for silicon, E = energy loss of the charged particle, and Q_D = detector charge.

The magnitude of the output pulse is given by

$$V = \frac{Q_D}{C_T}$$

$$C_T = C_D + C_S + C_A$$

$$V = \frac{Q_D}{C_D + C_S + C_A}$$

where C_T = total apparent capacity, C_D = detector capacity, C_S = stray capacity, and C_A = effective amplifier input capacity.

Since the detector charge Q_D is a constant, for a given energy particle, the output voltage is a function of the total apparent capacity C_T. The detector capacity C_D is a function of the bias voltage, which may not be particularly stable. If C_A can be made large compared to $C_D + C_S$, most of the signal (i.e., the charge) will be deposited on C_A.

$$Q_D = V(C_D + C_S + C_A)$$

$$= VC_D + VC_S + VC_A$$

$$Q_D = Q_{CD} + Q_{CS} + Q_{CA}$$

If C_A has a value close to or less than C_D, V will be dependent on C_D (i.e., a voltage amplifier).

In an amplifier configuration as shown in Fig. 4, the effective input capacity C_A is equal to

$$C_f(|A_1 + 1|)$$

where A_v is the gain of the amplifier and is large, and C_f is the feedback capacitor. Typical values might be:

Voltage amplifier	Charge amplifier
C_D = 10–300 pf	C_D = 10–300 pf
C_S = 10–15 pf	C_S = 10–15 pf
C_A = 30–50 pf	C_A = 25,000 pf (for C_f = 5 pf, A = 5,000)

This shows that if every effort is made to minimize the stray capacities, maximum signal transfer to the amplifier will be obtained by using a charge sensitive amplifier. In this way most of the charge will be deposited on the feedback capacitor, and the output voltage of the preamplifier will be nearly independent of the detector capacity. These, then, are the gross interface conditions that must be met between the detector and the amplifier.

The following are the design criteria of the amplifier:

a. Performance.

(1) *Noise resolution.* As low as practical using solid-state devices and limited primarily by the input device (should be less than 40 Kev with a 300-pf detector).

(2) *Gain stability and differential linearity.* Limited to 1% so that analog-to-digital converter (ADC) will dominate and temperature calibration will be minimized. ADCs typically have differential linearity of 2%.

The differential linearity of a system is a measure of its ability to reproduce faithfully the distribution of input pulse amplitudes. It is intimately related to the uniformity of the differential gain G_d, which is a function of output voltage V_0. Differential gain is defined as:

$$G_d(V_0) = \frac{\Delta V_0}{\Delta V_{in_{V_0}}}$$

If the average differential gain over the operating range of the system (V_{min} to V_{max}) is given by

$$G_{d_{ave}} \equiv \frac{\int_{V_{min}}^{V_{max}} G_d(V_0)dV_0}{V_{max} - V_{min}}$$

then differential linearity D is the worst-case value of

$$D = \frac{G_d(V_0) - G_{d_{ave}}}{G_{d_{ave}}} \times 100$$

(3) *Conversion gain.* 1 v/Mev for detector signals of 0–10 Mev is a convenient number for many applications. This gives levels high enough to work conveniently into discriminators or ADCs.

(4) *Versatility.* Stable operation for detector capacities varying from 10 to 300 pf. Ability to change time constants so that optimum trade-off can be made between noise resolution and repetition rate.

b. Environmental.

(1) Ability to operate and meet the above condition over a temperature range of −60 to +60°C.

(2) Meet all other normal spaceflight and launch conditions, i.e., shake, shock, magnetic, pressure, etc.

Table 1. Low-noise system characteristics

Parameter	Preamplifier	2-μsec shaper
Input signal handling capability	0–100 Mev	0–12 Mev
Conversion gain	5 pf	1 v/Mev
Differential linearity over signal range	0.5%	0.5%
Gain stability		
−60 to +100° C	1.0%	1.0%
−60 to +60° C	0.2%	1.0%
Maximum repetition rate for less than 1% tail pile up		30,000 count/sec
Output drive capability	1K, 300 pf	1K, 300 pf
Power consumption	60 mw	80 mw
Supply voltage	+24 v, − 6 v	+24 v, −12 v

	Detector capacity		
	10 pf	30 pf	40 pf
Noise resolution FWHM	8.3 Kev	9.7 Kev	10.4 Kev
For values above 40 pf	6.6 Kev + 0.096 Kev/pf		

		Detector capacity		
		50 pf	100 pf	300 pf
Noise resolution FWHM vs temperature	(−60° C)	11.13 Kev	15.23 Kev	34.6 Kev
	(+25° C)	12.16 Kev	16.20 Kev	35.4 Kev
	(+70° C)	13.15 Kev	18.59 Kev	40.6 Kev

c. Additional requirements.

(1) Small size.

(2) Low power consumption.

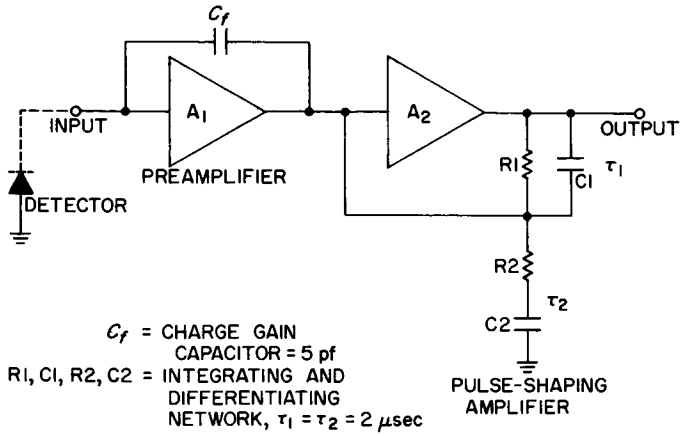


Fig. 4. Amplifier configuration

(3) Reliability.

(4) Component interchangeability selection limited to improving reliability, i.e., not necessary for desired operation.

(5) Careful theoretical and experimental consideration given to major and minor loop frequency and phase stability, dc stability, power dissipation, and reliability.

Table 1 describes the resulting characteristics of the amplifier systems developed to meet these criteria. Fig. 4 is a block diagram of the system, and Figs. 5 and 6 are schematics of specific systems. The pulse-shaping amplifier is optimized for detector capacity of 100 pf and detector leakage of 0.1 μamp; i.e., 2 μsec. Fig. 7 is a resolution versus capacity comparison of this flight-type amplifier and one of the best vacuum-tube laboratory amplifier systems (Ortec Model 103).

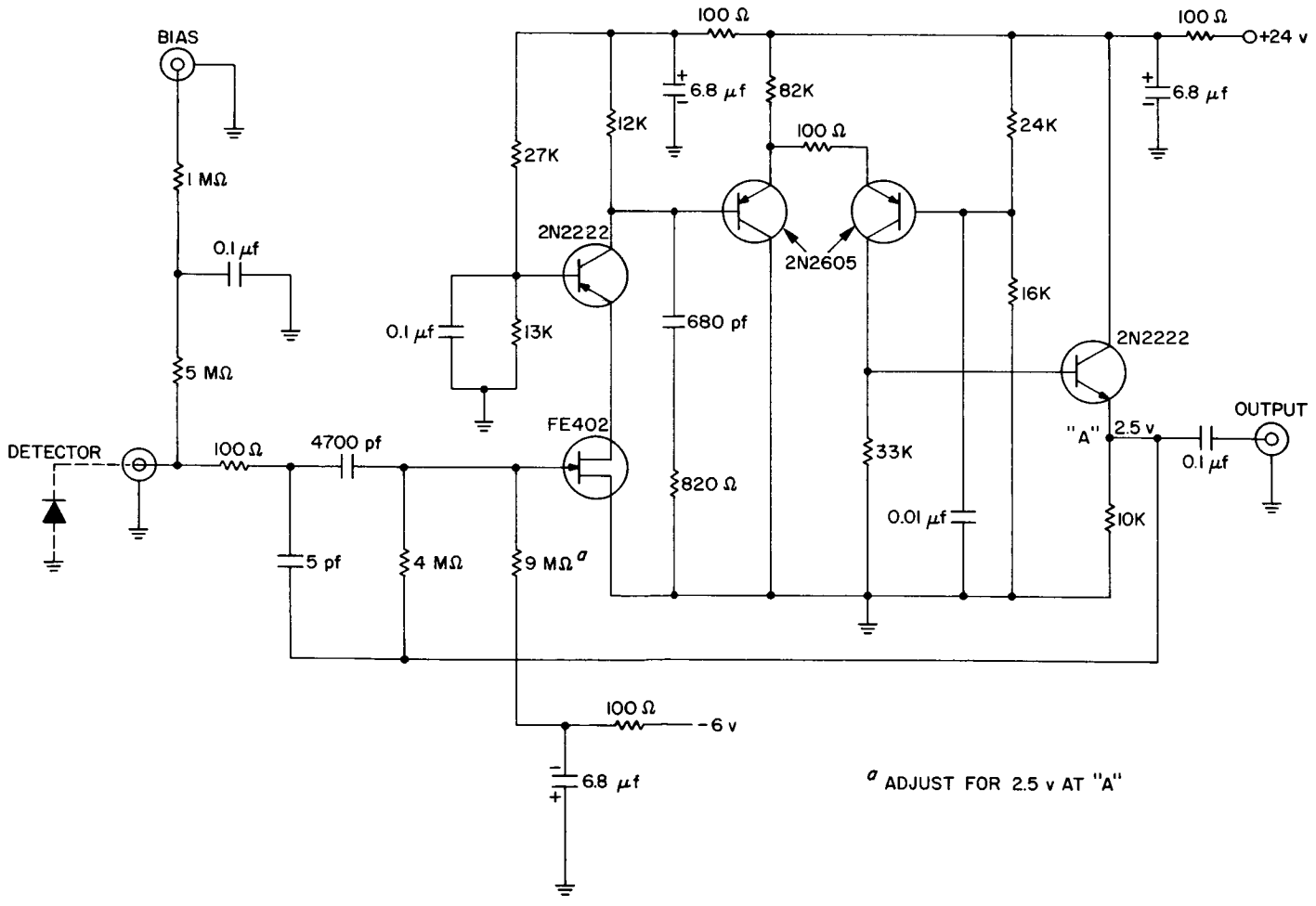


Fig. 5. Charge sensitive preamplifier

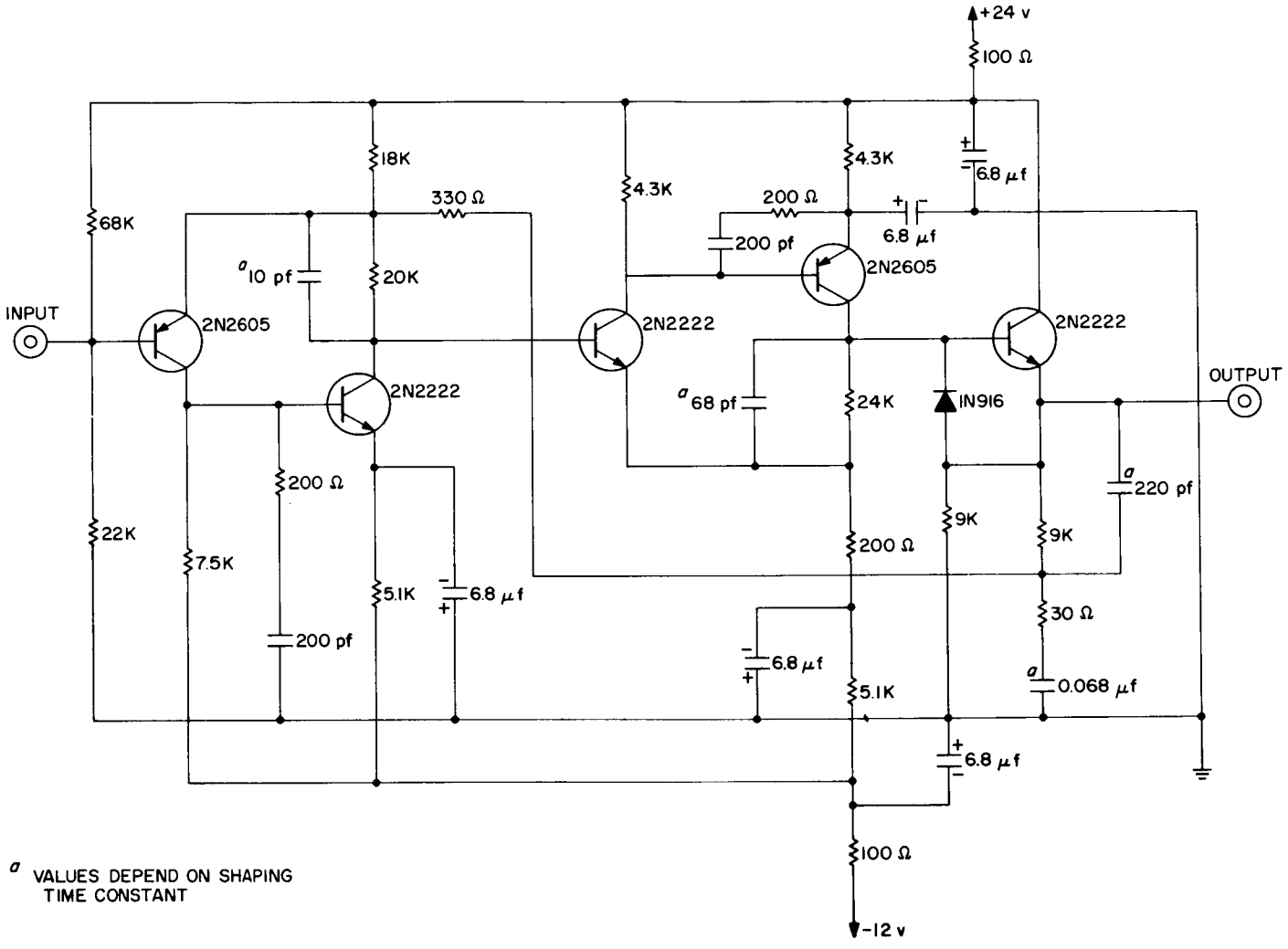


Fig. 6. 2-μ-sec pulse shaper

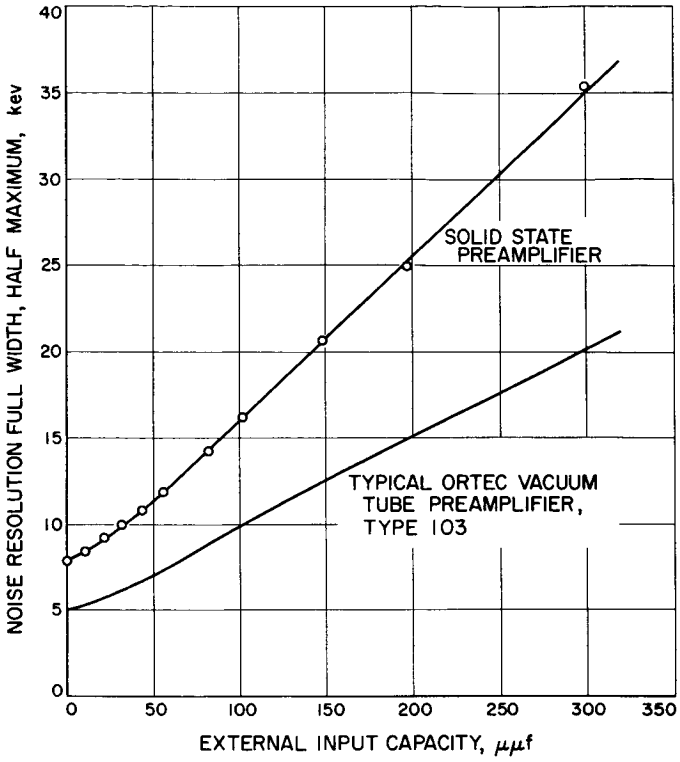


Fig. 7. Resolution spread vs external input capacity

D. Relativistic Rigid Frames in Gravitational Fields

H. D. Wahlquist and F. B. Estabrook

It has long been known that the motions of a rigid body in special relativity are more restricted than in Newtonian mechanics (Refs. 7 and 8). In flat space-time the rigid body has at most *three* degrees of freedom; one point of the body can move arbitrarily, but the motion of all other points is then in general completely determined. A corresponding study of rigid motion in curved space-time manifolds, that is in gravitational fields satisfying the field equations of general relativity, has been undertaken as an application of the dyadic formulation of general relativity reported previously (Refs. 9 and 10). This contribution summarizes some results of that study.

It is found that rigid frames in exterior (vacuum) gravitational fields have at most *one* degree of freedom, and are thus even more restricted than in flat space-time. They can exhibit this single degree of freedom *only* if they are non-rotating and immersed in one of a very special class of gravitational fields. Otherwise, no degrees of freedom at all are permitted.

Our result, showing that in gravitational fields *rotating* rigid frames have *no* degrees of freedom, generalizes the work of Herglotz and Noether (Refs. 7 and 8), who demonstrated this fact for flat space-time. It was obtained by deriving a complete set of first-order time differential equations for the motions. These equations and their significance were given in a previous article (Ref. 11). However, the set of equations which had been obtained at that time was not complete, and a new discussion based on the complete set will appear in a future contribution.

The entire class of exterior gravitational fields admitting rigid frames with a *single* degree of freedom, and the totality of allowed motions therein, have been explicitly obtained. The fields are all degenerate static vacuum fields of Class B in the terminology of Ehlers and Kundt (Ref. 12), or are limiting cases of these. The canonical form of the space-time metrics for these Class B fields can be written as

$$ds^2 = -y^2 h(x)\,dt^2 + y^2 dx^2 + \frac{1}{g(y)}\,dy^2 + g(y)\,dz^2\,, \tag{1}$$

with

$$g(y) = \epsilon - \frac{2m}{y} - \frac{1}{3}\Lambda\, y^2$$

where Λ is the cosmological constant, arbitrary as to sign and magnitude, m is a positive constant, $\epsilon = \pm 1$, and the function $h(x)$ can assume three possible forms:

$$h(x) = \sin^2 x \qquad \text{with} \quad \epsilon = +1 \tag{2a}$$

$$h(x) = \sinh^2 x \qquad \text{with} \quad \epsilon = -1 \tag{2b}$$

$$h(x) = \cosh^2 x \qquad \text{with} \quad \epsilon = -1 \tag{2c}$$

These metrics survive if the cosmological constant Λ vanishes.

The limiting cases, however, exist only for non-zero Λ; their metrics can all be written in the form

$$ds^2 = \frac{1}{|\Lambda|}\{-h(x)\,dt^2 + dx^2 + dy^2 + h(y)\,dz^2\}\,, \tag{3}$$

with again three choices for $h(x)$ as in Eq. (2). For the first choice of $h(x)$, Λ now must be positive; for the second and third choices, Λ must be negative. These then are the *only* exterior fields admitting rigid frames with any degrees of freedom.

The simplest way of exhibiting the rigid motions admitted by these fields is to transform the metrics to a system of coordinates comoving with the rigid frame. Let $(\bar{x}, \bar{y}, \bar{z}, \bar{t})$ be such a system with $(\bar{x}, \bar{y}, \bar{z})$ labeling the points of a rigid frame immersed in the field so that the lines on which $\bar{x}, \bar{y}, \bar{z}$ are held constant and only $\bar{t}$ varies are the world lines of the points of the frame. In these comoving coordinates we have shown that the Class B metrics take the forms

$$ds^2 = -\bar{y}^2\,\bar{h}(\bar{x},\bar{t})\;d\bar{t}^2 + \bar{y}^2 d\bar{x}^2 + \frac{1}{g(y)}d\bar{y}^2 + g(\bar{y})\,d\bar{z}^2 \tag{$\bar{1}$}$$

with

$$\begin{aligned} \bar{h}(\bar{x},\bar{t}) &= \sin^2[\bar{x} + f(\bar{t})]\,, \quad \epsilon = +1 \\ &= \sinh^2[\bar{x} + f(\bar{t})]\,, \quad \epsilon = -1 \\ &= \cosh^2[\bar{x} + f(\bar{t})]\,, \quad \epsilon = -1\,, \end{aligned} \tag{$\bar{2}$}$$

and that the limiting metrics similarly become

$$ds^2 = \frac{1}{|\Lambda|}\{-\bar{h}(\bar{x},\bar{t})\,d\bar{t}^2 + d\bar{x}^2 + d\bar{y}^2 + h(\bar{y})\,d\bar{z}^2\}\,. \tag{$\bar{3}$}$$

In Eq. ($\bar{2}$), $f(\bar{t})$ is an arbitrary function of $\bar{t}$ and so represents explicitly the single degree of freedom possessed by the rigid frames. Its physical significance is revealed by an expression for the square of the acceleration vector, **a**, of the points of the rigid frame:

$$\mathbf{a}^2 = \pm\frac{1}{\bar{y}^2\bar{h}(\bar{x},\bar{t})} - \frac{2m}{\bar{y}^3} - \frac{1}{3}\Lambda\,, \tag{4}$$

where the sign of the first term depends on which form of $\bar{h}(\bar{x},\bar{t})$ is considered. It is clear that the *magnitude* of the acceleration of *one* point of the rigid frame can be specified arbitrarily in time. This then fixes $f(\bar{t})$ so that the motion of the entire frame is determined.

A detailed exposition of the application of the dyadic formalism to the rigid body problem in relativity has been submitted for publication.

E. Thermoelectric Effect in Superconducting Lead

A. F. Hildebrandt and D. D. Elleman

When a hollow tube of pure lead at rest is cooled through the superconducting transition in zero external field, a non-zero magnetic field is found to be trapped in the hollow of the tube. This effect was first observed for lead electroplated on brass and was considered possibly due to ferromagnetic impurities (Ref. 13). The experimental apparatus is described elsewhere (Ref. 14). Also, in order that tubes and solenoids might be switched normal in the helium bath, a warm helium gas switch (Fig. 8) made of glass was attached to the length of the superconducting tubes with either Ducco cement, epoxy cement, styrofoam, or finally DC 11 vacuum grease.

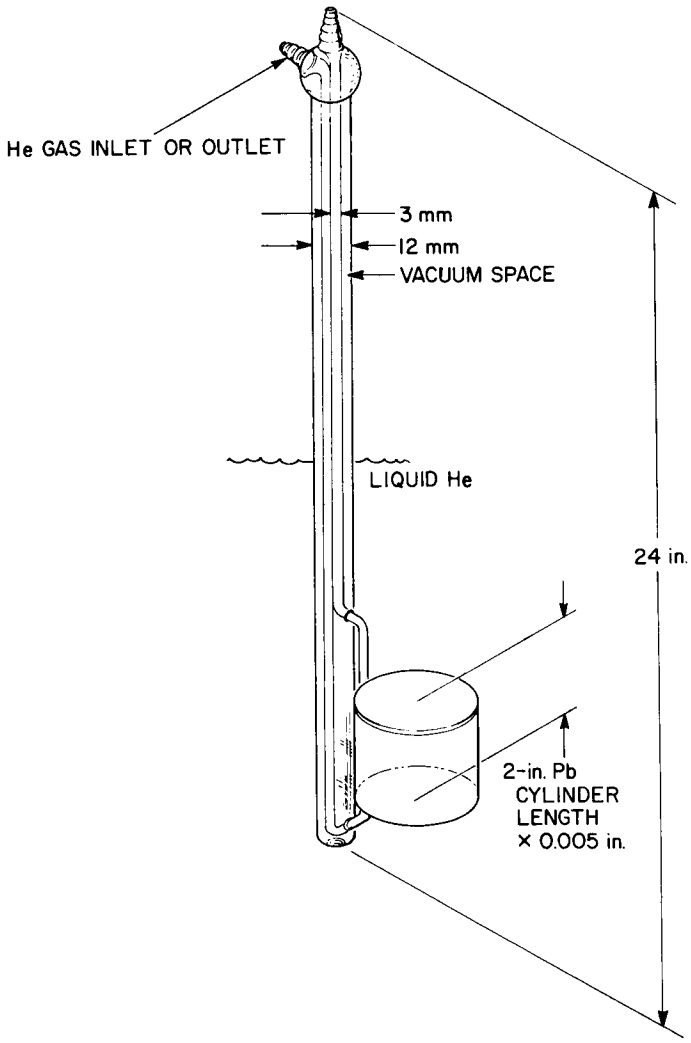


Fig. 8. Helium gas superconductor switch (pyrex glass)

The lead tubes were of five-nines purity material and nominally 5.0-cm diameter, 7.5-cm length, and 0.013-cm wall thickness. The seams (for 15 tubes) were either fused with the heat of a hot pure copper iron or crimped with pressure. One tube was seamless, slip-cast on aluminum. It was found in the earlier experiment that the external field of approximately $\pm 10^{-5}$ gauss (or one gamma) would be measured within the cylinder after transfer. This result is now attributed to a particular rate of transfer and associated rate of cooling through the transition, since initial fields as large as ± 100 γ have been observed for many of the samples, including several of the original specimens and the seamless tube. In experiments subsequent to the original transfer, where the sample is raised out of the helium bath until normal (i.e., where the sample exhibits no field and will not shield an externally applied field) and then immediately lowered into the bath again, the external field ± 1 γ is trapped with a much improved consistency.

When the gas switch is activated either after initial transfer or subsequent to raising of the experiment out of the bath until normal, a field as large as ± 250 γ can be trapped as the switch is closed. The maximum field generated in general is not equal to the field finally trapped, although when the switch is activated there is no shielding against externally applied dc fields. The final field configuration varies from being axial and nearly perfectly homogeneous to having a perpendicular component comparable to the axial component. It is also observed that the field can be modified by reversing the direction of gas flow. This effect can be enhanced by canting the switch relative to the axis of the tube. The possibility of this being a probe effect has been eliminated both by keeping the probe external to the experiment while switching and by finding the same results with three separate probes.

In summary, for the pure lead tubes it is found that the trapped field can deviate considerably from the applied field. The initial trapped value after transfer varies from the same sample. The field that is trapped after raising the sample out of the bath until normal and again lowering is very nearly the applied field. The value of the trapped field for a given switching procedure is in general non-zero but reproducible provided the switch is untouched and kept below liquid nitrogen from day to day. All lead tubes give the same result within an order of magnitude.

It should be noted that the effect reported here was virtually absent when the tubes were cooled through the transition while rotating at constant angular velocity (Ref. 14). There are two considerations which make this reasonable. First, the sample was cooled through the transition while at rest, then warmed just until normal, and then rapidly cooled while rotating. This should reduce the magnitude of the temperature fluctuations. Second, rotating while cooling should provide an azimuthal averaging of temperature.

The effect has not been observed for lead evaporated on either glass or quartz. The outer dimensions of the substrate were the same as that of the lead tubes, and the thickness of the coating was about 10,000 Å. Also the effect was absent within ± 1 γ for an electron-beam-welded pure niobium tube of the same outside dimensions.

No such effect has been found for a large superconducting lead shield used in the above experiments. The shield is 15 cm in diameter and 30 cm long; it has a 0.013-cm thick wall with a hemispherical bottom. The purity of the lead is the same as used previously. This shield has never been switched, but it has trapped ± 0.25 γ absolute for approximately zero applied field. In fact it exhibits a desirable Meissner effect, inasmuch as 40 γ applied results in less than 1 γ trapped.

The experiments show qualitatively that the effect is dependent on the dynamics of cooling and is not simply a property of the lead seam. To investigate this further, a 30-turn solenoid of 0.078-cm diameter wire of six-nines-purity lead was made. The diameter of the solenoid was again 5.0 cm and the length 7.5 cm. The ends were either crimped or fused with heat in an inert atmosphere. In order qualitatively to study temperature gradient effect, a 5-cm length of wire was attached along the length of one of the helium gas switches with DC 11 vacuum grease.

First it was observed that the magnitude of the trapped field was of the same order of magnitude (± 70 γ max) as observed for the tubes. The only suggestion for the similarity is that the current density in the bulk material at the transition is about the same in both cases. Further it was observed that there was an optimum switching time, the trapped field being less for longer and shorter times than the optimum. Also it was found that the field could be reversed by reversing the direction of gas flow. The direction of the field could be made to reverse by rotating the coil 180 deg while the switch remained fixed, consistent with the idea that the effect is due to the temperature gradient in the switch. The direction and magnitude of the field was reproducible for a given switching procedure for three runs over a period of as

many days. These results were borne out with a second coil which qualitatively gave the same result. The field in all cases was in agreement with the electron flow being in the same direction as the gas flow.

A large solenoid of 12.5-cm diameter, 12.5-cm length, and only 12 turns gave a field of less than 1 γ. When the turns per unit length were made the same as for the preceding solenoids, fields of ± 20 γ resulted.

These results show that thermally-induced currents are possible in superconducting lead. The determining features appear to be the cross-sectional area of the conductor and the temperature gradients at the transition temperature. It is suggested that the effect is not the ordinary thermoelectric effect but rather a quantum effect similar to the fountain effect in helium below the λ point where the temperature gradient provides a driving force for the normal fluid. This is under further investigation.

References

1. Barut, A. O., Muzinich, I., and Williams, D. N., "Construction of Invariant Scattering Amplitudes for Arbitrary Spins and Analytic Continuation in Total Angular Momentum," *The Physical Review*, Vol. 130, p. 442, 1963.
2. Zmuidzinas, J. S., "Internal Automorphisms of a Two-Particle System," *Space Programs Summary, No. 37-27*, Vol. IV, p. 141, Jet Propulsion Laboratory, Pasadena, California, June 30, 1964.
3. Zmuidzinas, J. S., "Relativistic Many-Particle States," *Space Programs Summary No. 37-26*, Vol. IV, p. 174, Jet Propulsion Laboratory, Pasadena, California, April 30, 1964.
4. Joos, H., "Zur Darstellungstheorie der inhomogenen Lorentzgruppe als Grundlage quantenmechanischer Kinematik," *Fortschritte der Physik*, Vol. 10, p. 65, 1962.
5. McCracken, K. G., "The Cosmic Ray Flare Effect," *Journal of Geophysical Research*, Vol. 67, pp. 423–458, 1962.
6. Bryant, D. A., Cline, T. L., Desai, U. D., and McDonald, F. B., "Explorer-12 Observations of Solar Cosmic Rays and Energetic Storm Particles after the Solar Flare of September 28, 1961," *Journal of Geophysical Research*, Vol. 67, pp. 4982–5000, 1962.
7. Herglotz, G., *Annalen der Physik*, Vol. 31, p. 393, 1910.
8. Noether, F., *Annalen der Physik*, Vol. 31, p. 919, 1910.
9. Wahlquist, H., and Estabrook, F., "Dyadic Formalism and the Problem of Rigidity in Relativity," *Space Programs Summary No. 37-25*, Vol. IV, pp. 265, 266, Jet Propulsion Laboratory, Pasadena, California, February 29, 1964.
10. Estabrook, F., and Wahlquist, H., "Dyadic Analysis of Space-Time Congruences," *Journal of Mathematical Physics*, Vol. 5, p. 1629, 1964.
11. Wahlquist, H. D., and Estabrook, F. B., "Rigid Motions in Einstein Space," *Space Programs Summary No. 37-32*, Vol. IV, pp. 233–235, Jet Propulsion Laboratory, Pasadena, California, April 30, 1965.
12. Ehlers, J., and Kundt, W., "Gravitation: An Introduction to Current Research," L. Witten, Ed., Chap. 2, John Wiley & Sons, New York, N.Y., 1962.
13. Hildebrandt, A. F., *Physical Review Letters*, Vol. 12, p. 190, 1964.
14. Hildebrandt, A. F., *Proceedings of the IX International Conference of Low Temperature Physics*, 1964 (to be published).

TELECOMMUNICATIONS DIVISION

XIX. Communications Elements Research

A. Microwave Electronics

W. H. Higa

1. Hydrogen Maser Frequency Standard

In anticipation of future requirements for precision frequency standards for tracking deep space probes it is desirable to evaluate new and improved devices. One such device which appears worthy of serious consideration is the atomic hydrogen maser developed by Professor Ramsey of Harvard (Ref. 1). This machine is currently being manufactured by Varian Associates (Ref. 2). Two of these frequency standards are being purchased by the Telecommunications Division of JPL for study and evaluation. A qualitative description of the hydrogen maser is given below, and various operational aspects of the device will be discussed in future articles.

The hydrogen maser is somewhat similar to the original ammonia beam maser: an atomic beam of hydrogen is passed through a state selector (focuser) which allows only excited atoms to enter a microwave cavity where the atom can make a radiative transition and emerge in the lower state. The principal difference here is that a magnetic dipole rather than an electric dipole transition is involved.

Fig. 1 is a schematic of the maser. Molecular hydrogen (H_2) gas, controlled by valves, leaks into an RF discharge bulb, where it is thermally converted to atomic hydrogen. Atomic hydrogen effuses out of the nozzle as a narrow beam into the magnetic state separator or focuser where upper state atoms are focused and lower state atoms are defocused. The excited atoms thus enter a resonant cavity tuned appropriately to the transition frequency.

A characteristic of the magnetic dipole transition utilized in the hydrogen maser is that it is roughly one thousandth the strength of the electric dipole transition utilized in the familiar ammonia maser. To alleviate this weak transition problem, Professor N. Ramsey made a major contribution to maser technology by introducing the teflon coated quartz bulb in the resonant cavity which allows the excited atoms to bounce back and forth many times before emerging from the cavity. Hence, the small transition probability is compensated for by a much longer interaction time.

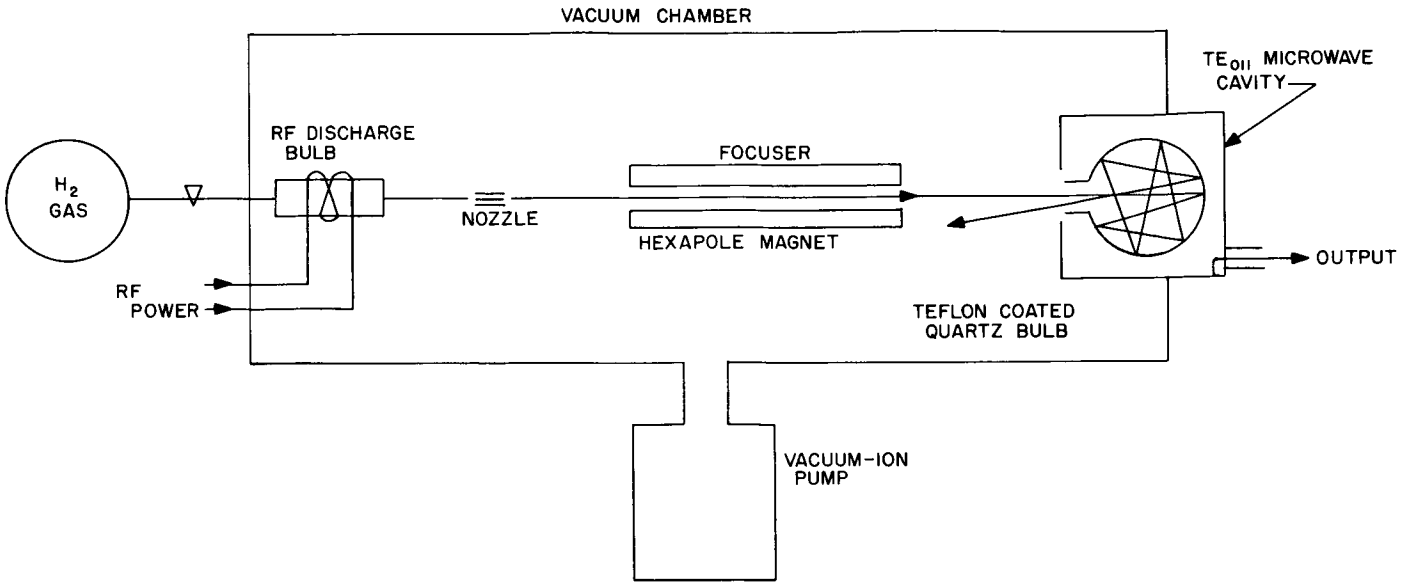


Fig. 1. Schematic diagram of atomic hydrogen maser. A typical trajectory for an upper state atom is shown

Teflon on quartz provides a surface which has virtually no effect on the hyperfine energy levels of the hydrogen atom. The long interaction time between atoms and radiation field results in a usable power level and a very narrow line width (or high Q_H) for the atoms. Thus, a very stable source of oscillation is achieved.

The quantum mechanical transition involved in this maser is the electron-spin resonance of the valence electron in the field due to the nuclear magnetic moment of hydrogen. Fig. 2 shows the energy level scheme for hydrogen as a function of applied magnetic field. It is noted that for weak fields the $F = 1$, $m_F = 0$ state and the ground state $F = 0$ have energies which exhibit only second order dependence on the applied field. The transition between these two levels is, therefore, used for good frequency stability. It is also noted that at high field strengths the two states mentioned above have energies which vary in opposite senses for increasing fields. This fact is used in the state selector (focuser) to separate the ground state atoms from the excited atoms.

Alnico permanent magnets are used in a hexapole configuration to produce a radially nonuniform magnetic field for the focuser. Hence, the unpredictable arcing which occurred in the high voltage electrostatic focuser of the ammonia maser is completely eliminated.

The spent gas is finally pumped out of the system by means of vacion pumps which need servicing or replacement once a year.

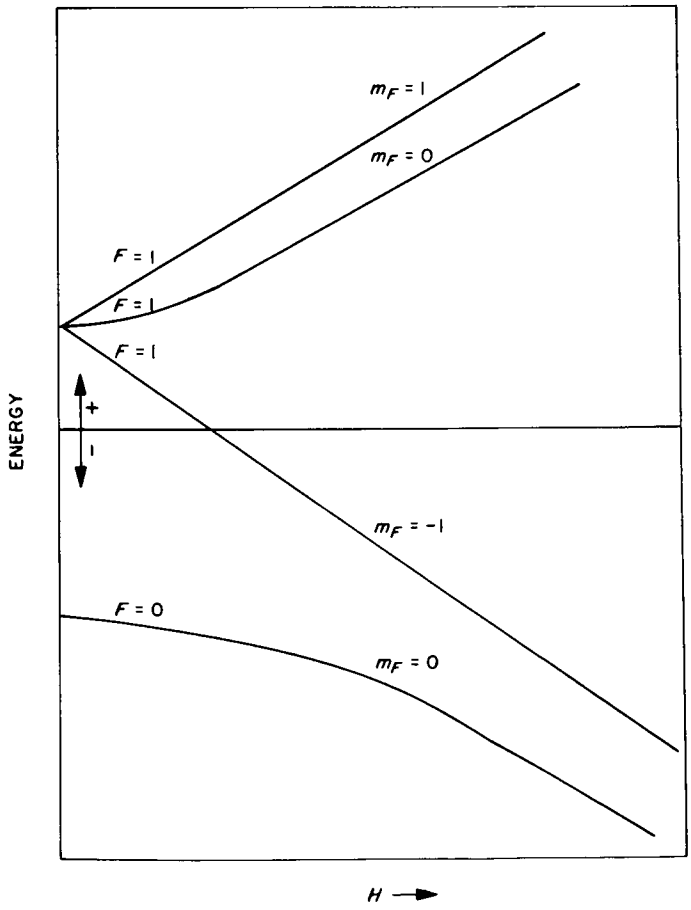


Fig. 2. Diagram for hyperfine energy levels for atomic hydrogen in a magnetic field

The resonant cavity detunes the atomic oscillator according to the familiar expression

$$f_{osc} - f_o = \frac{Q_{cav}}{Q_{atom}} (f_{cav} - f_o)$$

where

f_{osc} = frequency of oscillation

f_o = frequency of atomic resonance

f_{cav} = center frequency of resonant cavity

Q_{cav} = loaded Q of cavity (typically 4×10^4)

Q_{atom} = Q of atoms (typically 1.4×10^9)

Thus, for a detuning of a part in 10^{13} the cavity tuning must be held constant to around one part in 4×10^{-9} or to within 6 cps. This is achieved in the Varian hydrogen maser by means of a silver plated fused-quartz cavity with temperature regulation. In addition, the ferrite isolator in the output circuit is also held at constant temperature in order to minimize the frequency pulling effect.

B. Solid State Circuits

R. Brantner

1. Summary

A system consisting of a low noise ac amplifier, demodulator, and dc amplifiers has been designed and constructed for use in a millimeter wave radiometer.

2. Amplifier-Demodulator System

A system was required having the following characteristics:

(1) 37 cps tuned amplifier with equivalent input noise of approximately 0.2 μv,

(2) a reference channel utilizing an externally generated square wave signal to drive a phase-adjustable demodulator,

(3) a total power amplifier,

(4) a fixed time-constant output dc amplifier,

(5) a variable time-constant output dc amplifier.

Fig. 3 is a block diagram of the resulting instrument.

The reference input is filtered by a resistance-capacitance tuned Q-multiplier (Ref. *SPS 37-34*, Vol. IV) to produce a sinusoidal output signal. This signal passes through two parallel channels to produce four output signals. These signals, displaced in phase in 90 deg steps, are applied to a phase control network. The resulting signal is amplified and limited to make the level independent of the input amplitude. It then passes through a regenerative saturating amplifier. The two outputs of the reference channel, symmetrical square waves of constant amplitude and 180 deg phase difference, are applied to a series-shunt demodulator (Fig. 4).

The signal input is an amplitude-modulated IF signal that is detected by a broad-band detector (Telonics type XD-3A), which is loaded by a resistor empirically determined for the best possible square-law response. The detected signal is applied directly to a dc amplifier with a fixed time constant to produce a total power output.

The same signal is capacitively coupled to the low noise preamplifier (Fig. 5), consisting of two fixed-tuned amplifier stages and an impedance-transformation output stage. The amplifier has a bandwidth of about 5 cps. Its voltage gain was set to about 200 to prevent saturation at the highest expected input signal level. Maximum voltage gain is about 3000. The signal passes through an adjustable attenuator and into an untuned postamplifier, which applies it to the demodulator, two insulated-gate field effect transistors used as switches. The demodulated signal is then fed to the two output dc amplifiers.

The signal channel has an equivalent input noise of 0.20–0.23 μv and, at maximum sensitivity, requires a minimum input signal of about 80 μv for full dc output. The system is linear within $\pm$ 1%. Maximum AF input signal is about 10 mv at minimum sensitivity.

Meters are provided for both signal and reference channels to facilitate setting the appropriate signal levels of each channel. The reference output signal can be varied from 0 to 360 deg with respect to the reference input. A modification to permit continuously variable tuning of the signal channel is being considered.

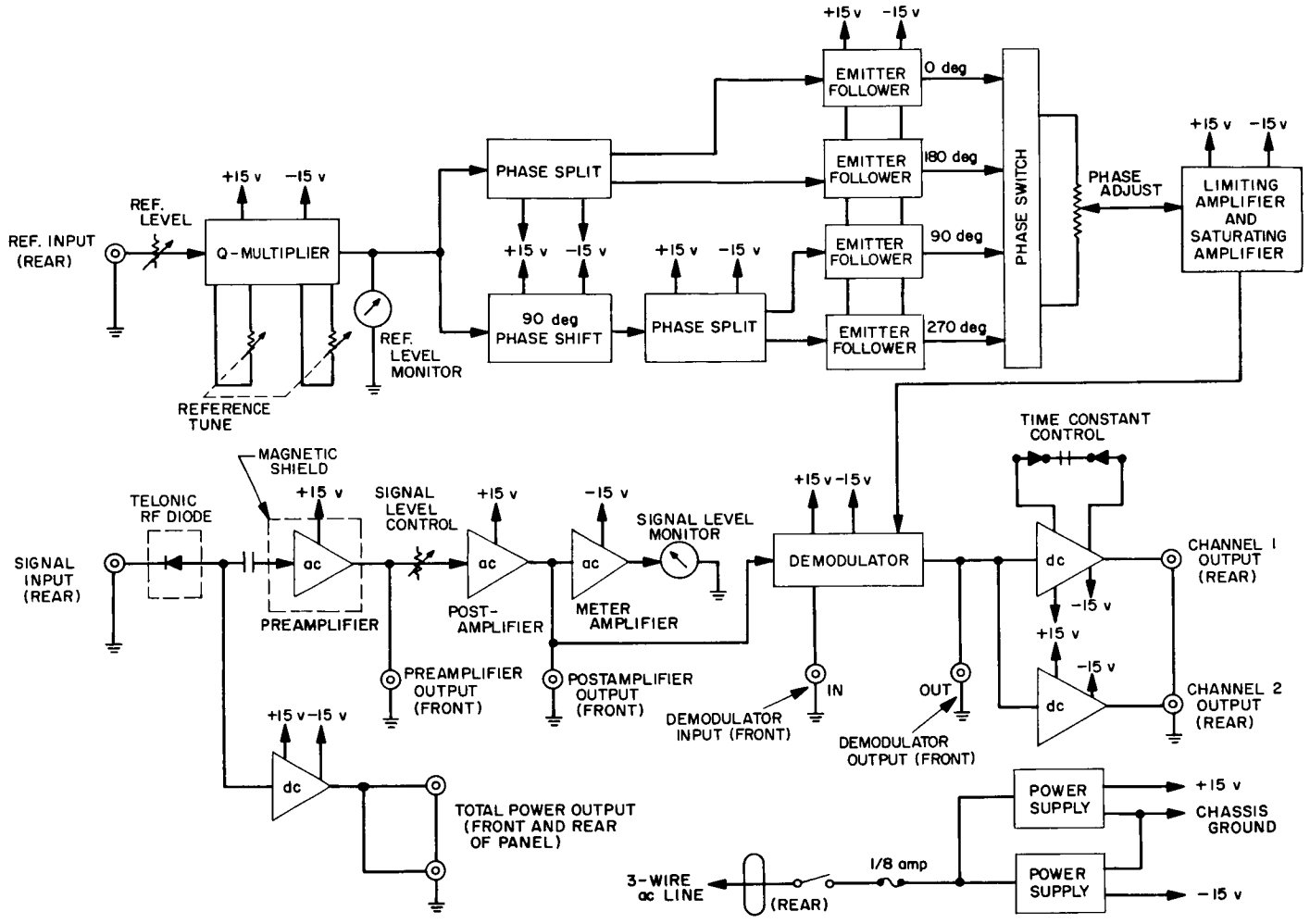


Fig. 3. Amplifier-demodulator, block diagram

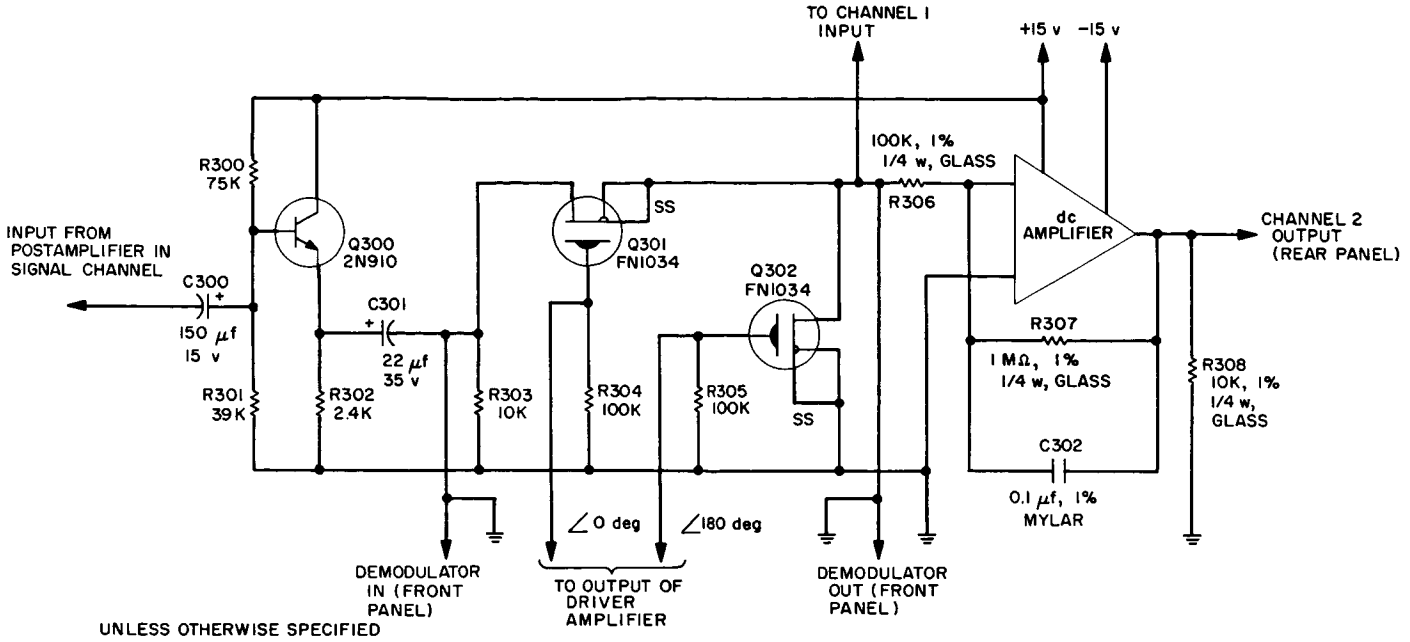


Fig. 4. Demodulator and dc amplifier

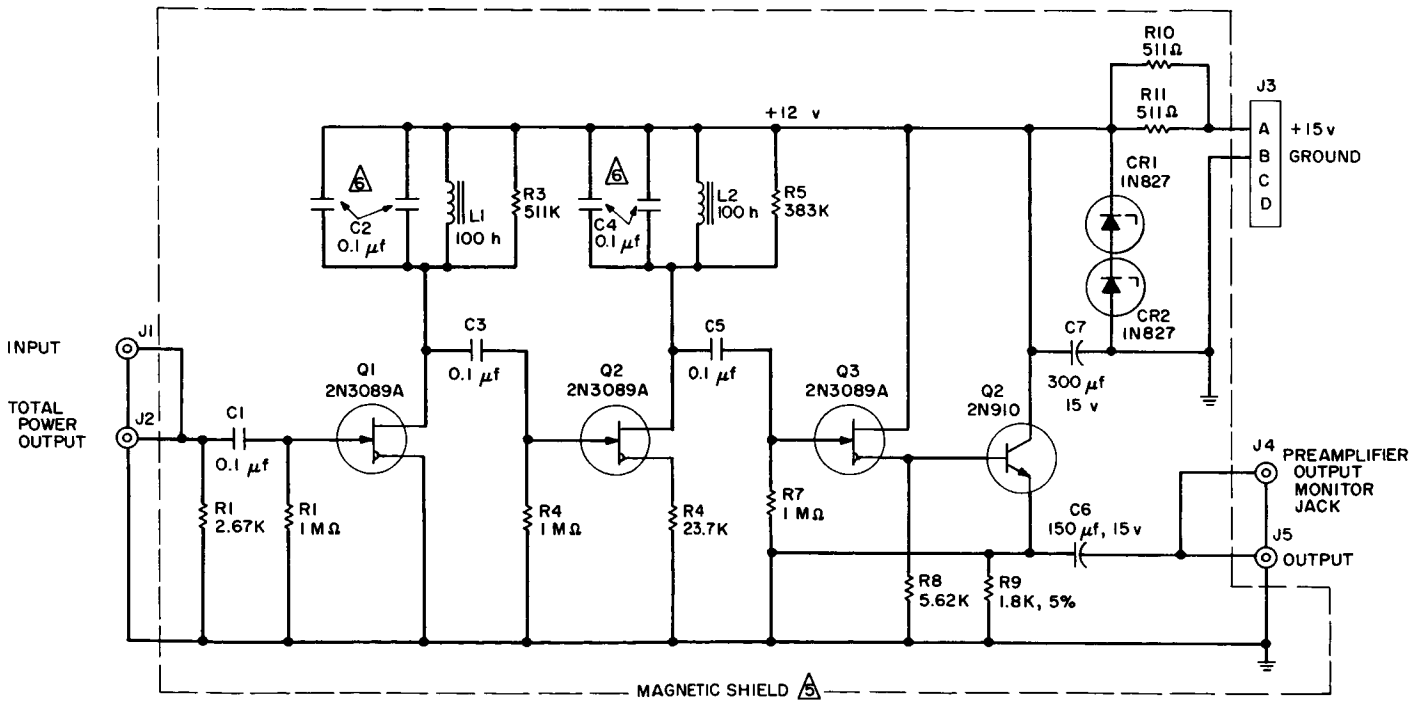


Fig. 5. Low noise preamplifier

C. Optical Communications Components

W. H. Wells

1. Modes of a Tilted Mirror Resonator

As reported previously in Vol. IV of *SPS 37-29, 37-30* and *37-33*, pp. 200, 231 and 254, respectively, we have been solving the diffraction integral equation to find the normal modes of an optical resonator that consists of two tilted plane mirrors facing one another. The mode studies per se were completed in *SPS* 37-33, but it seemed advisable to extend the study to find the shape of the output waveform that spills over the edge of the shorter mirror. Fortunately the output is a form that can be collimated or focused to reasonable accuracy by cylindrical optics without special corrections.

The output is found by repeating the basic integration [Eq. (7) of *SPS* 37-29, Vol. IV, p 203] for negative values of y, which correspond to positions past the edge of the one-dimensional mirror. The previous mode solutions (eigenfunctions ψ and eigenvalues) were used as inputs.

The amplitude distributions are shown in Figs. 6–8 for three values of the tilt parameter β and for modes TEM_0, TEM_1, and TEM_2. In order to make the plots fall close together, to easily interpolate β, a new dimensionless measure of distance was introduced in these graphs, namely

$$z = \frac{y}{2\pi\beta^{3/2}} = \frac{x}{\lambda^{1/4}\, b^{3/4}\, \alpha^{1/2}}$$

where

λ = wavelength,

b = mirror separation,

α = tilt angle (radians),

and

x = distance from output mirror's edge.

The shape of the output of TEM_2 for $\beta = 0.01$ is not as well behaved as it would seem from the plot of Fig. 8, which extends out to $z = -20$. For larger values of

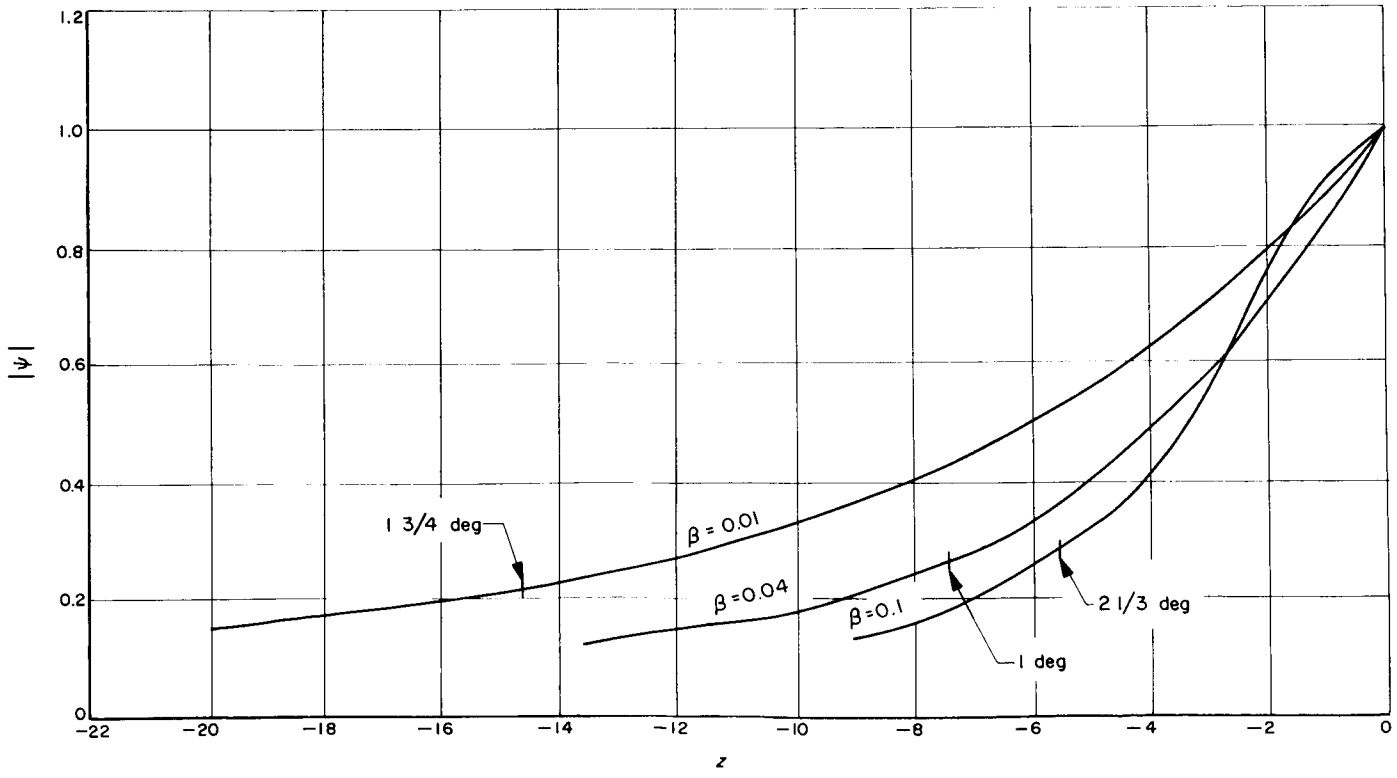


Fig. 6. Output amplitude distribution for the TEM_0 mode and for $\beta = 0.1, 0.04$, and 0.01

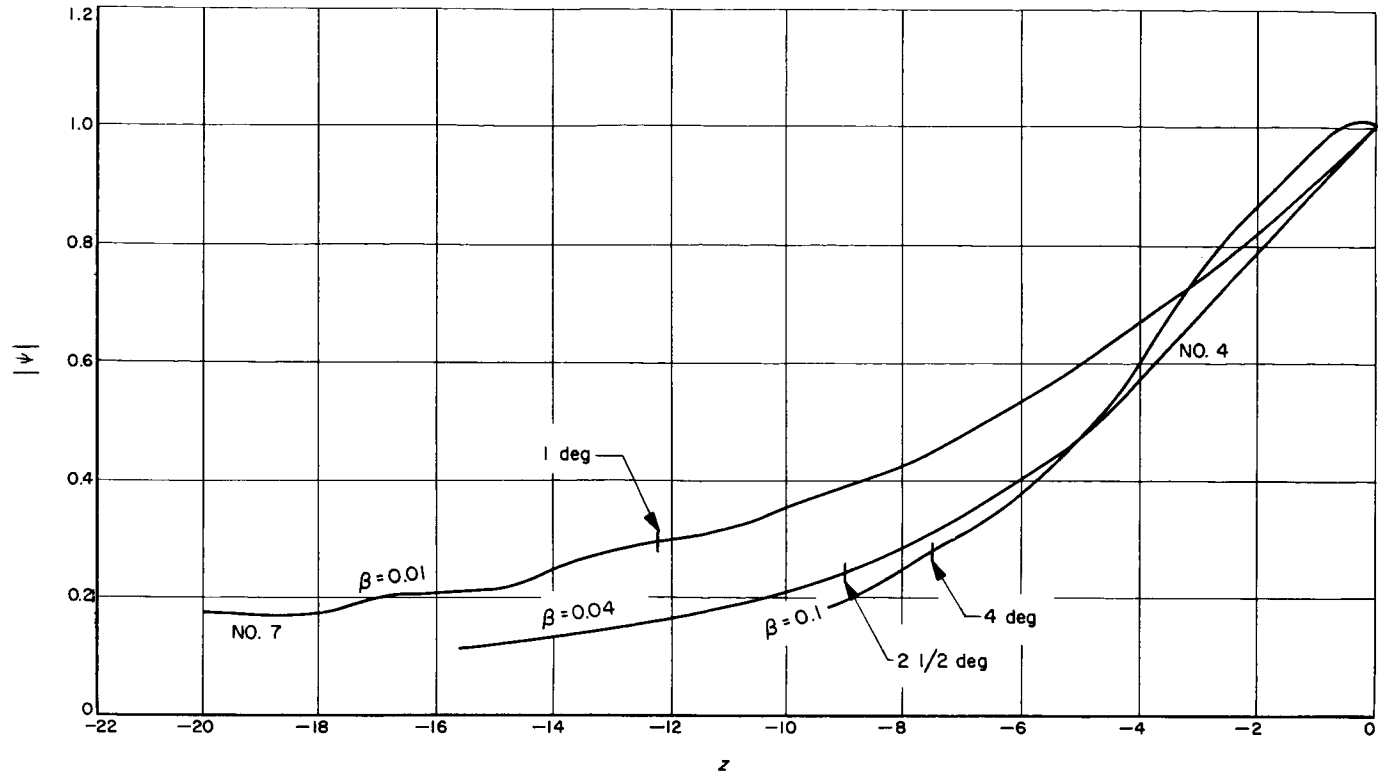


Fig. 7. Output amplitude distribution for the TEM_1 mode and for $\beta = 0.1, 0.04$, and 0.01

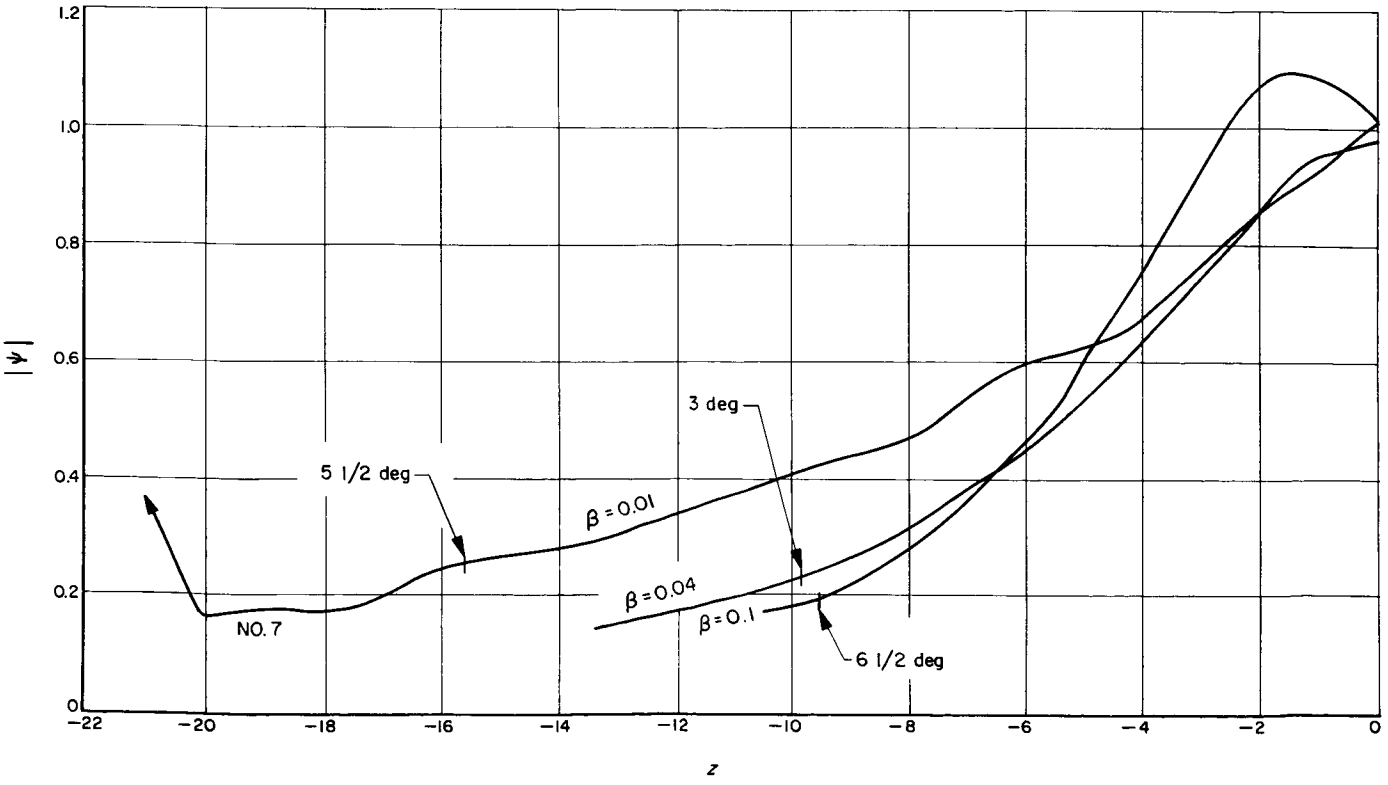


Fig. 8. Output amplitude distribution for the TEM_2 mode and for $\beta = 0.1, 0.04$, and 0.01

$-z$ the amplitude oscillates; then it climbs to a value of 0.64, near $z = 55$, before decaying again.

The phases of the output waves closely fit the quadratic form $\phi = C + Bz + Az^2$ out to values of $-z$, at which the intensity $|\psi(z)|^2$ falls off to 10% or less. This means that the wavefronts are nearly cylindrical and that they can be collimated to good accuracy by cylindrical optics. Figs. 6–8 include markers where the wavefront begins to break away from its cylindrical approximation. The markers are labeled with the maximum phase error caused by the approximation for lower values of $-z$.

The cylindrical wavefronts are described by the position of the virtual line source from which they appear to radiate. Fig. 9 gives $X(\lambda b)^{-1/2}$, where X is the source position in the x dimension. Fig. 10 gives R/b, where R is the source distance behind the output mirror, i.e., the radius of curvature of the wavefront where it spills over the output mirror.

The scatter of points shown on the TEM_2 plots is not due to inaccuracies in the solution of the basic problem but is due, rather, to the lack of a rigorous criterion for curve fitting the actual phase to the cylindrical approximation. Least squares is inappropriate, because the importance of accuracy diminishes with the intensity. A three-point fit was used near $z = 0$, the half intensity point, and the tenth intensity point. Scatter would be less for the smoother TEM_0 and TEM_1 modes.

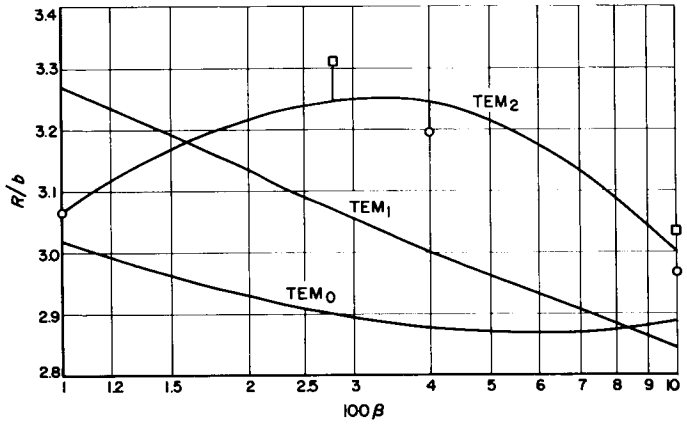


Fig. 9. Output's apparent source: position X in x direction

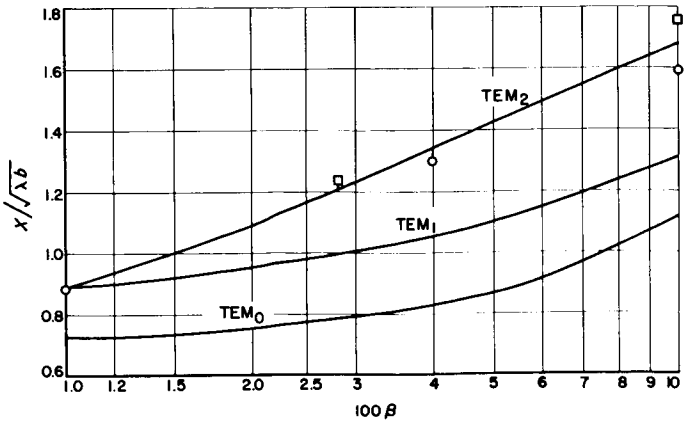


Fig. 10. Output's apparent source: radius of curvature at spillover point

D. Antennas for Space Communications

A. C. Ludwig and R. E. Cormack

1. Shaped Reflector Cassegrainian Antennas,

A. C. Ludwig

a. Summary. A technique for improving Cassegrainian antenna performance by shaping both of the reflector surfaces is evaluated quantitatively. A comparison is made with conventional Cassegranian antenna performance.

b. Introduction. A technique based on geometrical optics has been devised for obtaining uniform amplitude and phase illumination in the aperture of a Cassegrainian antenna (Ref. 3). Efforts have been made to experimentally realize the predicted increase in performance (Ref. 4 and 5). Data on the configuration of one such antenna have been obtained from the WDL Division of the Philco Corporation in Palo Alto, California (Ref. 6) and have been evaluated by JPL computer program.

c. Shaped-Reflector Antenna Performance. The shape of the subreflector synthesized by this method primarily redistributes the energy radiated by the feed to provide uniform amplitude illumination of the main reflector. The shaped subreflector for the Philco antenna is compared to a hyperbola used in conventional antenna geometry in Fig. 11. The phase front of the scattered field is distorted from the nearly spherical phase front scattered from a hyperbola, and the phase must be corrected

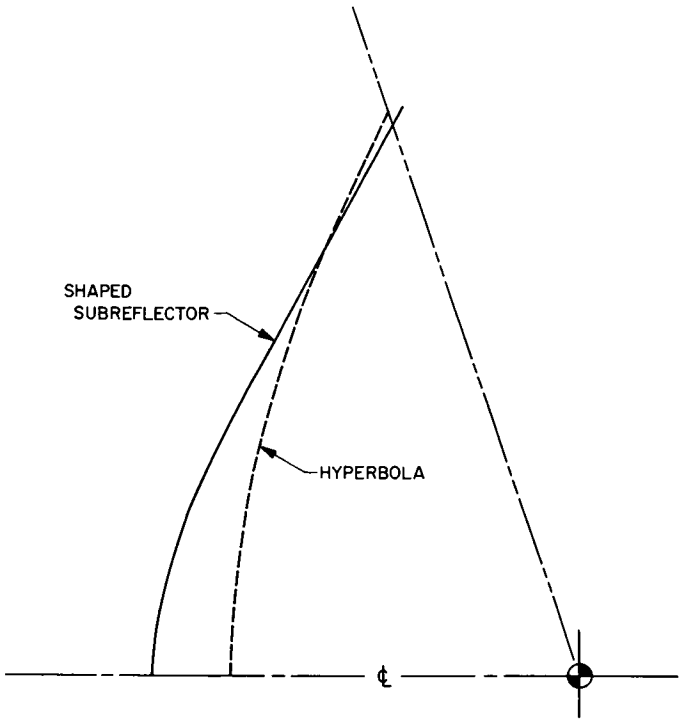


Fig. 11. Shaped subreflector

by the main reflector. The deviation of the main reflector from a best-fit parabola is shown in Fig. 12. It is significant that the distortion is most severe at the edge of the reflector. The Philco antenna has a f/D ratio of 0.35. (The best-fit paraboloid weighed by area has a f/D ratio of 0.33.) For a higher f/D ratio, like the 0.42 f/D of the NASA/JPL advanced antenna system (AAS), the main reflector would be even closer to a paraboloid.

The fields scattered from a 47.75 wavelength diameter shaped subreflector were calculated using a JPL developed computer program (Ref. 7), which has demonstrated outstanding agreement with experimental patterns in past calculations. The scattered fields are shown in Fig. 13 compared to the scattered fields for the AAS feed system (the angle subtended by the main reflector is different in the two cases).

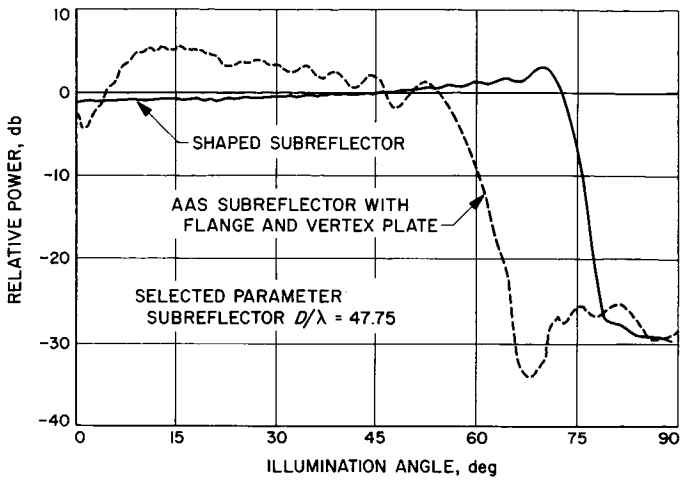


Fig. 13. Fields scattered from subreflector

The scattered fields were then input to another JPL developed computer program (*SPS* 37-26, Vol. IV, pp. 200–208) which evaluates antenna efficiency. The results are shown in Figs. 14 and 15, compared to a similarly calculated performance of the AAS. The angle subtended by the main reflectors was optimized in both cases, for maximum gain in Fig. 14, and for maximum figure of merit in Fig. 15. Figure of merit is defined as the ratio of antenna gain to antenna system temperature; system temperature is taken as antenna noise temperature + 25° K.

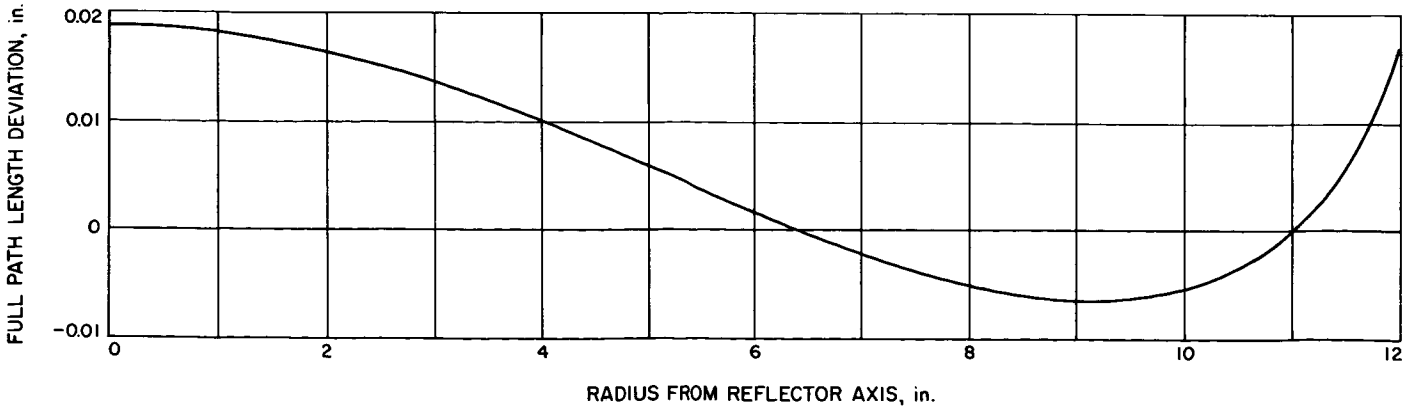


Fig. 12. Deviation of main reflector from best-fit paraboloid

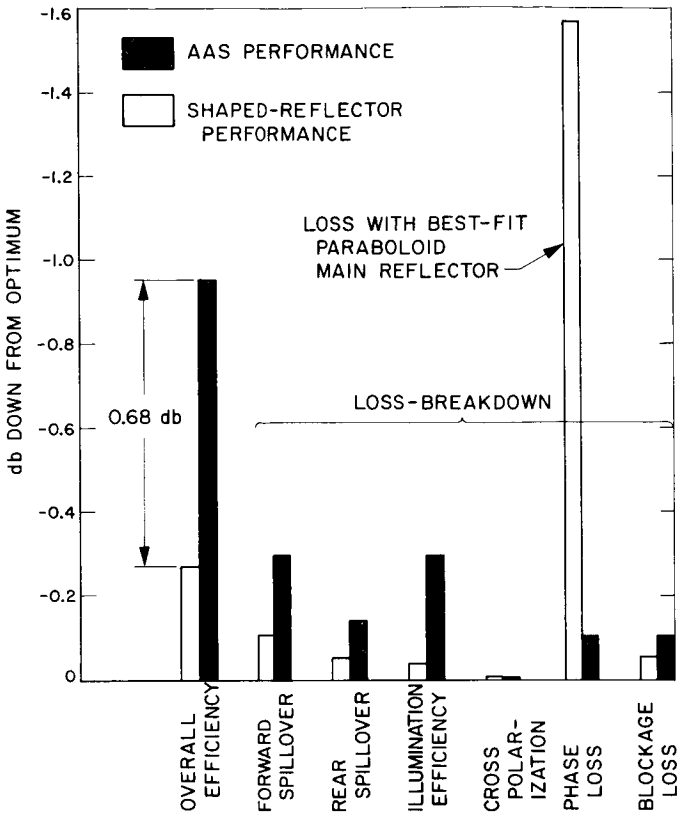


Fig. 14. Antenna performance at maximum gain

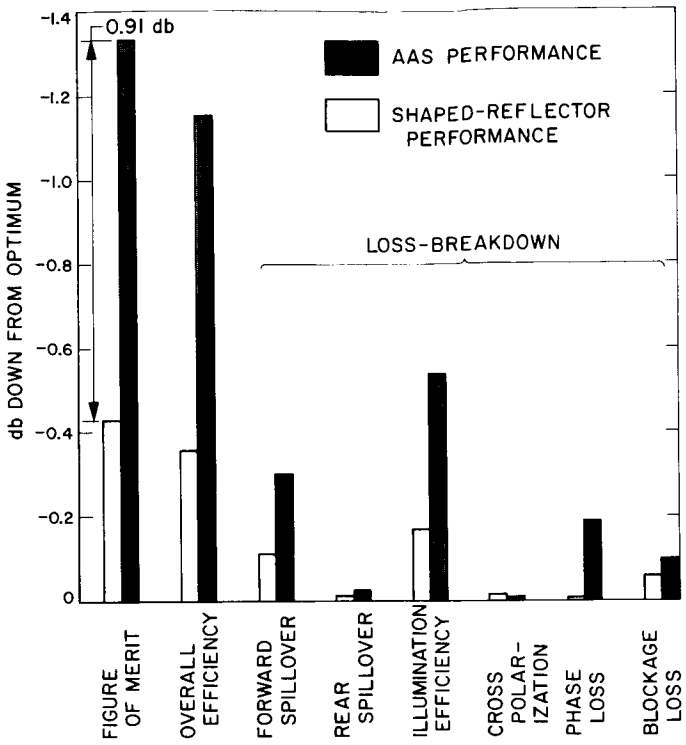


Fig. 15. Antenna performance at maximum figure of merit

The phase error in the shaped reflector system could not be computed with the existing programs, and has been assumed to be zero. In principle an experimentally determined phase pattern of the fields scattered from the subreflector could be used to define the surface of the main reflector, approaching zero phase loss very closely. Blockage loss, as shown in Figs. 14 and 15, was computed on the basis of 7 deg vertex blockage and does not include quadripod blockage. With these assumptions, and neglecting loss due to manufacturing and erection tolerances for the surface, the efficiency of the shaped reflector system at maximum gain is 94%, an extremely high level of performance.

d. Conclusion. The shaped-reflector Cassegrainian antenna has been found to provide an increase in performance of 0.68 db at maximum gain and of 0.91 db at maximum figure of merit when compared to a very well designed conventional Cassegrainian antenna. The calculated efficiency is very near the theoretical maximum for nonsupergain antennas. The shaped reflector technique appears to be an extremely promising method of achieving the maximum potential of a reflector type antenna.

2. R & D Listening Cone for AAS, R. E. Cormack

a. Summary. A multiple purpose R&D listening feed cone is being developed for evaluation and possible long-term use with the AAS. This cone is designed to provide a high level of performance for planetary radar applications (2388 Mc) and for DSIF/Manned Space Flight Network backup applications (2100–2300 Mc) not requiring angle tracking capability.

b. Performance and Characteristics. This cone is designed to support, without modification, the following possible applications.

(1) DSIF diplexed operation (no angle tracking)

(2) Unified S-band *Apollo* support (no angle tracking)

(3) Low noise DSIF listening

(4) *Mariner IV* postencounter listening

(5) Bistatic Mars-Venus planetary radar (2388 Mc)

(6) Monostatic Mars Station planetary radar (2388 Mc)

(7) Depolarization experiments (2388 or 2295 Mc)

(8) Simultaneous DSIF listening and 2388-Mc listening

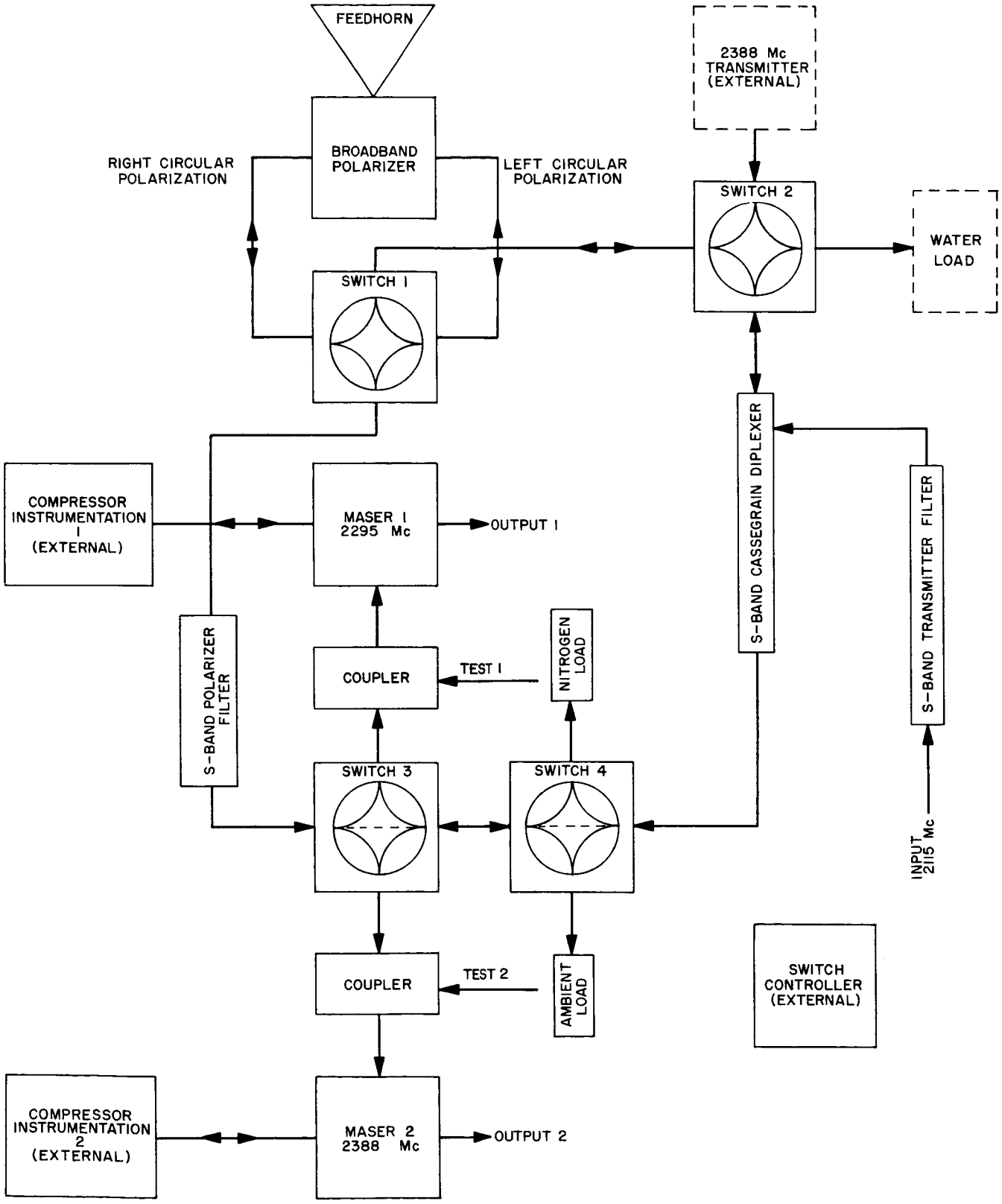


Fig. 16. Mars DSIF R&D listening cone

Fig. 16 is a block diagram of the cone; the basic concept is that there are two reception paths, a low-noise path (via switches 1 and 3) usable at either 2295 or 2388 Mc and a medium noise path (via all four switches) for normal DSIF applications. For special application, either maser may be used with either path.

Potentialities exist to easily increase the diversity and quality of the cone even more in the future in that

(1) High power transmission is available via switches 1 and 2

(2) Space is provided for inclusion of a Dicke switch

(3) With the possible advent of tunable dual frequency masers, replacement of the presently considered single frequency masers will also enhance dual frequency operation by providing simultaneous dual maser operation and some additional polarization diversity

While not obvious from the block diagram, this cone is interchangeable with the standard tracking cone.

c. Some details of components. The feedhorn will be a newly designed dual frequency, dual mode horn with optimum performance at 2295 and 2388 Mc. The broadband polarizer employs a modified (adaptable to circular waveguide) Hughes Aircraft Company orthomode transducer together with a newly designed circular waveguide quarter-wave plate differential phase shifter. All of these items are designed so as not to exclude possible future use at a 400-kw power level.

Both the nitrogen and ambient loads are new R&L designs acceptable for both 2295- and 2388-Mc operation The 2115-Mc reject filter (S-band polarizer filter) procured from Rantec Corporation, is similar in design to the standard DSIF 2295-Mc reject filter [S-band transmitter filter (STF)] with the same low-loss characteristics being obtained at 2295 and 2388 Mc rather than at 2115 Mc, as in the STF. Switch control equipment will involve the addition of external conversion units to the DSIF switch controller, since the cone's functional requirements are in some instances different from those of the standard S-band systems.

The estimated (not a specification) performance of this cone is given in Table 1.

Table 1. Estimated performance

Parameter, Mc	2100	2300	2388
Zenith system noise temperature, low noise channel, °K	—	20–25	20–25
Zenith system noise temperature, medium noise channel, °K	—	35–45	Unknown, not a normal mode
Antenna gain, low noise channel, db	—	61.3	61.5
Antenna gain medium noise channel, db	59.9	61.0	Unknown, not a normal mode
Antenna ellipticity, db	2	0.5	1

E. RF Techniques

W. V. T. Rusch, C. T. Stelzried, B. Seidel and G. S. Levy

1. 90-Gc mm Wave Work, *W. V. T. Rusch and C. T. Stelzried*

a. Summary. The objective of the millimeter wave work is to investigate millimeter wave components and techniques to ascertain the future applicability of this frequency range to space communications and tracking. This involves the development of instrumentation for accurate determination of insertion loss, VSWR, power and equivalent noise temperature of passive elements. Further, in order to bring together state-of-the-art millimeter wave circuit elements and evaluate their use in a system, a radio telescope consisting of a 60-inch antenna and a superheterodyne radiometer has been built (*SPS* 37-33, Vol. IV). The radio telescope was used to observe the 90-Gc temperature of the Moon during the December 30, 1963 eclipse (*SPS* 37-26, Vol. IV, p. 181) and the most recent eclipse, December 18, 1964. These experiments were the joint efforts of personnel from JPL and the Electrical Engineering Department of the University of Southern California.

b. Observations. Radiometric observations of the sub-Earth point on the Moon were carried out daily from August 3 to September 12. The data, which were of considerable scientific interest, provided an excellent test for the long term stability of the radiometer. Thermal calibration of the gas tube before and after each series of

observations yielded information on the long term output of the gas tube. Another useful by-product of the observations was the data on atmospheric absorption. There were indications that the absorption could change significantly during periods of roughly one-hour duration.

All of the data are being studied.

2. *Simultaneous Lobing Radiometric Tracking System,* B. Seidel and G. S. Levy

The S-band systems for the DSIF 85-ft and future 210-ft antennas will use simultaneous lobing angle tracking feed systems. A radiometer, which could be used with the tracking feed, would be a useful device for angle pointing and gain calibrations of the antenna systems using radio star sources. This procedure will become quite important when the 210-ft antenna is completed because this antenna will not have a collimation tower.

Testing of the simultaneous lobing radiometric tracking system has been resumed (*SPS* 37-26, Vol. IV and *SPS* 37-27, 28, 30, 31, Vol. III). On September 1, 1965 the system was installed and used at the Echo Station of the Goldstone Tracking Station using the 85-ft antenna and the S-band receiver system. On this, and subsequent nights, various radio stars were tracked and tracking data were taken. These data are currently being reduced. Fig. 17 shows a sample of the hour angle (HA) residuals obtained by tracking Omega Nebula on the nights of September 22 and 24.

Omega *Nebula* is presently at a declination very close to that of *Mariner IV* and the sidereal hour angle is about 3 hrs behind the *Mariner*. This source will, therefore, provide not only repeatability data and tracking jitter data, but it will also provide a comparison against the *Mariner* tracking data. This comparison will be reported in a future *SPS* as reduced data become available.

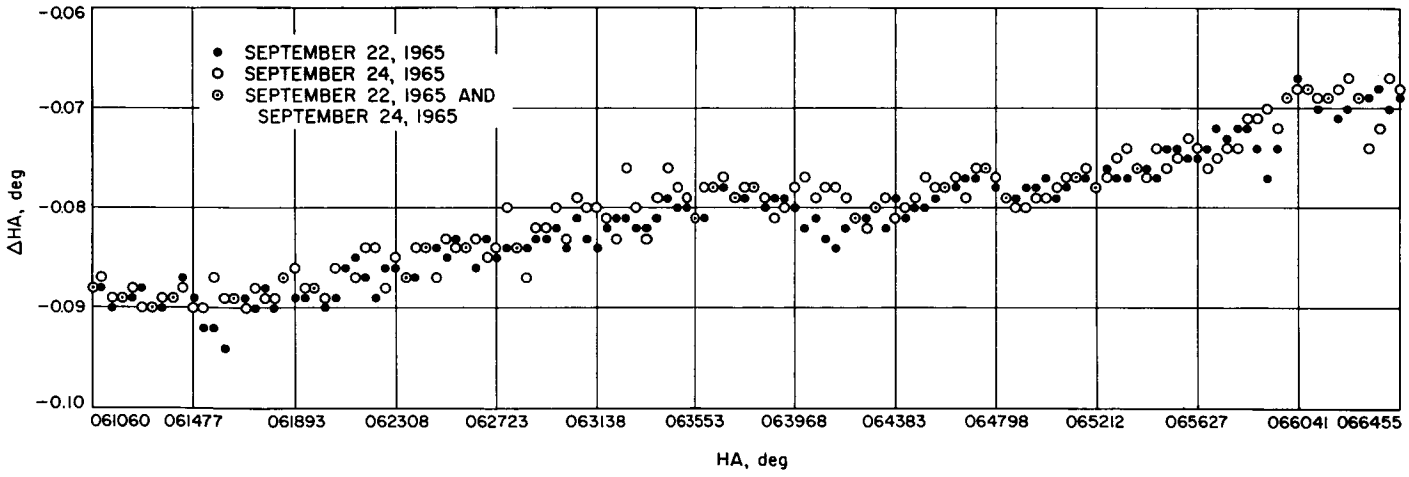


Fig. 17. Auto track on Omega Nebula

References

1. Kleppner, D., Goldenberg, H. M., and Ramsey, N. F., *Physics Review*, 126, 603, 1962.

2. *Hydrogen Maser Program*, Interim Technical Report, Varian Associates, Beverly, Massachusetts (NASA Contract NAS 8-2604), January 1964.

3. Galindo, V., *Synthesis of Dual Reflector Antennas*, Report No. 64-22, Electronics Research Laboratory, University of California, Berkeley, California, July 1964.

REFERENCES (Cont'd)

4. Rouffy, F., "High Efficiency Design Technique for Dual Reflector Antennas," PTGAP International Symposium, pp. 172–180, 1964.

5. Williams, W. F., "High Efficiency Antenna Reflector," *Microwave Journal*, pp. 79–82, July 1965.

6. Zobel, F., private communication, 1965.

7. Rusch, W. V. T., *Scattering of a Spherical Wave by an Arbitrary Truncated Surface of Revolution*, Technical Report No. 32-434, Jet Propulsion Laboratory, Pasadena, May 1963.

XX. Spacecraft Radio

A. Large Aperture Antenna Study

R. M. Dickinson

Large aperture antennas will be required for future spacecraft missions where increased spacecraft antenna gain will be needed to communicate over longer distances or to increase the communications bit rate capability.

This study is directed toward developing an analytical technique for evaluating the performance of large aperture spacecraft reflector type antennas and to obtain design information for performance optimization. A major portion of the study will be concerned with erectile type antennas wherein the reflector is not a simple surface of revolution. One particular form of erectile reflectors consists of segments of curved planes joined together to form the antenna reflector.

A measure of the performance of a large aperture antenna is aperture efficiency. Aperture efficiency can be defined as the ratio of the peak gain of the actual large aperture antenna to the peak gain of a theoretical uniformly illuminated aperture of the same physical size.

The peak gain of a uniformly illuminated aperture is given by

$$G_{max} = \frac{4\pi A}{\lambda^2} \tag{1}$$

where A is the area of the aperture and λ is the wavelength. There exist various methods of calculating the peak gain of reflector type antennas that consist of surfaces of revolution (e.g., paraboloid) illuminated by point source type feeds. However, because the diameter of the launch vehicle limits the diameter of such solid surface spacecraft antennas, the gain calculating method must be general enough to analyze surfaces such as those of erectile reflectors. Because of this requirement, the surface current integration technique, as outlined in Ref. 1, will be used to calculate gain.

The calculation of gain by the surface current integration technique requires information about the surface element of area, the surface normal, and the distance

from the focus to a point on the surface as a function of the coordinate system variables. These surface parameters may be obtained by use of differential geometry (Refs. 2 and 3).

For a surface given by the position vector

$$\boldsymbol{\rho} = \rho(u,v) \tag{2}$$

where the end point of $\boldsymbol{\rho}$ generates the surface, the element of area is

$$dA = (EG - F^2)^{1/2}\, du\, dv \tag{3}$$

and the surface unit normal is

$$\mathbf{n} = \frac{\dfrac{\partial \boldsymbol{\rho}}{\partial y} \times \dfrac{\partial \boldsymbol{\rho}}{\partial v}}{(EG - F^2)^{1/2}} \tag{4}$$

where

$$E = \frac{\partial \boldsymbol{\rho}}{\partial u} \cdot \frac{\partial \boldsymbol{\rho}}{\partial u},\ F = \frac{\partial \boldsymbol{\rho}}{\partial u} \cdot \frac{\partial \boldsymbol{\rho}}{\partial v},\ \text{and } G = \frac{\partial \boldsymbol{\rho}}{\partial v} \cdot \frac{\partial \boldsymbol{\rho}}{\partial v} \tag{5}$$

All position vectors in this discussion will be positioned from the origin of coordinates. In addition, the reflector antenna focus or caustic will be considered to be at the coordinate origin.

The surface position vector may be derived from standard vector techniques. For example, the equation in polar coordinates of a parabola whose focus is at the coordinate origin is

$$r(\theta) = \frac{2F}{1 - \cos\theta}. \tag{6}$$

The corresponding surface position vector in spherical polar coordinates for a paraboloid is

$$\boldsymbol{\rho}(\theta) = r(\theta)\, \mathbf{i}_\rho = \frac{2F}{1 - \cos\theta}\, \mathbf{i}_\rho, \tag{7}$$

where $\mathbf{i}_\rho$ is a unit vector in the radial direction. Additional surface equations for surfaces of revolution may be derived by this technique. However, to obtain segmented erectile reflector surface equations, a different approach must be employed.

The initial segmented surfaces to be evaluated are to consist of a flexible reflecting material pulled uniformly between purely radial ribs, as shown in Fig. 1. The portion of the surface material between a rib pair is called a gore. (Large aperture antenna reflectors can be described as a gored surface of radial ribs.)

To derive the equation of the gore surface refer to Fig. 1. The same z-height or elevation of the surface with respect to the x-y plane may be obtained by either

$$z = \rho \cos\theta, \tag{8}$$

or

$$z = r \cos\theta', \qquad r = |\mathbf{r}_1| = |\mathbf{r}_2|\,. \tag{9}$$

Equating Eqs. (8) and (9) and solving for ρ yields

$$\rho = \mathbf{r}\, \frac{\cos\theta'}{\cos\theta}. \tag{10}$$

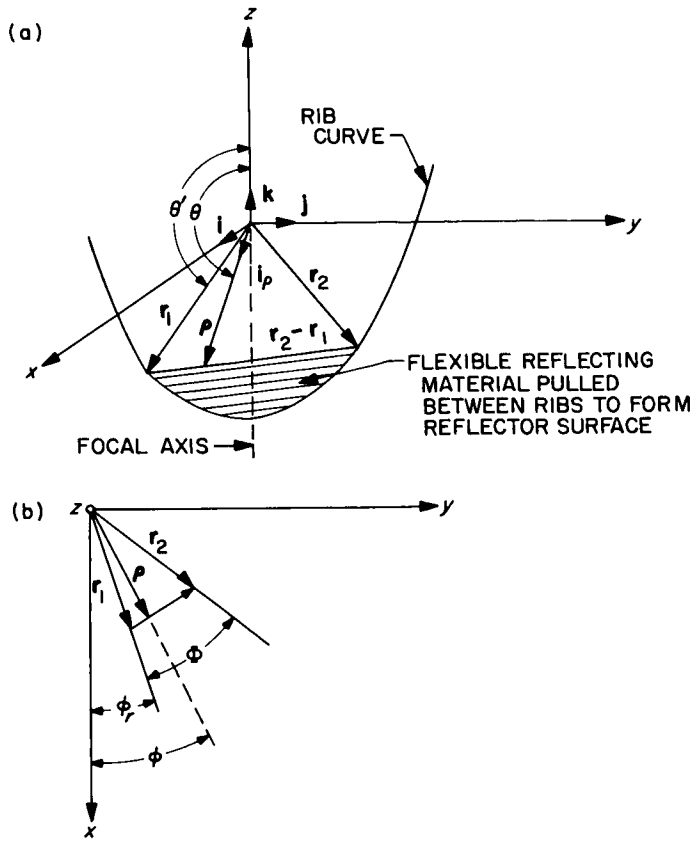


Fig. 1. Gore geometry of a gored surface

Thus, the magnitude of the surface position vector at angle θ is equal to the magnitude of the rib position vector at angle θ' times the ratio of the cosines of θ' and θ.

A connective relation between θ and θ' is given by the triple scalar product

$$\boldsymbol{\rho} \cdot \mathbf{r}_1 \times \mathbf{r}_2 \tag{11}$$

wherein the cross product of $\mathbf{r}_1$ and $\mathbf{r}_2$ defines the normal vector to the plane containing $\mathbf{r}_1$ and $\mathbf{r}_2$. Since $\boldsymbol{\rho}$ also lies in this plane, by definition of the surface, then

$$\boldsymbol{\rho} \cdot \mathbf{r}_1 \times \mathbf{r}_2 \equiv 0. \tag{12}$$

With reference to Fig. 1, substituting

$$\boldsymbol{\rho} = \rho (\sin\theta \cos\phi \mathbf{i} + \sin\theta \sin\phi \mathbf{j} + \cos\theta \mathbf{k}), \tag{13}$$

$$\mathbf{r}_1 = \mathbf{r} (\sin\theta' \cos\phi_r \mathbf{i} + \sin\theta' \sin\phi_r \mathbf{j} + \cos\theta' \mathbf{k}), \tag{14}$$

and

$$\mathbf{r}_2 = r [\sin\theta' \cos(\phi_r + \Phi) \mathbf{i} + \sin\theta' \sin(\phi_r + \Phi) \mathbf{j} + \cos\theta' \mathbf{k}] \tag{15}$$

into Eq. (12) yields

$$\sin\theta \cos\theta' [\sin(\phi_r - \phi) + \sin(\phi - \phi_r - \Phi)] + \sin\theta' \cos\theta \sin\Phi = 0 \tag{16}$$

where Φ is the gore angular width and ϕ is the general azimuth angle between ϕ_r and $\phi_r + \Phi$. θ' and ϕ_r are, in general, functions of s, the arc length along the rib curve.

From Eq. (16)

$$\tan\theta' = \tan\theta \frac{\sin(\phi - \phi_r) + \sin(\Phi + \phi_r - \phi)}{\sin\Phi} \tag{17}$$

Using

$$\cos\theta' = \pm (1 + \tan^2\theta')^{-1/2} \tag{18}$$

or

$$\theta' = \cos^{-1} \pm \left\{ 1 + \tan^2\theta \left[\frac{\sin(\phi - \phi_r) + \sin(\Phi + \phi_r - \phi)}{\sin\Phi} \right]^2 \right\}^{-1/2} \tag{19}$$

from which

$$\frac{\cos\theta'}{\cos\theta} = \left\{ \cos^2\theta + \sin^2\theta \left[\frac{\sin(\phi - \phi_r) + \sin(\Phi + \phi_r - \phi)}{\sin\Phi} \right]^2 \right\}^{-1/2} \tag{20}$$

or

$$\frac{\cos\theta'}{\cos\theta} = (\cos^2\theta + m^2\sin^2\theta)^{-1/2} \tag{21}$$

with

$$m = \frac{\sin(\phi - \phi_r) + \sin(\Phi + \phi_r - \phi)}{\sin\Phi} = \sec\Phi/2\cos(\phi_r + \Phi/2 - \phi) \tag{22}$$

Thus, Eq. (10) becomes in vector form

$$\boldsymbol{\rho} = \left[\frac{\mathbf{r}(s)\cdot\mathbf{r}(s)}{\cos^2\theta + m^2\sin^2\theta}\right]^{1/2}\mathbf{i}_\rho. \tag{23}$$

For the radial ribbed reflectors considered at present, most of the rib curves are monotonic functions of θ'. That is, the rib curve position vector magnitude is a single valued function of θ'. With these restrictions, the magnitude of the rib curve position vector $|\mathbf{r}(s)|$ may be given as a function of θ', rather than as a function of arc length s, along the rib curve. Then, by use of Eq. (19), the rib curve position vector may be obtained in terms of θ.

Eq. (23), with the value of m substituted, becomes

$$\rho(\theta,\phi) = r(\theta')\left[\cos^2\theta + \sin^2\theta\sec^2\Phi/2\cos^2(\phi_1 + \Phi/2 - \phi)\right]^{-1/2}\mathbf{i}_\rho \tag{24}$$

with $\phi_1 = \phi_r$, a constant for radial ribs.

For example, the polar equation of a parabolic rib is given by

$$r(\theta) = \frac{2F}{1 - \cos\theta} \tag{25}$$

where F is the focal length measured from the origin of coordinates of Fig. 1 in the $-Z$ direction. θ is measured from the $+Z$ axis.

Using Eq. (19),

$$r(\theta') = \frac{2F}{1 \pm \left[1 + \sec^2\Phi/2\cos^2(\phi_1 + \Phi/2 - \phi)\tan^2\theta\right]^{-1/2}} \tag{26}$$

and substituting Eq. (26) in Eq. (24) yields

$$\boldsymbol{\rho}(\theta,\phi) = \frac{\pm 2F\sec\theta}{\left[1 + \sec^2\Phi/2\cos^2(\phi_1 + \Phi/2 - \phi)\tan^2\theta\right]^{1/2} \pm 1}\mathbf{i}_\rho \tag{27}$$

as the surface position vector equation for the gore of a reflector whose ribs are purely radial and have a parabolic section.

Similarly, as further examples, for the straight rib shown in Fig. 2

$$r(\theta) = \frac{F \tan \beta}{\sin \theta - \cos \theta \tan \beta}, \tag{28}$$

and

$$\rho(\theta,\phi) = \frac{\pm F \tan \beta}{\sec \Phi/2 \cos(\phi_1 + \Phi/2 - \phi) \sin \theta \pm \cos \theta \tan \beta} i_\rho. \tag{29}$$

For the circular rib of radius a, shown in Fig. 3

$$r(\theta) = (a - F) \cos \theta + [a^2 - (a - F)^2 \sin^2 \theta]^{1/2}, \tag{30}$$

and

$$\rho(\theta,\phi) = \frac{\pm (a - F) + (\sec^2 \Phi/2 \cos^2(\phi_1 + \Phi/2 - \phi) \tan^2 \theta [a^2 - (a - F)^2] + a^2)^{1/2}}{\pm \cos \theta [1 + \sec^2 \Phi/2 \cos^2(\phi_1 + \Phi/2 - \phi) \tan^2 \theta]} i_\rho. \tag{31}$$

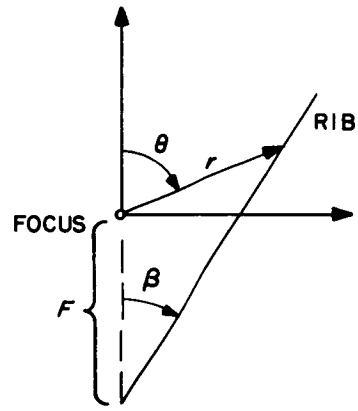


Fig. 2. The straight rib geometry

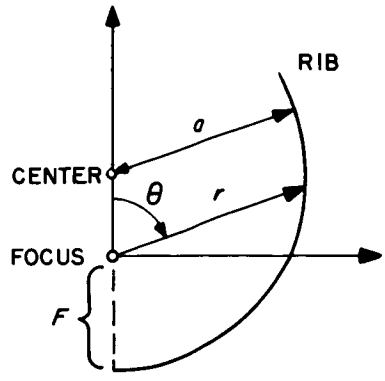


Fig. 3. The circular rib geometry

The distance from the focus to a point on the reflector surface is given by

$$|\rho| = (\rho \cdot \rho)^{1/2}. \quad (32)$$

Eq. (32) along with Eqs. (3), (4), (5), (19), (23), and the surface current integration formulas of Ref. 1 will be employed in a computer program for calculating the aperture efficiency and secondary patterns of various proposed large aperture antennas.

B. Telecommunications Capsule Antenna Systems Study

J. F. Boreham

1. Objective

The capsule antenna systems study was initiated to investigate the problems associated with communications antennas which fulfill the requirements for (1) planetary atmosphere entry missions and (2) planetary landing survival missions. Some of the particular problems to be studied are

(1) the definition of required antenna radiation coverage

(2) effect on antenna characteristics of the capsule near environment, such as impact limiters, capsule structure, rocket motors, parachutes and parachute cases, heat shields, etc.

(3) high impact survival

(4) landed erection and pointing techniques

(5) high power capacity

(6) weight and volume

(7) material and fabrication;

(8) associated transmission line hardware capable of high impact survival.

The study will include investigations into both S-band antennas for direct link communications to Earth and VHF/UHF antennas for relay link communication to Earth via an orbiting spacecraft.

2. History

Since previous work on this project has not been reported in the *SPS* all work to date will be reported here. The initial effort was a study of the antenna implementation of a spinning, planetary atmosphere capsule. This study lasted approximately three months and culminated in the development and testing of a prototype antenna with the required characteristics.

The next phase of the study, an investigation of problems associated with antennas capable of surviving presently projected high impact missions (10,000 g's), is continuing. This is to be a low gain (quadrisphere to hemisphere coverage) circularly polarized S-band antenna.

The first antenna selected for investigation was the split balun, cupped turnstile. Since this type antenna was used as an element of the *Mariner IV* high gain antenna feed array, a flight type turnstile element was chosen for the initial impact tests.

Information gained from these tests led to the design of a second version of the antenna, which is being fabricated. Another inherently rugged antenna, a side wall radiating waveguide which is formed into a one guide-wave length cylinder, or annular slot, is also in the design phase.

3. Spinning Orbiter Capsule Antenna Study

a. Introduction. The objective of this study was to determine the optimum high gain antenna for a spinning orbiter type planetary capsule. The design requirements for this particular implementation were

(1) That the mechanization be highly reliable (mission duration could range up to 400 days)

(2) That the peak gain be greater than 10 db

(3) That both type I and type II trajectories for a typical Mars mission be considered.

This last requirement necessitates that the peak of the beam be optimized at approximately 42 deg from the Sun oriented spin axis of the capsule.

A prime decision made early in the study was to eliminate consideration of all types of implementations which required either excessive complexity, weight or power, if a simple mechanization fulfilling the above requirements was determined feasible. Some of these

more complex implementations involve phased and adaptive arrays and mechanically rotated devices.

An antenna meeting the above requirements is described in Fig. 4. If one half of a parabolic curve, *a c b* (Fig. 4) from the vertex *a* to an end point *b*, is rotated about an axis intersecting the focus *f*, but at an angle α to the parabolic axis, a figure of revolution is generated.

The resulting surface, shown in cross section in Fig. 5, has two symmetric sections above and below point *c*, where the axis of rotation intersects the parabolic curve. Now, if the section of the surface *a′ c a* is eliminated and if the axis of rotation is considered colinear with the spin axis of the capsule, a single surface of revolution *b c b′* exists which will focus energy radiated from the focus into a conical beam whose axis is at an angle α to the spin axis.

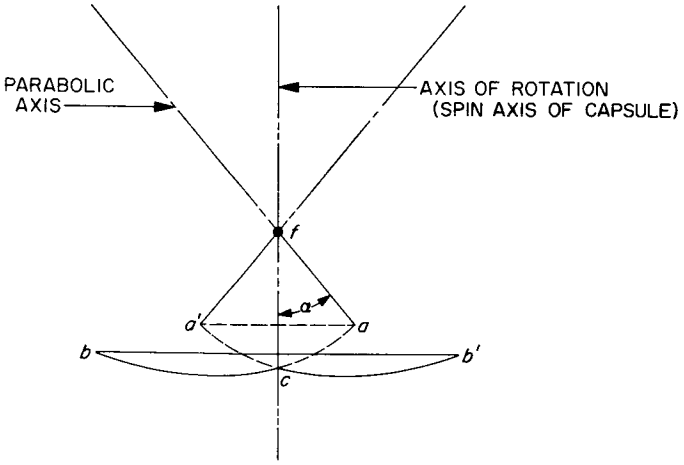


Fig. 4. Antenna nomenclature

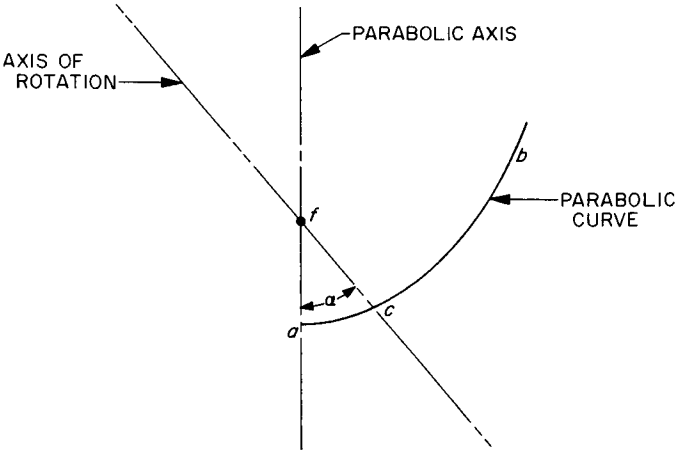


Fig. 5. Surface of revolution of two parts

The 3-db beamwidth of the beam in a plane containing the antenna axis of symmetry (the cone plane) may be related to the reflector diameter (D) by the relation

$$\theta_1 = \frac{116\lambda}{D \cos\alpha}$$

where

λ = wavelength

α = peak of the beam position in the cone plane

and

$$\alpha > \theta_1$$

The peak gain of the antenna may be calculated from the following relation

$$G_p = \frac{k}{\theta_1 \theta_2}$$

where

G_p = power gain

k = an empirical constant

= 31,500 deg^2

θ_2 = 360 deg sin$_\alpha$

hence

$$G_p = \frac{31{,}500}{\theta_1 \times 360 \sin_\alpha}$$

For the 65-in. diameter antenna selected, which meets the look angle requirements for a Mars 1969 mission, for α = 41.5 deg, θ_1 = 12.2 deg, hence the peak power gain

$$G_p = \frac{31{,}500}{12.2 \times 360 \sin 41.5°}$$

$$G_p = 10.82$$

$$G_p = 10.34 \text{ db}$$

The above parameters were obtained empirically from design curves generated from measured data of the two prototype antennas constructed. One of these prototypes, with an α of 39.5 deg and a diameter of 65 in., is shown in Fig. 6. Fig. 7 shows a typical 360 deg cone-cut radiation pattern of the α = 39.5 deg, 65-in. diameter antenna. The radiation pattern shown in Fig. 7 is relative to right hand circular polarization.

Note in Fig. 6 that the reflector surface is constructed by joining two conical surfaces with a flat plate. It is practical to construct the reflector in this manner without deviating from a true parabola by more than 0.3 in. because the reflector is only 3 in. deep, and because the radius of curvature is relatively large. This deviation represents a phase error of less than $\lambda/16$ and, hence, has a negligible effect on gain or the peak of the radiation pattern beam.

Since the reflector can be made of large simply-curved or flat surfaces and since the axis of symmetry of the antenna faces the Sun, the reflector surface can also serve as a mounting surface for solar cells. An investigation into the circuit patterns and the cell filter effects

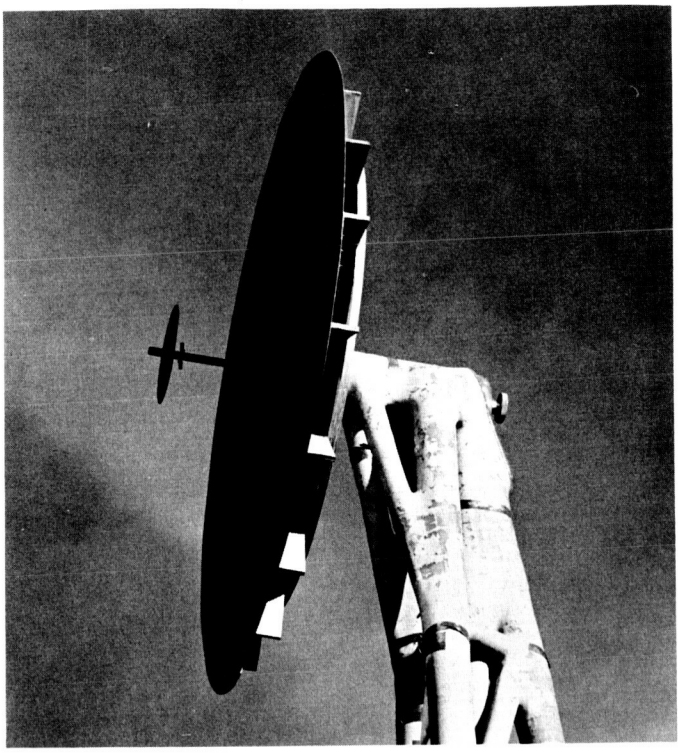


Fig. 6. Conical beam high gain antenna

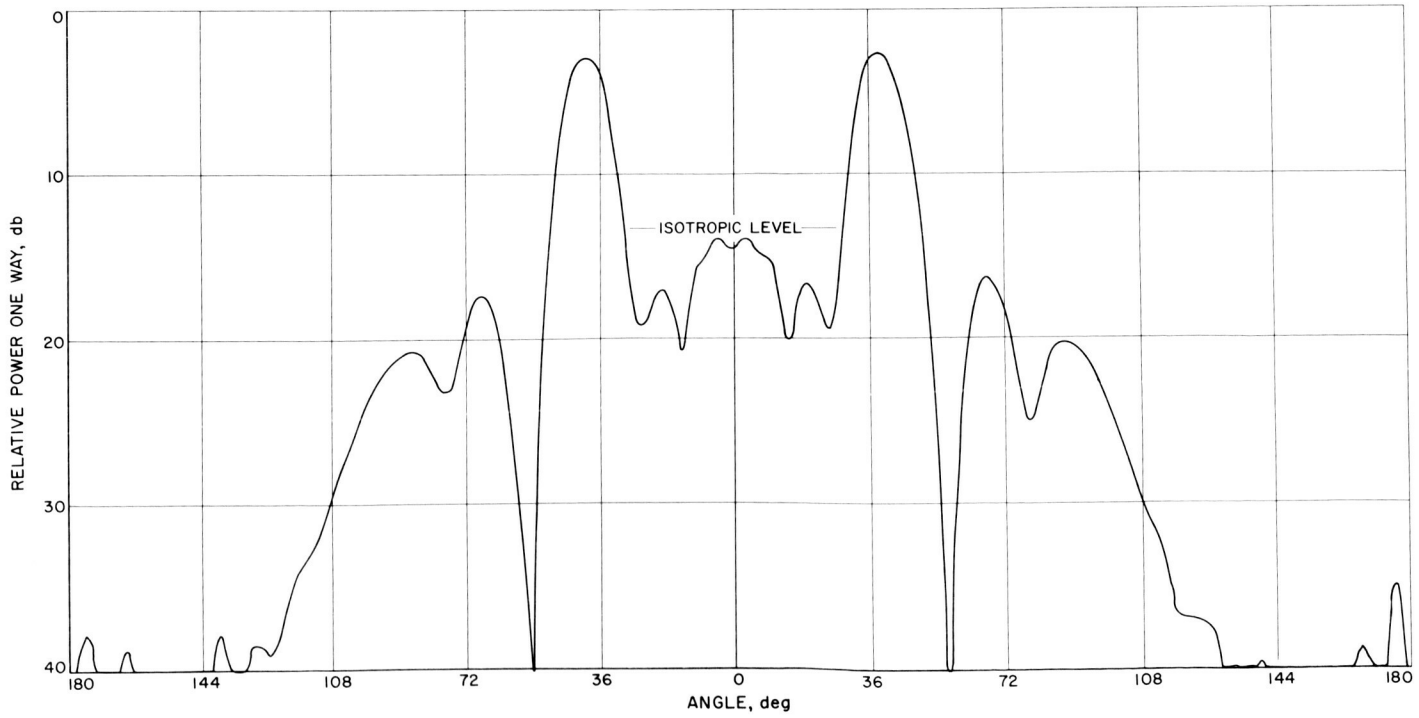


Fig. 7. 65 in., 39.5 deg conical beam antenna radiation pattern

on the RF characteristic of the antenna should, however, be made.

The feed used in the prototype work on this antenna is a simple split balun turnstile over a 2-λ ground plane. This is not an optimum feed design for this antenna since the feed radiation pattern is not perfectly symmetrical circularly. This asymmetry causes a 1.5-db level variation around the peak of the beam. Also, the minimum ellipticity obtainable on the peak of the beams was approximately 3 db.

It is suggested that a complementary cupped ground plane of about 0.7-λ diameter by 0.25-λ deep be substituted for the 2-λ diameter planar ground plane. This type ground plane equalizes the E- and H-plane patterns of the turnstile and hence gives good circular symmetry and low ellipticity over the portion of the beam illuminating the reflector.

4. High Impact Low Gain Antenna Study

a. Introduction. The purpose of this study is to obtain design information for the development of low gain antenna systems with certain electrical characteristics that are capable of surviving the extremely high impact shocks predicted for the early planetary or lunar capsules.

b. Discussion. In order to obtain preliminary design data for a ruggedized cupped turnstile, a *Mariner IV* flight type turnstile was high impact tested. The tests were conducted at JPL's high impact test fixture.

The antenna element survived two 3,000-g shocks oriented along and orthogonal to the center column support tube without visible damage. When the antenna was similarly shocked at 6,000-g's in the direction along the center column no effect was noted. However, when shocked orthogonally to the center column, a small permanent distortion of the center column occurred. Fig. 8 shows the center column deformation. Fig. 9 is a view of the same antenna at 45 deg to Fig. 7. Electrically this damage would have degraded performance by producing a slightly asymmetric beam; yet, it would not have been catastrophic.

This information has been used to design a more rugged turnstile. Work on the design of compatible transmission line components will be started in the near future.

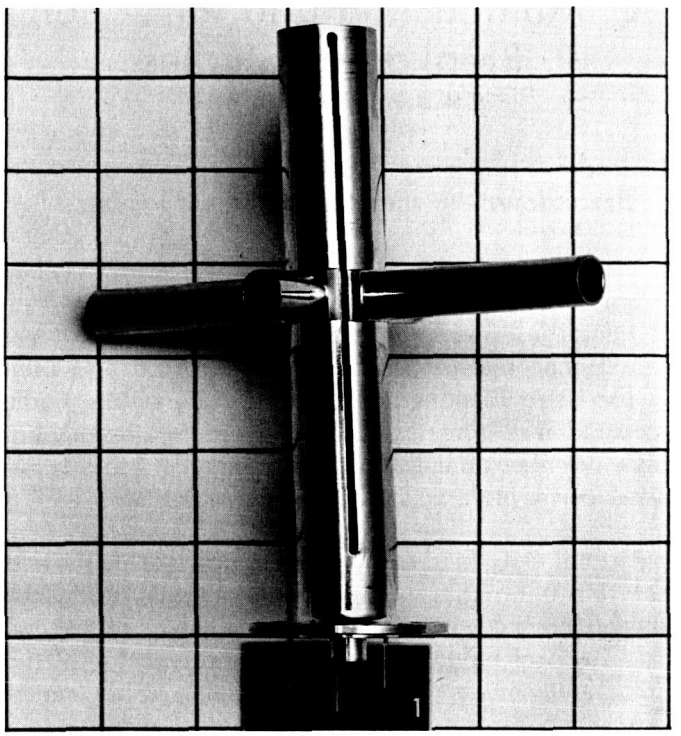


Fig. 8. Turnstile element after 6,000-g shock test

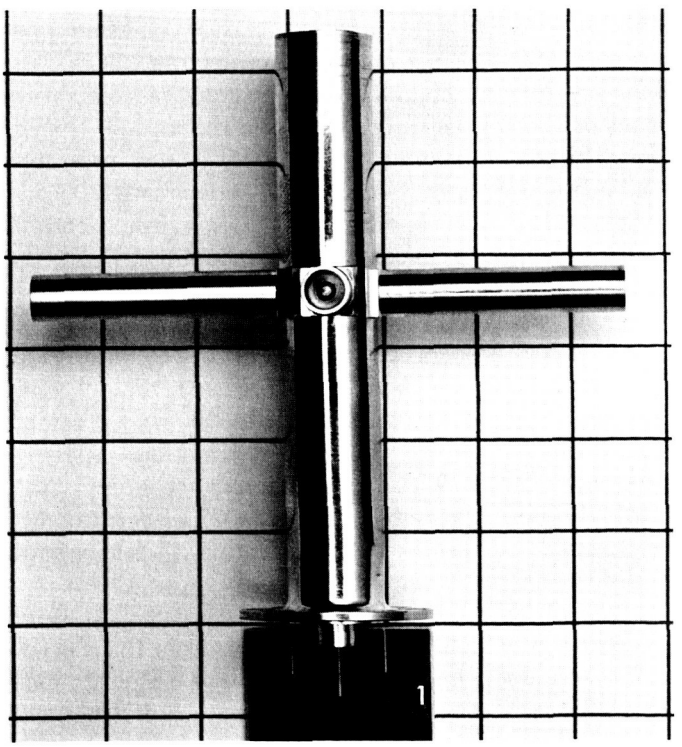


Fig. 9. High impact tested antenna viewed at 45 deg to Fig. 8

C. Multipacting and Ionization Breakdown Studies

S. A. Brunstein and R. Woo

1. Radio Frequency Multipactor Voltage Breakdown on the Ranger Lunar Probe,

S. A. Brunstein

The *Ranger* hard-landing lunar probe was designed to transmit television pictures of the lunar surface from six television cameras for a 10-min period prior to lunar impact. The television system was designed and built by the Astro Electronics Division (AED) of the Radio Corporation of America (RCA). In order to accomplish this mission, communication constraints dictated a transmitter power of 60 w, at a frequency of 960 Mc.

Vacuum tests of the radio frequency system showed that this transmitter power caused a radio frequency voltage breakdown phenomenon called multipacting. The effects of multipacting are large attenuation of the radio frequency signal and physical damage to the microwave components involved. The physical damage renders the components incapable of proper operation, if the breakdown is sustained for a sufficiently long period of time.

Multipacting is a resonant, secondary emission phenomenon that occurs only when electron mean free paths are long compared to the spacing between electrodes or spacing between walls of a microwave component. Thus, multipacting appears only at low pressures. In the case of *Ranger* it occurred at pressures of 10^{-3} torr or less. When this condition exists stray free electrons occurring within a microwave component, upon being accelerated by the electric field, will impact the electrodes or walls. If the electron has gained sufficient energy before impact it will generate secondary electrons, which can be accelerated toward the opposite electrode if properly phased. Thus, if the pressure, electric field intensity, component electrode spacing, and electric field phase are correct, secondary electrons will multiply until space charge effects modify the electric field sufficiently to cause a steady state condition.

The energy which the electrons gain from the electric field is lost upon impact, thus causing a large attenuation of the incident radio frequency energy. Further, the oscillating cloud of electrons causes an impedance mismatch which reflects considerable radio frequency energy back toward the source. The electron impacts with the walls of the component liberate considerable heat, which causes physical damage to the component involved.

A test program, carried out jointly by JPL and RCA-AED, on the *Ranger* microwave system, showed that the type "N" coaxial connectors on the components were the major offenders. Two exceptions were the transmitter power amplifier cavities, and a four port hybrid, where multipacting took place in areas other than the connectors. In the remaining components, even though voids existed where electron oscillation could take place, breakdown occurred only in the type N connector area.

Empirical investigation by JPL and RCA-AED over a period of several months brought about several changes in the system. The type N coaxial cable connectors were changed to JPL modified type "SC" dielectric filled connectors. The transmitter cavity design was changed by RCA-AED to include pressurization. The four port hybrid design was changed by RCA-AED to a solid dielectric stripline version. The connector design on the directional coupler, rotary joint, and parabolic antenna feed was changed by JPL to be compatible with the modified SC cable connectors and to eliminate all voids where multipacting had been shown to occur.

Further testing showed the design changes to be sufficient for the elimination of multipaction for this mission. There still exists, however, a great lack of basic and engineering information on this phenomenon that could cause further severe communication problems with other space flight programs.

2. Multipacting Breakdown Study, R. Woo

a. Introduction. As a result of the multipacting problem encountered in the *Ranger* Project (see preceding article by S. A. Brunstein and also *SPS* 37-21, Vol. I) a theoretical and experimental study has been started. With a better theoretical understanding of the phenomena, we will be able to design microwave components for high power space applications that will not be susceptible to this type of breakdown.

The study phase that is reported here was carried out with the assistance of Dr. Akira Ishimaru of the University of Washington, consultant to JPL under contract DH5-351654.

b. History. Secondary electron resonance breakdown was observed and investigated as early as 1924 (Refs. 4 and 5). However, a significant mathematical formulation was not made until 1948 when Gill and Von Engel (Ref. 6) introduced the assumption of a constant v_f/v_i ratio in their analysis. (v_i is the initial electron emission velocity while v_f is the arrival velocity at the opposite electrode.) Their argument was that both v_i and v_f were approximately constant and consequently their ratio was also constant. A cutoff law was introduced and general agreement between experiment and theory was shown.

Hatch and Williams (Refs. 7 and 8) used the constant v_f/v_i assumption in extending Gill and Von Engel's results. However, along the minimum breakdown voltage curve, the relationship that v_i is proportional to frequency f and the electrode separation d was employed. Since fd varies, Gill and Von Engel's justification for the constant v_f/v_i assumption was no longer valid. In a later paper, Hatch (Ref. 9) indicated that the assumption was made for mathematical simplicity. Despite the lack of a physical explanation for the constant v_f/v_i assumption, Hatch and Williams' analysis was quite successful. In terms of the external parameters of breakdown voltage, applied frequency and separation distance, simply described breakdown curves resulted. Moreover, these curves could be satisfactorily fitted to the experimental data, as seen in Fig. 10.

Criticism of the controversial constant v_f/v_i assumption led to a subsequent study by Miller, Williams and Theimer (Ref. 10) of computer-calculated trajectories. They demonstrated that the electrons preferred certain values of v_f/v_i, but no light was shed on the physical reason.

Krebs and his associates (Refs. 11 and 12) and others (Refs. 13–15) conducted multipacting discharge studies without making the constant v_f/v_i assumption. Instead, either a constant velocity or a distribution of velocity was assumed and more general and detailed calculations for various values of the parameter of time phase angle of secondary emission were made. In terms of the experimental parameters, the results showed no significant deviation from those obtained by Hatch and Williams. More important, however, was the loss of Hatch and Williams' attractive feature of simplicity.

It was felt that some basic physical reason underlaid the success of the constant v_f/v_i assumption. It was also believed that this assumption, along with the simple

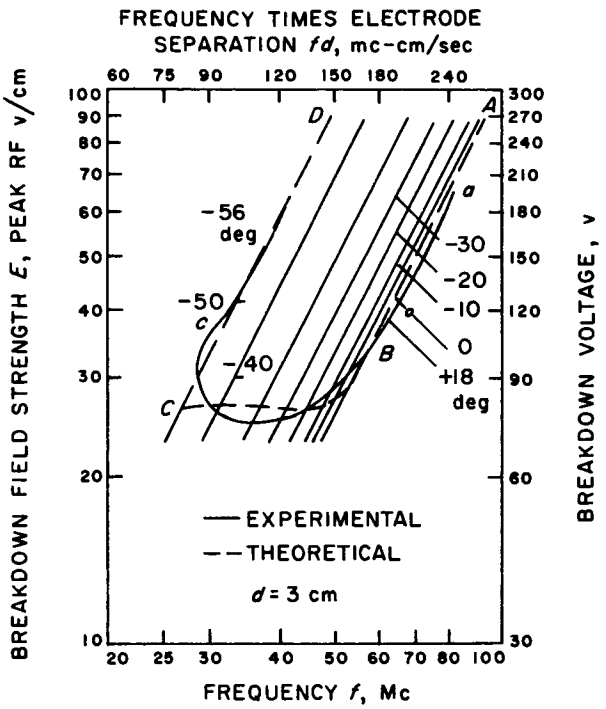


Fig. 10. Two-sided multipacting region for parallel plates geometry

results of Hatch and Williams, were not restricted to just parallel plates. Of course, of special interest is the coaxial case for microwave circuits.

The portion of study reported here was undertaken with two objectives. First, to establish some basic physical reason for the success of predicting breakdown regions using the constant v_f/v_i assumption. Second, to seek a means of extending this simple analysis to geometries other than parallel plates. To accomplish this, a similarity principle was introduced.

c. A similarity principle for multipacting. Similarity relations have been quite popular for the analysis of gas discharges (Ref. 16) as well as for analysis in other branches of physics and engineering. In deriving these relations, very often a complex analysis is avoided and much pertinent information is gained through a minimum of effort. In multipacting discharges, such relations would prove valuable.

Since the phase relationship between the electron and the electric and magnetic fields is a prime controlling factor of the multipacting mechanism, we will look for similarity conditions in this vein. The phase relationship

of the electron with the fields is governed by the equation for charged particle motion

$$m\frac{d^2\mathbf{r}}{dt^2} = e\,(E + \mathbf{v} \times \mathbf{B}) \tag{1}$$

where

$\mathbf{E}$ is the electric field vector

$\mathbf{B}$ is the magnetic flux density vector

$\mathbf{v}$ is the velocity vector

e is the electron charge

m is the electron mass

$\mathbf{r}$ is the displacement vector

t is the time

In general, both $\mathbf{E}$ and $\mathbf{B}$ are given as a sum of dc and RF components. The dc electric and magnetic fields can be obtained independently, but the RF components are related to each other through Maxwell's equations. We write

$$\begin{aligned}\mathbf{E} &= \mathbf{E}_{dc} + \mathbf{E}_{rf}\\ \mathbf{B} &= \mathbf{B}_{dc} + \mathbf{B}_{rf}\end{aligned} \tag{2}$$

The ratio of magnetic intensity H_{rf} to E_{rf} is of the order of $\left(\frac{\epsilon}{\mu}\right)^{1/2}$ (where ϵ is the dielectric constant and μ is the permeability). We also know that B_{rf} is related to H_{rf} by $B_{rf} = \mu H_{rf}$. Therefore, the RF component of the second term of Eq. (1), as compared with the first term, is of the order of

$$\frac{|\mathbf{v} \times \mathbf{B}_{rf}|}{|\mathbf{E}_{rf}|} = \frac{|\mathbf{v} \times (\mu\epsilon)^{1/2}\mathbf{E}_{rf}|}{|\mathbf{E}_{rf}|} = \frac{|\mathbf{v}|}{c}$$

where c is the velocity of light.

Therefore, unless $\frac{|\mathbf{v}|}{c}$ is significantly close to unity, we can neglect $\mathbf{B}_{rf}$ in which case we would be left with

$$m\frac{d^2\mathbf{r}}{dt^2} = e\,(\mathbf{E} + \mathbf{v} \times \mathbf{B}_{dc}) \tag{3}$$

This equation can be rewritten in dimensionless form by normalizing each variable. We let

$$\begin{aligned}\mathbf{r} &= r_c\mathbf{r}^* & t &= t_c t^*\\ E &= E_c\mathbf{E}^* & \mathbf{B} &= B_c\mathbf{B}_{dc}\end{aligned} \tag{4}$$

where the variables with subscript c have the same dimension as the unstarred variables, and their values can be chosen appropriately. The starred quantities are dimensionless. Substitution of Eq. (4) into Eq. (3) yields

$$\frac{d^2\mathbf{r}^*}{dt^{*2}} = \frac{e}{m}\,[C_1\mathbf{E}^* + C_2\mathbf{v}^* \times \mathbf{B}^*_{dc}] \tag{5}$$

where C_1 and C_2 are both dimensionless quantities and are given by

$$C_1 = \frac{t_c^2 E_c}{r_c} \tag{6}$$

$$C_2 = B_c t_c \tag{7}$$

and the dimensionless boundary conditions are determined from

$$\mathbf{r}^* = \frac{1}{r_c}\mathbf{r} \tag{8}$$

$$\frac{d\mathbf{r}^*}{dt^*} = v^* = \frac{t_c}{r_c}\frac{d\mathbf{r}}{dt} \tag{9}$$

For identical solutions of Eq. (5), it is clear that the boundary conditions determined from Eqs. (8) and (9), in addition to Eqs. (6) and (7), must remain unchanged. Thus, Eqs. (6)–(9) determine the similarity relations for the various variables.

The implications of Eqs. (6)–(9) are interpreted in terms of voltage V, velocity v, magnetic flux density B_{dc}, frequency f, and characteristic length l. Voltage is defined as $V = \int \mathbf{E} \cdot \mathbf{dr}$ and may be RF, dc or a combination of both; characteristic length l is any length that characterizes one of a family of configurations that differ only in scale. The normalized time variable is the phase angle of the electron.

There are two different situations, one in which the variables are unprimed and the other in which they are primed. If Eq. (6) is invariant,

(1) Voltage must be proportional to $(fl)^2$

Therefore,

$$\frac{V}{(fl)^2} = \frac{V'}{(f'l')^2} = \text{constant} \tag{10}$$

If Eq. (7) is invariant,

(2) dc magnetic flux density must be proportional to f

Therefore,

$$\frac{B_{dc}}{f} = \frac{B'_{dc}}{f'} = \text{constant}$$

We interpret the boundary conditions in terms of the two multipacting surfaces and the initial and final velocities of the electrons v_i and v_f, respectively. From Eq. (7) we have

(3) the two separate configurations must be geometrically similar.

From Eq. (9) we have

(4) the velocity must be proportional to fl

$$\frac{v}{fl} = \frac{v'}{f'l'} = \text{constant} \tag{11}$$

In particular, at the boundary

$$\frac{v_i}{fl} = \frac{v_{i'}}{f'l'} = \text{constant} \tag{12}$$

$$\frac{v_f}{fl} = \frac{v_{i'}}{f'l'} = \text{constant} \tag{13}$$

or

$$\frac{v_f}{v_i} = \frac{v_{f'}}{v_{i'}} = \text{constant} \tag{14}$$

These relations may be alternately expressed in terms of kinetic energy. In particular, Eq. (13) may be written as

$$\frac{\left(\frac{1}{2} m v_f^2\right)}{(fl)^2} = \frac{\left(\frac{1}{2} m v_{f'}^2\right)}{(f'l')^2} = \text{constant} \tag{15}$$

If the two situations are to be similar in electron arrival energy as well as in phase, there is the additional condition that

$$fl = f'l' \tag{16}$$

We are now in a position to state our similarity principle for multipacting discharges. Multipacting discharges similar in phase are defined as multipacting discharges in geometrically similar configurations where electrons at geometrically similar points possess the same phase relationship with the electric and magnetic fields. The similarity conditions are that the quantities $\frac{V}{(fl)^2}$, $\frac{B_{dc}}{f}$ and $\frac{v_f}{v_i}$ (where V may be RF, dc or both) must remain invariant. If the multipacting discharges are to be similar in electron arrival energy also, we have the additional similarity condition that fl must be invariant.

It is proper here to point out the difference between the scaling relationships obtained by Stratton (Ref. 17) in scaling electromagnetic fields and those of our similarity principle. In scaling electromagnetic fields, if μ and ε are unchanged and conductivity σ is zero, fl must be invariant. Of course this is the well known condition used in scaling antenna models. In scaling multipacting discharges, however, the electron trajectories must also be considered. The result is scaling information not only for two physical structures but for a single structure with frequency varied as well.

In the following sections, the similarity principle is applied to two-sided multipacting in parallel plates geometry and the implications are discussed.

d. Two-sided multipacting in parallel plates geometry. Let us consider a sinusoidal voltage with frequency $f = \omega/2\pi$ applied to two parallel plates located at $s = 0$ and $s = d$, where s is perpendicular to the plates. The equation of motion of an electron is given by

$$\frac{d^2s}{dt^2} = \frac{e}{m} \cdot \frac{V}{d} \sin \omega t \tag{17}$$

and the boundary conditions are as follows:

$$\text{Initial} - \text{ at } \omega t = \phi_0,\ s = 0 \text{ and } \frac{ds}{dt} = v_i$$
$$\text{Final} - \text{ at } \omega t = \phi_1,\ s = d \text{ and } \frac{ds}{dt} = v_f \tag{18}$$

The desired phase synchronization for multipacting occurs when

$$\phi_1 = \phi_0 + (2n - 1)\,\pi \tag{19}$$

where $n = 1, 2, 3 \ldots$.

According to the procedure described in the previous section, we normalize s and t as follows:

$$\frac{s}{d} = x \tag{20}$$

$$\omega t = \phi \tag{21}$$

This normalization results in phase similar solutions $x(\phi)$, if both $\frac{V}{(fd)^2}$ and $\frac{v_f}{v_i}$ remain invariant. If the solutions are to be similar in electron arrival energy as well, fd must also remain invariant. As a convenience for plotting the solution in the phase plane depicted in Fig. 11, we normalize v as $v/v_i = y$. In this manner, all similar solutions are represented by one plot only.

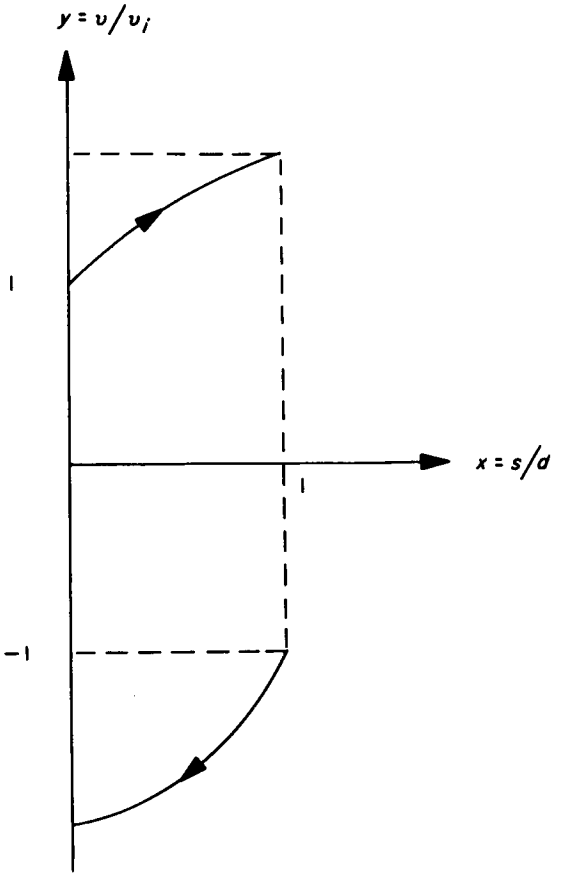


Fig. 11. Phase plane plot for two-sided multipacting in parallel plates geometry

Fig. 10 shows Hatch and Williams' theoretically-fitted, experimental, closed curve (Ref. 4). The specific relationships they derived and used were

$$\frac{V}{(fd)^2} = \frac{(2\pi)^2}{\frac{e}{m}\left[\frac{k+1}{k-1}\,\pi\cos\phi_0 + 2\sin\phi_0\right]} \tag{22}$$

and

$$v_f = \frac{k}{k-1}\,\frac{2e}{m}\,\frac{V}{\omega d}\cos\phi_0 \tag{23}$$

where $k = \frac{v_f}{v_i}$

In Fig. 10, BC represents the starting breakdown voltages corresponding to an electron arrival energy of 60 ev. For voltages higher than these defined by BC, the electron arrival energies are correspondingly higher and therefore sufficient for multipacting. For this region then, the boundaries AB and CD are determined from phase considerations. These two types of boundaries will be discussed separately.

(i) Phase-controlled boundary. Along AB and CD or any other constant ϕ_0 line, Hatch and Williams used Eq. (22) for fitting. The relationship for initial velocity can be obtained by substituting Eq. (22) into Eq. (23); the result is that the initial velocity is proportional to fd. It is now clear that these relations are precisely the phase similarity conditions of our similarity principle. Needless to say, where fd is similar the electron arrival energy is also similar. Some comments applicable to all similar situations will be made.

From the similarity principle, it is obvious that similar points have the same constant k. However, there is no reason that k should be the same constant throughout the entire region. For example, on one hand two separate points along AB are similar and have the same

k, $k = k_1$. On the other hand, two separate points along CD are also similar and have the same k, $k = k_2$. However, since points along AB are not similar to those along CD, k_1 and k_2 need not be the same. In fact, as shown by the computer studies of Miller and his coworkers (Ref. 10), $k_1 \sim 2$ while $k_2 \sim 3$. It should also be mentioned that since k need not have the same value throughout the entire multipacting region, including the higher order multipacting modes, the overlapping regions criticized by Kossel and Krebs (Ref. 12) could be eliminated.

Hatch and Williams assigned to AB and CD single semiempirical values of ϕ_0 and k. In a later paper, Hatch (Ref. 9) termed this the "thin sheet" assumption, since the participating electrons may be considered as crossing the gap in a very thin sheet without phase spreading. Actually there is phase spreading among the electrons such that a phase distribution (Ref. 18) and, consequently, a k distribution also exist. However, since the similarity principle applies to each individual electron in the phase distribution, the entire phase distribution remains the same for similar breakdown points. This, indeed, is the most significant result of the similarity principle; for in terms of the external parameters, the only condition is that $V/(fd)^2$ remain constant. That the phase distributions are similar for similar breakdown points is also demonstrated in the computer studies of Miller and his coworkers (Ref. 10).

According to the similarity relations, the entire initial velocity distribution of the electrons participating in the multipacting process is proportional to fd. This is in contrast to the fact that the secondary electron emission velocity distribution is generally independent of fd and the energy of the primary electrons. In view of the good correlation of experiment and theory for AB and CD, as shown in Fig. 10, it can be concluded that the phase similarity conditions, so important to the multipacting process, take precedence over the velocity distribution of secondary electron emission. In other words, when fd is varied, the multipacting process will sort out participating electrons whose v_i is proportional to fd, in order to maintain similar proper phase relations.

(ii) Minimum electron arrival energy boundary. Separate points along the minimum electron arrival energy boundary BC are not similar and consequently need not possess the same value of k. However, the assumption of a constant k is still relatively satisfactory, since v_f is constant and v_i is approximately constant. This then represents the situation close to that studied by Gill and Von Engel (Ref. 6).

The similarity principle may be applied to each point along BC, since the electron arrival energy is constant for an invariant fd. In particular, when applied to point C, Gill and Von Engel's important cutoff condition, that the cutoff wavelength be proportional to ϕ, is obtained.

In determining each point of BC, Hatch and Williams assigned single values of ϕ_0 and k to compute the voltage level corresponding to the minimum energy requirement. Since a phase distribution actually exists for each point, it is realized that the minimum energy value is effectively assigned to the entire phase distribution. The thin sheet assumption is therefore more substantial in this case. This along with the assumption of a constant k value for all of BC makes it obvious why correlation of the experimental data for BC was not as satisfactory as that for AB and CD.

Although our attention has been confined to the 1/2 cycle mode, all discussions in (i) and (ii) apply to the higher order modes as well (Ref. 8).

To summarize, we have introduced a similarity principle for multipacting. This yields a physical interpretation of a controversial assumption made by previous workers in the field. More significant, however, is its applications to geometries other than parallel plates.

References

1. Silver, S., *Microwave Antenna Theory and Design*, McGraw-Hill Book Co., New York, New York, 1949, pp. 144–155.

2. Lass, H., *Vector and Tensor Analysis*, McGraw-Hill Book Co., New York, New York, 1950, Chapter 3.

References (Cont'd)

3. Eisenhart, L. P., *A Treatise on the Differential Geometry of Curves and Surfaces*, Ginn & Co., Boston, Massachusetts, 1909.

4. Gutton, C., *Compt. Rend.* 178, 467 (1924); Gutton, C., and Gutton, H., *Compt. Rend.* 186, 303 (1928); Gutton, H., *Ann. Physik* 13, 62, 1930.

5. Kirschner, F., *Ann Physik* 77, 287 (1925); Kirschner, F., *Ann. Physik* 7, 798, 1930.

6. Gill, E. W. B., and Von Engel, A., "Starting Potentials of Electrodeless Discharges," *Proceedings of the Royal Society of London*, Vol. 192, pp. 446–463, 1948.

7. Hatch, A. J., and Williams, H. B., "The Secondary Electron Resonance Mechanism of Low Pressure High Frequency Breakdown," *Journal of Applied Physics*, Vol. 25, pp. 417–423, 1954.

8. Hatch, A. J., and Williams, H. B., "Multipacting Modes of High Frequency Gaseous Breakdown," *Physical Review*, Vol. 112, pp. 681–685, 1958.

9. Hatch, A. J., "Electron Bunching in the Multipacting Mechanism of High Frequency Discharge," *Journal of Applied Physics*, Vol. 32, pp. 1086–1092, 1961.

10. Miller, A., Williams, H. B., and Theimer, O., *Journal of Applied Physics*, Vol. 34, pp. 1673–1679, 1962; Miller, A., "Experimental and Theoretical Investigations of the Secondary Electron Emission Phase Angle Distributions in a Low Pressure Multipacting Electrical Discharge," Ph.D. dissertation, New Mexico State University, University Park, New Mexico, 1961.

11. Krebs, K., and Meerbach, H., "Die Pendelvervielfachung von Sekundärelektronen," *Annalen der Physik*, Vol. 15, pp. 189–206, 1955.

12. Kossel, F., and Krebs, K., "Die Pendelvervielfachung von Sekundärelektronen (Multipactoreffekt) als Zündbedingung von Hochfrequenzentladungen," *Zeitschrift für Physik*, Vol. 175, pp. 382–290, 1963.

13. Tamagawa, H., *Electrotechnical Journal of Japan*, Vol. 3, pp. 42–46, June 1957; Tamagawa, H., "A Theory of High-frequency Discharge in High Vacuum," *Electrotechnical Journal of Japan*, Vol. 3, pp. 93–97, September 1957.

14. Fang, F., *Secondary Electron Multiplication in Microwave Field*, Ph.D. dissertation, University of Illinois, Urbana, Illinois, 1959.

15. Abraham, W. G., *Interactions of Electrons and Fields in Cavity Resonators*, Ph.D. dissertation, Stanford University, Palo Alto, California, 1950.

16. Francis, G., *Ionization Phenomena in Gases*, Butterworths Scientific Publications, London, 1960.

17. Stratton, J. A., *Electromagnetic Theory*, McGraw-Hill Book Co., Inc., New York, p. 488, 1941.

18. Francis, G., and Von Engel, A., "The Growth of the High Frequency Electrodeless Discharge," *Philosophical Transactions of the Royal Society of London*, Series A, Vol. 246, pp. 143–180, 1953.

XXI. Communications Systems Research: Mathematical Research

A. A Generalization of Two-Level Autocorrelation

R. J. McEliece

1. Introduction

A (v, k, λ) difference set D is a set of k distinct residues $\{a_1, a_2, \cdots, a_k\}$ modulo v such that every residue $b \not\equiv 0 \pmod{v}$ can be expressed in exactly λ ways in the form $b \equiv a_i - a_j \pmod{v}$. With each difference set we may associate a binary periodic sequence $(s_1, s_2, \cdots)$ with $s_i = 1$, if $i \pmod{v}$ is in D, and $s_i = 0$ otherwise. Since this sequence is periodic of period v, only one cycle s from the sequence is considered. Such cycles we agree to call (binary) difference cycles.

Difference cycles (equivalently, difference sets) have been studied intensively (Refs. 1 and 2). They have important applications to digital communications, mainly because they have 2-level autocorrelation. In this discussion we shall point out certain other[1] (equivalent) properties of difference cycles, which seem susceptible to immediate generalization but show that these generalizations are vacuous.

2. Motivation

If s is a difference cycle, then the defining property of difference sets tells us that the number of ordered pairs (s_i, s_{i+b}) from s (subscripts taken mod v) of the form $(1, 1)$ is λ for all values of $b \not\equiv 0 \pmod{v}$. More generally, let the number of ordered pairs (s_i, s_{i+b}) from s of the form (ϵ_1, ϵ_2) be denoted by $p_{\epsilon_1,\epsilon_2}(b)$. Thus, $p_{1,1}(b) = \lambda$ for all $b \not\equiv 0 \pmod{v}$. A simple enumeration now shows that, in addition, $p_{0,1}(b) = p_{1,0}(b) = k - \lambda$; $p_{0,0}(b) = v - 2k + \lambda$, whenever $b \not\equiv 0 \pmod{v}$. Let us represent s and its bth translate s_b schematically as

$$
\begin{array}{cccccc}
 & \overbrace{}^{k} & & \overbrace{}^{v-k} & & \\
s: & 11\cdots1 & 11\cdots1 & 00\cdots0 & 00\cdots0 & \\
s_b: & \underbrace{11\cdots1}_{\lambda} & \underbrace{00\cdots0}_{k-\lambda} & \underbrace{11\cdots1}_{k-\lambda} & \underbrace{00\cdots0}_{v-2k+\lambda} &
\end{array}
$$

Since exactly λ of the ones from s match up with the 1's from s_b, the remaining $k - \lambda$ ones from s must be paired with zeros from s_b. This shows $p_{1,0}(b) = k - \lambda$. The other relations may be similarly verified.

Now let s be any binary cycle of length v (not necessarily associated with a difference set). The (unnormalized) autocorrelation of s, $R_s(b)$, has been defined as follows (Ref. 1): $R_s(b) = A_b - D_b$, where A_b is the number of agreements between s and s_b (i.e., the number of components in which s and s_b have the same entry), and D_b is the number of disagreements. With the notation introduced above, $R_s(b) = p_{1,1}(b) + p_{0,0}(b) - p_{1,0}(b) - p_{0,1}(b)$. Our remarks then show in particular that for a difference cycle s, $R_s(b)$ is independent of b, if $b \not\equiv 0 \pmod{v}$, and so s has two-level autocorrelation. (Of course $R_s(0) = v$). Conversely, it is easy to show that, except for

[1]The author wishes to thank Professor E. S. Selmer, University of Bergen, Norway for suggesting this problem.

pulse cycles (i.e., $s = (1, 0, 0, \cdots, 0)$ or $(0, 1, 1, \cdots, 1)$ or a translate of one of these), any cycle with two-level autocorrelation is associated with a difference set (Ref. 1).

3. Generalization

The preceding discussion motivates the following formal generalization of a difference cycle, which does not depend on the notion of autocorrelation function. If $s = (s_1, s_2, \cdots, s_v)$ is any binary cycle, and if b is an integer satisfying $0 < b < v$, we define a bigram $M(s,b)$ as follows: $M(s,b) = \{(s_i, s_{i+b}): i = 1, 2, \cdots, v\}$, multiplicity included. We have seen that if s is a binary difference cycle, then $M(s,b) = M(s,b')$, whenever $0 < b < v, 0 < b' < v$. (Equality means that the two collections contain the same pairs with the same multiplicity. Generally, we define an m-gram for an n-ary cycle $s = (s_1, s_2, \cdots, s_v)$ as follows:

Let $b = (b_1, b_2, \cdots, b_{m-1})$ be an ordered $(m-1)$-tuple of integers with $0 < b_1 < b_2 < \cdots < b_{m-1} < v$. We define the m-gram $M(s,b)$ to be the collection $\{(s_i, s_{i+b_1}, s_{i+b_2}, \cdots, s_{i+b_{m-1}}), i = 1,2, \cdots, v\}$, multiplicity included. If $M(s,b) = M(s,b')$ for all such b and b', we say that s is *m-tuply regular*.

Thus, in the new terminology an ordinary difference cycle becomes a doubly-regular binary cycle. We have formally generalized in two directions; we allow both the degree of regularity of s and the number of symbols in s to increase. m-tuply regular unary cycles and singly regular n-ary cycles are trivial, and a pulse binary cycle is m-tuply regular in a trivial way. It is surprising that binary difference cycles and the above trivial examples are the only examples of m-tuply regular n-ary cycles possible. We prove this result now.

4. Nonexistence

Theorem 1: If s is an m-tuply regular n-ary cycle, then one of the following alternatives holds:

(1) $m = 2, n = 2$ and s is a difference cycle

(2) $n = 2$ and s is a pulse

(3) $m = 1$

(4) $n = 1$.

The proof is in two parts: (1) We assume $m > 2$ and conclude $n = 1$, unless s is a pulse; (2) We assume $n > 2$ and show $m = 1$.

Suppose, then, that s is an m-tuply regular n-ary cycle with $m > 2$. It is clearly sufficient to prove this is impossible for $n = 2$, since if $n > 2$ we may identify certain of the symbols to obtain an m-tuply regular binary cycle. We now digress in order to place the problem in a wider context.

Definition: (Ref. 3.) Let S be a set of v distinct objects. A *tactical configuration* $C = C\,[k,m,\lambda,v]$ is a collection of b subsets (called blocks) $B_i, i = 1,2, \cdots, b$, of S such that each block contains exactly k objects from S, and each (unordered) m-tuple from S occurs in exactly λ blocks. C is *symmetric*, if $b = v$ and if each object in S occurs in exactly k blocks.

Our plan is to show that the existence of an m-tuply regular binary cycle implies the existence of a symmetric $C\,[k,m,\lambda,v]$. Theorem 2 shows that such configurations are trivial; we prove Theorem 2 first.

Theorem 2. There are no nontrivial symmetric $C\,[k,m,\lambda,v]$ configurations, if $m \geq 3$. (*Trivial* means that k does not satisfy $m \leq k \leq v - m$.)

Proof. It is clear that a $C\,[k,m,\lambda,v]$ configuration is also a $C\,[k,m',\lambda,v]$ configuration for $m' \leq m$, since each unordered m'-tuple from S is a subset of exactly $\binom{v-m'}{m-m'}$ unordered m-tuples from S, and so each m'-tuple occurs $\lambda\binom{v-m'}{m-m'} \Big/ \binom{k-m'}{m-m'}$ times in the configuration. Consequently, it will be sufficient to prove Theorem 2 for $m = 3$.

For the moment let $\lambda = \lambda_3$, and let λ_2 represent the number of times each unordered pair from S occurs in C. Then counting the number of times a triple involving a given pair occurs in the configuration, we see that

$$\lambda_2(k-2) = \lambda_3(v-2)\,. \tag{1}$$

Note that Eq. (1) holds for any $C\,[k,3,\lambda,v]$ configuration, symmetric or not.

If C is symmetric, let us count the number of times a pair involving a given element occurs

$$k(k-1) = \lambda_2(v-1)\,. \tag{2}$$

We now perform the standard trick (Ref. 4) of deleting from C an arbitrary block, and all objects occurring in that block. Since C is in particular a symmetric block design, the derived design C' will also be a $C\,[k',3,\lambda',v']$ configuration (but no longer symmetric) with $k' = k - \lambda_2, \lambda'$

$= \lambda_3' = \lambda_3$, $\lambda_2' = \lambda_2$, $v' = v - k$. Eq. (1) will now apply to the derived parameters; i.e.,

$$\lambda_2(k - \lambda_2 - 2) = \lambda_3(v - k - 2). \qquad (3)$$

Combining Eq. (3) with Eq. (1) $\lambda_2^2/k = \lambda_3$. Thus, $\lambda_2/k = (k-2)(v-2)$; and so, from Eq. (2), $(k-1)/(v-1) = (k-2)/(v-2)$, which implies $k = v$. But $k = v$ is a trivial design, and this completes the proof of Theorem 2.

To complete the first part of the proof of Theorem 1, it remains to show that the existence of an m-tuply regular binary cycle ($m \geq 3$), which is not a pulse, implies the existence of a nontrivial symmetric $C\,[k,m,\lambda,v]$ configuration. First of all, it is clear that if $1 < k < m$ then no such cycle exists; since for certain b's, $M(s,b)$ will contain $(\overbrace{11\cdots1}^{k}\overbrace{00\cdots0}^{m-k})$, while others will not. Similarly, $v - m > k > v - 1$ is impossible. Thus, except for pulses, all m-tuply regular binary sequences with k 1's satisfy $m \leq k \leq v - m$.

If now s is an m-tuply regular binary cycle of length v (assuming the two symbols are 0 and 1), let $S = \{a_1, a_2, \cdots a_v\}$ be any set containing v distinct objects. Blocks B_i, $i = 0,1,2, \cdots, v-1$ are defined as follows: $a_j \epsilon B_i$ if and only if $s_{i+j} = 1$. To show that these blocks form a symmetric $C\,[k,m,\lambda,v]$ configuration, we need only verify the m-tuple condition.

Thus, let $(a_{i_1}, a_{i_2}, \cdots, a_{i_m})$ be an m-tuple from S, and assume $i_1 < i_2 < \cdots < i_m$. Let $b_j = i_{j+1} - i_j$, $j = 1,2, \cdots, m-1$, and set $b = (b_1, b_2, \cdots, b_{m-1})$. Since s is m-tuply regular, the m-tuple $(1,1, \cdots, 1)$ will occur in $M(s,b)$ a certain number of times, say λ_m, and λ_m is independent of b. It is clear that if $(s_i, s_{i+b_1}, \cdots, s_{i+b_{m-1}})$ is such an m-tuple from $M(s,b)$ then B_{i-i_1} will contain the m-tuple $(a_{i_1}, a_{i_2}, \cdots, a_{i_m})$ and conversely. Hence, every m-tuple from S occurs in exactly λ_m blocks, and so the blocks B_i do form a (nontrivial) symmetric $C\,[k,m,\lambda,v]$ configuration. But this is impossible by Theorem 2, and so every m-tuply regular binary cycle is a pulse. This completes the first part of the proof of Theorem 1.

Our attention has recently been drawn to the fact that the second half of Theorem 1 was proved independently by R. Titsworth (Ref. 5) several years ago. (Titsworth's "perfect" sequences are precisely the doubly regular sequences discussed here.) We present here a new proof which makes use of the stronger theory of difference sets.

In order to prove the second half of Theorem 1, we will assume that s is a doubly regular n-ary cycle and show that $n > 2$ is impossible. It will be sufficient to prove that there are no doubly regular ternary cycles, since a doubly regular n-ary cycle ($n > 3$) can be transformed into a doubly regular ternary cycle by a simple identification of certain symbols.

If s is a doubly regular ternary cycle in the symbols 0,1,2, let s contain k_0 zeros, k_1 ones and k_2 twos. Observe that each k_i must be ≥ 2 since if, say $k_0 = 1$, then some bigrams would contain (0,1) but not (0,2), while others would contain (0,2) but not (0,1). Let us now identify the symbols 0 and 1; s then becomes a doubly regular binary cycle (which is not a pulse by the observation above), and so the set $D_2 = \{i : s_i = 2\}$ is a difference set. Similarly, D_0 and D_1 are also difference sets. But also $D_{0,1} = \{i : s_i = 0$ or $s_i = 1\}$, being the complement of D_2, is a difference set. $D_{0,2}$ and $D_{1,2}$, defined similarly, are also difference sets.

We remark at this stage $D_0 \cup D_1 = D_{0,1}$, $D_0 \cap D_1 = \phi$. This means that if we could prove that the union of two disjoint difference sets is never a nontrivial difference set, the second part of Theorem 1 would follow as an immediate corollary. Although there is strong evidence that this is so (discussed at the end of this article), we are as yet unable to prove it. So we must go another route.

To continue with the proof, write $D_0 = \{a_1, a_2, \cdots, a_{k_0}\}$ and $D_1 = \{b_1, b_2, \cdots, b_{k_1}\}$. We define polynomials θ_0 and θ_1 (Ref. 2) as follows:

$$\theta_0(x) \equiv x^{a_1} + x^{a_2} + \cdots + x^{a_{k_0}} \pmod{x^v - 1}$$

$$\theta_1(x) \equiv x^{b_1} + x^{b_2} + \cdots + x^{b_{k_1}} \pmod{x^v - 1}.$$

Since D_0 and D_1 are difference sets, we have, as in Ref. 2,

$$\theta_0(x)\,\theta_0(x^{-1}) \equiv n_0 + \lambda_0\,T(x) \pmod{x^v - 1}$$

$$\theta_0(x)\,\theta_1(x^{-1}) \equiv n_1 + \lambda_1\,T(x) \pmod{x^v - 1}, \qquad (4)$$

where

$$n_0 = k_0 - \lambda_0$$

$$\lambda_0 = k_0(k_0 - 1)/(v - 1)$$

$$n_1 = k_1 - \lambda_1, \lambda_1 = k_1(k_1 - 1)/(v - 1)$$

$$T(x) \equiv 1 + x + x^2 + \cdots + x^{v-1} \pmod{x^v - 1}.$$

Now since s is doubly regular, the pair (1,0) occurs in each bigram $M(s,b)$ equally often, say μ times. This says that every residue $b \not\equiv 0 \pmod{v}$ can be written in exactly μ ways in the form $b \equiv a_i - b_j \pmod{v}$. In terms of the θ's, this condition becomes

$$\theta_0(x)\,\theta_1(x^{-1}) \equiv -\mu + \mu T(x) \pmod{x^v - 1}. \qquad (5)$$

Multiplying both sides of Eq. (5) by $\theta_1(x)$, and using Eq. (4), we obtain

$$n_1\theta_0(x) + \mu\theta_1(x) \equiv (\mu k_1 - \lambda_1 k_0)\, T(x) \pmod{x^v - 1}. \tag{6}$$

(Observe that $R(x)T(x) \equiv R(1)T(x) \pmod{x^v - 1}$ (Ref. 3).) But the left-hand side of Eq. (6) cannot contain powers of n higher than $v - 1$, and so

$$n_1\theta_0(x) + \mu\theta_1(x) = (\mu k_1 - \lambda_1 k_0)\, T(x). \tag{7}$$

But this is impossible: the left-hand side of Eq. (7) cannot contain all powers of x less than v, since some s_i are equal to 2. This contradiction completes the proof of Theorem 1.

5. Conclusion

We remark finally that there is a certain amount of arbitrariness in our definition of multiple regularity. With hindsight, at least, we might regard the fact that the pairs (0,1) and (1,0) occur evenly distributed among the bigrams of a difference cycle as a fluke, peculiar to the case of binary cycles. We could then define multiple regularity by requiring that only m-tuples of the form $(aa \cdots a)$ be evenly distributed. If this had been done, the first part of the proof of Theorem 1 would still have worked, since we only needed the even distribution of $(11 \cdots 1)$ anyway. But a new proof of the second part of the theorem would be needed; in fact, our conjecture that the union of two disjoint difference sets can never be a difference set is exactly what is required, at least for ternary cycles. Some numerical evidence that this conjecture is true is available. For example, we require $(v - 1) \mid k_0(k_0 - 1)$, $(v - 1) \mid k_1(k_1 - 1)$, $(v - 1) \mid k_0 + k_1)(k_0 + k_1 - 1)$, and Miller (Ref. 6) has found all triples (v, k_0, k_1) which satisfy these conditions and with $k_0 \neq k_1$. It is easy to show that $k_0 = k_1$ is impossible. Using Miller's list, the author has been able to verify the conjecture for $v \leq 300$; there are 32 triples (v, k_0, k_1) in this range. There appears to be no particular difficulty in pushing these numerical results even farther.

B. The Notion of a (p,q,r) Automaton

R. P. Stanley

1. Summary

This work continues the discussion on the algorithmic complexity of finite-state machines, or automata as they are called here. A (p,q,r) automaton is essentially an automaton on a p-symbol alphabet, each stage having at most q "fan-ins" and at most r "fan outs;" that is, a (p,q,r) automaton is defined as one being built with a fixed line of hardware and no amplifiers or expanders. Results are given connecting the algorithmic complexity of (p,q) automata (no fan out restriction) with the complexities of (p,q,r) automata that compute the same function. Use is made of a normal-form theorem for automata.

2. Introduction

In this article the definition of a (p,q) automaton is extended by limiting the number of fan-outs from the stages of a (p,q) automaton to be no more than r. The fan-ins are limited by q, and the alphabet has p symbols.

a. Definition. A (p,q,r) automaton is a (p,q) automaton having at most r stages depend on any given stage.

Of course, any given (p,q) automaton is a (p,q,r) automaton for some r, since the number of stages is necessarily finite. But the point of the definition is that r is bounded in advance by the line of hardware.

It is clear that a (p,q) automaton cannot always be replaced by a $(p,q,1)$ automaton which computes the same function, for one of the input stages may be related to two other input stages in two different ways, thus requiring at least two fan-outs from the given input. For instance, the function $ac(a + b)$ could never be computed on a $(p,2,1)$ automaton, since the computations $a + b$ and ac require two fan-outs from the input a. If, however, we were allowed to input the variable a into two different stages, then the function $ac(a + b)$ could be computed on a $(p,2,1)$ automaton, as illustrated in Fig. 1. These considerations motivate the following definition.

b. Definition. A (p,q,r) automaton with repeated inputs [abbreviated (p,q,r) r.i.-automaton] is a (p,q,r)

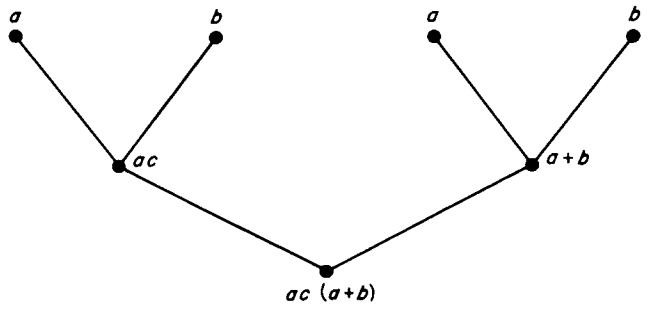


Fig. 1. An automaton with repeated inputs

automaton in which the same input bit is allowed to be put into more than one stage initially.

We shall show later that any function f: $V_p^k \rightarrow V_p^n$ realizable on a (p,q) automaton is also realizable on a (2,2,1) r.i. – automaton. If repeated inputs are not to be allowed, then the repetition of the inputs can be incorporated into the automaton by using stages with two fan-outs and one fan-in which double the multiplicity of an input bit at successive times, until the desired number of repetitions is reached. Therefore, as a consequence of the fact that a (p,q) automaton can be replaced by a (2,2,1) r.i. – automaton, we have that any function f: $V_p^k \rightarrow V_p^n$ realizable on a (p,q) automaton is realizable on a (2,2,2) automaton (without repeated inputs) such that at most three "wires" are connected to any one stage. In other words, each stage has at most three leads coming out of it, input leads plus output leads. Such a (2,2,2) automaton will be denoted a (2,2,2)* automaton. In Fig. 2 the automaton of Fig. 1 is illustrated without repeated inputs.

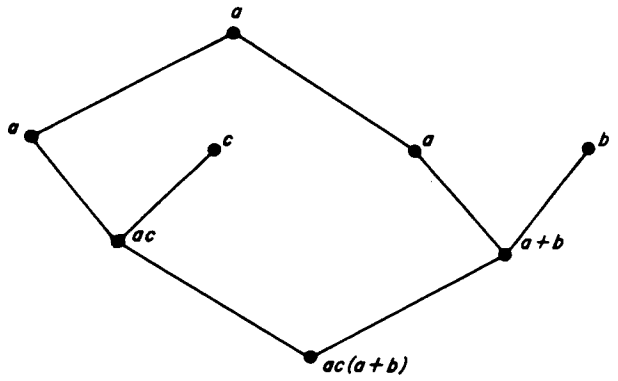


Fig. 2. An equivalent automaton with repeated inputs

3. Converting to (p,q,1) r.i.-Automata

There is a simple procedure for converting a (p,q,r) automaton into a $(p,q,1)$ r.i. – automaton. We need consider only the case of one output bit ($n = 1$), since the procedure can be repeated separately for each output bit. We simply label each stage, copy down the one output stage, connect it to the, at most, q stages on which it depends, label these, at most, q stages to correspond to the original automaton, and repeat this process until only stages with input labels remain unconnected to stages "above" them. Stages with the same label compute the same function g: $V_p^q \rightarrow V_p^1$. The entire procedure will be clear after examining the illustration in Fig. 3. We call the final $(p,q,1)$ r.i. – automaton obtained from this procedure the normal form of the original automaton.

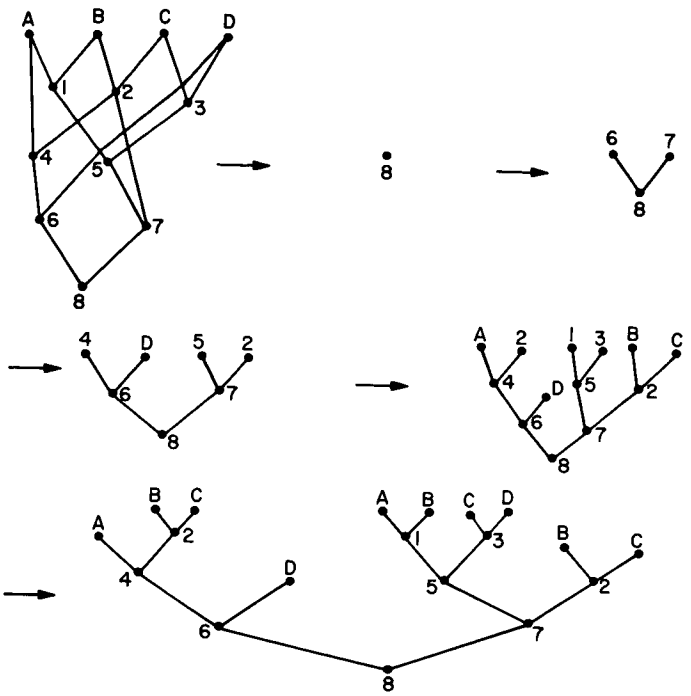


Fig. 3. Conversion of a (2,2) automaton into a (2,2,1) r.i.-automaton

4. Redundant Automata

Although it is very difficult, in general, to decide whether a given (p,q) automaton is the simplest possible for computing a given function, i.e., whether the automaton corresponds to a point on the algorithmic complexity, there, however, is a special case in which this is easily determined. This occurs whenever there is a stage A which depends on not more than q stages where at least one stage intervenes between the q stages and the stage A itself which suffices to determine A. This intervening stage is clearly extraneous, since A can have q fan-ins and thus depend directly on the not more than q stages.

We shall call a (p,q) automaton redundant if this phenomenon occurs, otherwise irredundant. An irredundant automaton does not necessarily correspond to a point on the algorithmic complexity. If, however, it does not, then the degree and method of simplification depend explicitly on the function being computed.

We call a set $\mathcal{S}$ of stages a cover of the stage S if specifying the states of the members of $\mathcal{S}$ at some particular time (not necessarily the same time for all members of $\mathcal{S}$) suffices to determine the state of S at some later time. Equivalently, $\mathcal{S}$ covers S if every "path" from

S to any input stage on which S depends passes through some member of $\mathcal{S}$.

We require one further definition. The index of a stage S is the number of times the stage S is repeated in the normal form of the automaton. Thus, in Fig. 3 some covers of stage 6 are {6}, {4,D}, {4,D,1}, {A,2,D}, while, e.g., {A,B,D}, {1,2,3,4,A,B,C} are not covers. The index of stage 6 is one, of 2 is two, of B is three, etc. The automaton in Fig. 3 is irredundant.

5. Normal Form Theorems

An interesting topological connection between an automaton and its normal form follows immediately from the construction of the normal form.

a. Theorem 1. The index of a stage S is equal to the number of distinct paths from the output stage to S, such that two distinct fan-ins or two distinct fan-outs of any given stage are not both part of the path, and such that the path does not trace back over any line segment (wire); in other words, the index equals the number of "upwards directed" paths.

Since we have previously proved (*SPS 37-34*, Vol. IV, p. 298) that any (p,q) automaton can be reduced to a (2,2) automaton which computes the same function, for the remainder of the article we shall be concerned primarily with the cases $p = q = 2$. When $q = 2$, there is a fairly simple procedure for determining whether or not a given automaton is redundant. Note that there is always a finite procedure for answering this question for any q; we simply check all possible covers. This process can, however, be quite time-consuming, and it may be difficult to go about it systematically for sufficiently complicated automata. In such cases the following theorem is useful. We always assume that every nonoutput stage has at least one fan-out.

b. Theorem 2. A $(2,p)$ automaton is irredundant if and only if in its normal form the following simultaneously hold (Fig. 4).

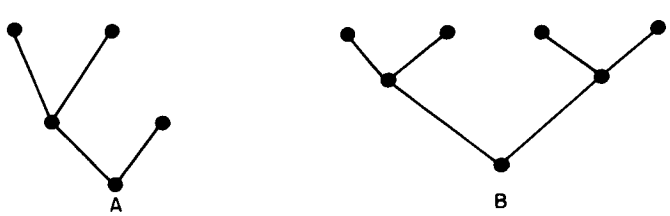


Fig. 4. Subautomata

(1) Every noninput stage has precisely two fan-ins

(2) The five stages of every subautomaton of the form of Fig. 4A (or its mirror image) have distinct labels

(3) The four upper stages of every subautomaton of the form of Fig. 4B have at least three distinct labels

(See Fig. 5 for examples.)

Proof. ($\Leftarrow$) In Fig. 5Bα, stage 3 is clearly extraneous, for in the original automaton stage 4 can depend on 1 and 2 directly. Moreover, stage 3 of Fig. 5Bβ and stage 3 and 4 of Fig. 5Bγ are clearly extraneous, for in the original automata stage 4 of β can depend directly on 1 and 2. It is easily seen that any other labeling violating the conditions of the theorem similarly leads to extraneous stages, whereby the implication in one direction is established.

($\Rightarrow$) Assume the original automaton is redundant and that stage 1 is determined by the states of stages 3 and 4, although 2 intervenes. Then at some place in the normal form stage 1 is covered by stages 3 and 4 with stage 2 in between. Since each stage has only one fan-out, this implies that either

(1) The stages 3 and 4 cover of 1 consists of just two stages in the normal form, in which case a subautomaton of the form of Fig. 5Bα must occur; or (2) the stages 3 and 4 cover of 1 consists of more than two stages.

Assume no subautomata of the form 5Bα occur; we want to prove that there is then a subautomaton of the form 5Bβ or 5Bγ. Let T be the greatest time at which any stage 3 or 4 is removed from the output in the normal form. Say that this time is achieved by stage labeled 3. Call this stage S_1. Then S_1 has one fan-out leading to a stage S_3. The other fan-in to stage S_3 must be connected to a stage S_2 labeled 4, since T is the greatest time of removal, by definition. Stage S_3 has one fan-out leading to a stage S_4. S_4 has one other fan-in. If this comes from a stage labeled 3 or 4, we have a subautomaton of the form 5Bβ. Otherwise it comes from a stage S_5. Since 3,4 cover stage 1 and since T is maximal, S_5 must depend on stages labeled 3,4. This introduces a subautomaton of the form 5Bγ; the proof of Theorem 2 is completed.

The statement of Theorem 2 gives an algorithm for removing irredundancies from a $(p,2)$ automaton. We convert the automaton to normal form, locate the subautomata of the form of Fig. 5B, remove the irredundancies from them (relabeling the stages to show that the

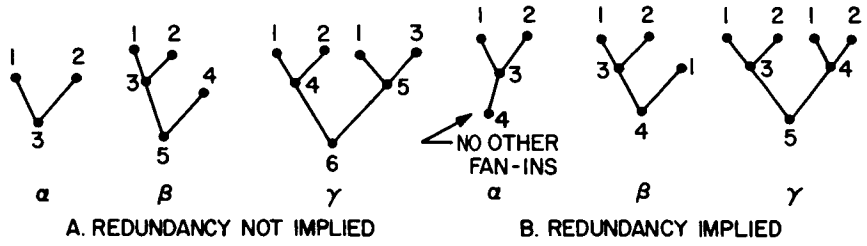


Fig. 5. Determining redundancy from the normal form

function f: $V_2^2 \rightarrow V_2^1$, which a particular stage computes, may have changed), and reconvert the normal form. Fig. 6 illustrates this procedure. The crucial subautomata of the normal form are indicated by broken circles. The primes (′) indicate a relabeling of the stages, as mentioned above.

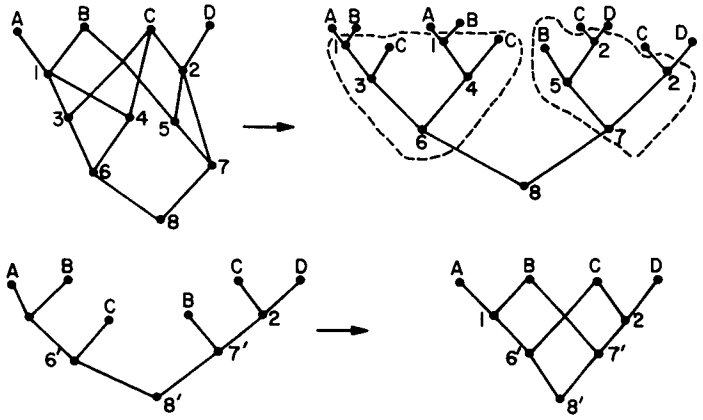


Fig. 6. Removing irredundancies

There are, of course, alternative methods for determining redundancy. We leave it to the reader to verify that a $(p,2)$ automaton is redundant if and only if in its diagram (not its normal form) the following do not occur (still assuming that each stage except the output has at least one fan-out. (1) A closed triangle connecting three stages (2) Two closed quadrilaterals, each connecting four stages, with three stages in common, (3) A stage with only one fan-in.

The criteria of Theorem 2 are, however, simpler to apply.

6. (2,2,1) Automata

There are two quantitative questions concerning (2,2,1) automata which we shall discuss further.

(1) Given a (2,2) automaton with N stages and computing time T, what are the characteristics N_0, T_0 of the (2,2,1) r.i. – automaton which computes the same function?

(2) Given a function f: $V^k \rightarrow V^n$, what are the characteristics N and T of a (2,2,1) r.i. – automaton which computes f?

Theorem 3 is concerned with the first of these questions.

a. Theorem 3. Let N and T be the characteristics of an irredundant (2,2) automaton which realizes the function f: $V^k \rightarrow V^1$. Then there exists a (2,2,1) r.i. – automaton (namely, the normal form) which realizes f with characteristics N_0 and T_0 such that

$$\begin{cases} N_0 \leq 2^A - 1, \text{ where } A = \min\left[(N/3) + 1, T\right] \\ T_0 \leq T. \end{cases}$$

Conversely, there exists an irredundant (2,2) automaton with characteristics N,T such that its normal form satisfies

$$\begin{cases} N_0 = 2^T - 1; \\ T_0 = T, \end{cases}$$

where N can be any multiple of 3 and $T = N/3 + 1$.

Proof. The equality $T_0 = T$ always holds because of the method of construction of normal form. Since any (2,2) automaton of time T_0 with one output never has more than $2^{T_0} - 1$ stages (as each stage is connected with at most two stages "above" it), we have

$$N_0 \leq 2^{T_0} - 1 = 2^T - 1.$$

There remains only the inequality $N_0 \leq 2^{N/3+1} - 1$. The proof is by induction on T, the computing time of the (2,2) automaton. We say that a stage is at time $T-t$ if the maximum number of time units which separate it from the output is t. Thus, output is always at time T, while at least one input stage is at time 1. Since there is only one output stage, there is one stage at time T and two stages at $T-1$. If the automaton then has at least

three stages at time $T-t$ for all $2 \leq t \leq T-1$, then clearly $T \leq N/3+1$, so that $N_0 \leq 2^{N/3+1}-1$. The only case remaining is that the (2,2) automaton has only two stages A,B at some time $T-t_1, t_1 > 1$. The induction base for this case is $T = 1,2$; for then $N = 1,3$ and the theorem is satisfied.

For the case that we are considering, it follows that there is only one noninput stage C at $T-t_1+1$, namely the stage depending on A and B. We can then regard C as the output of the subautomaton M_2M_2, consisting of all stages at time $\leq T-t_1$ and stage C itself. Assume that this subautomaton has N_2 stages, and let $N_1 = N - N_2 + 1$ be the number of stages of the remaining subautomaton M_1. (The stage C is part of both automata.) It is easy to see that M_1 and M_2 are irredundant when the entire automaton M is. Let N_1' be the number of stages in the normal form M_i' of M_i. By the induction hypothesis,

$$N_1' \leq 2^{\frac{N_1}{3}+1}$$

$$N_1' \leq 2^{\frac{N_2}{3}+1}$$

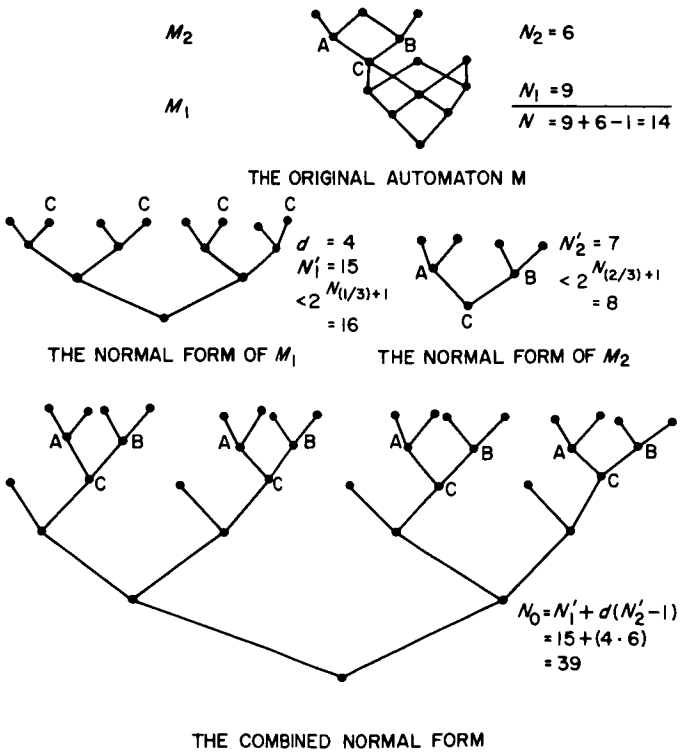


Fig. 7. The proof of Theorem 3

It is not hard to see that a normal (2,2) automaton with N' stages contains $\frac{1}{2}(N'+1)$ input stages (including repetitions). Moreover, it is not hard to see that the index d of a given input stage of the original automaton cannot be greater than $\frac{1}{4}(N'+1)$ without introducing irredundancies. Hence the stage C can have an index d of at most $\frac{1}{4}(2^{\frac{N_1}{3}+1}-1+1) = 2^{\frac{N_1}{3}-1}$. Therefore

$$N_0 = N_1' + dN_2' - d, \tag{1}$$

since the d repetitions of stage C have been counted twice. From Eq. (1) there follows

$$\begin{aligned}
N_0 &= N_1' + d(N_2' - 1) \\
&\leq 2^{\frac{N_1}{3}+1} - 1 + 2^{\frac{N_1}{3}-1}\left(2^{\frac{N_2}{3}+1} - 2\right) \\
&= 2^{(N_1+N_2)/3} + 2^{\frac{N_1}{3}} - 1 \\
&= 2^{\frac{N_1+N_2-1}{3}+1} - \left(2^{\frac{N_1+N_2-1}{3}+1} - 2^{\frac{N_1+N_2}{3}} - 2^{\frac{N_1}{3}}\right) - 1 \\
&= 2^{\frac{N}{3}+1} - \left[2^{\frac{N_1+N_2}{3}}\left(2^{\frac{2}{3}}-1\right) - 2^{\frac{N_1}{3}}\right] - 1 \\
&\leq 2^{\frac{N}{3}+1} - \left[2^{\frac{N_1+3}{3}}\left(2^{\frac{2}{3}}-1\right) - 2^{\frac{N_1}{3}}\right] - 1 \text{ (since } N_2 \geq 3) \\
&= 2^{\frac{N}{3}+1} - 2^{\frac{N_1}{3}}\left(2^{\frac{5}{3}} - 2 - 1\right) - 1 \\
&= 2^{\frac{N}{3}+1} - 2^{\frac{N_1}{3}}(0.1748\cdots) - 1 \\
&< 2^{\frac{N}{3}+1} - 1.
\end{aligned}$$

With this last inequality the proof follows by induction. Some of the concepts of the proof are illustrated in Fig. 7.

To prove the converse we need only exhibit a (2,2) automaton whose normal form has the desired characteristics. From the above proof we expect such an automaton to have three input stages and to have each time unit $\geq T-2$ occupied by exactly three stages. Fig. 8 illustrates this automaton for $N = 15$. The general case is immediate, and the reader can verify that the normal form does indeed have the desired characteristics.

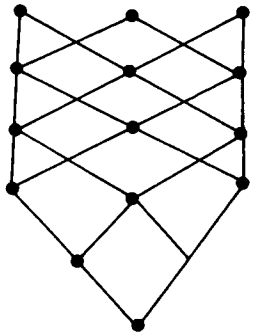


Fig. 8. An automaton whose normal form is maximal

As a consequence of the equality $T_0 = T$ we can improve a result obtained in *SPS 37-34*, Vol. IV, p. 298 concerning functions which take a long time to be computed:

b. Theorem 4. For any $n, k \geq 1$ there exists a function f: $V^k \to V^n$ such that every (2,2) automaton realizing f requires a computing time of

$$T > k - \log_2 \log_2 k.$$

Proof. It suffices to consider the case $n = 1$. Because of the inequality $T_0 \leq T$ in Theorem 3, it also suffices to assume that every automaton is in normal form. We proceed to get an upper bound on the number of automata in normal form realizing some function $V^k \to V^1$ in a time t. Clearly nothing is lost by extending the normal form so that there are exactly $2^{t-t'}$ stages at time t', for we can add extraneous stages which merely double the multiplicity of an input stage. This is illustrated in Fig. 9.

Each of the $2^{t-1} - 1$ stages at times $2 \leq t' \leq t$ is determined by one of the 16 functions g: $V^2 \to V$. (The stages which repeat input correspond to one of these functions g, e.g., $g(x_1, x_2) = x_1$.) The 2^{t-1} stages at time 1 are determined by which of the k input bits they admit. Thus, there are at most

$$16^{(2^{t-1}-1)} k^{2^{t-1}}$$

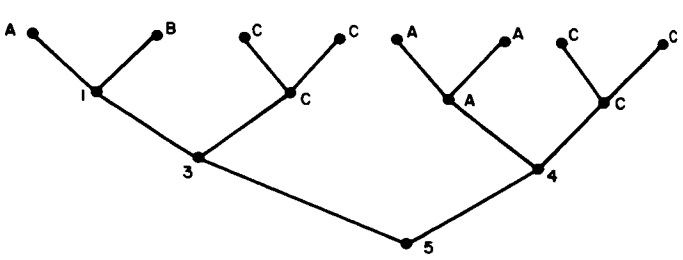


Fig. 9. Adding extraneous input stages

different functions f: $V^k \to V$ which can be realized in a time t on a (2, 2) automaton.

Now if

$$\sum_{t=1}^{T} 16^{(2^{t-1}-1)} k^{2^{t-1}} < 2^{2^k},$$

then there is a function f: $V^k \to V$ requiring a time $> T$. Now

$$\begin{aligned} S(T) &= \sum_{t=1}^{T} 16^{(2^{t-1}-1)}\, k^{2^{t-1}} \\ &< 2 \cdot 16^{2^{T-1}-1}\, k^{2^{T-1}}. \text{ Thus} \\ \log S(T) &< 1 + 4(2^{T-1} - 1) + 2^{T-1} \log_2 k \\ &< 2^{T+1} (1 + \tfrac{1}{4} \log_2 k), \text{ so that} \\ \log_2 \log_2 S(t) &< T + 1 + \log (1 + \tfrac{1}{4} \log_2 k) \\ &\leq T + \log_2 \log_2 k. \end{aligned}$$

We require

$$\log_2 \log_2 S(t) < k,$$

which thus holds for all

$$T \leq k - \log_2 \log_2 k.$$

The largest such T satisfies

$$1 + T > k - \log_2 \log_2 k,$$

and, hence, there is indeed a function requiring time at least

$$T > k - \log_2 \log_2 k.$$

This completes the proof of Theorem 4.

The next theorem deals with the second question about (2, 2, 1) automata.

c. Theorem 5. (A) any function f: $V^k \to V^n$ can be realized on a (2, 2, 1) r.i. − automaton with characteristics

$$\begin{cases} N \leq n\,(k2^{k+1} - 1), \\ T \leq k + \log_2 k + 1. \end{cases}$$

(B) Any function f: $V^k \to V^n$ can be realized on a (2, 2, 2)* automaton with characteristics

$$\begin{cases} N^* \leq 3n\,(\mathrm{k}2^k - 1) \\ T^* \leq 2\,(k + \log_2 k). \end{cases}$$

Proof. (A) It was proved in *SPS 37-34*, Vol. IV, p. 298 that any function f: $V^k \to V^n$ can be realized on a (2, 2) automaton satisfying

$$T' \leq k + \log_2 k + 1 .$$

We can regard an automaton mapping $V^k \to V^n$ as a set of n automata mapping $V^k \to V^1$. Thus, by Theorem 3, we have for a (2, 2, 2) r.i. – automaton that N and T satisfy

$$\begin{aligned} N &\leq n(2^{T'} - 1) \\ &= n(2^{k+1\log_2 k+1} - 1) \\ &= n(k2^{k+1} - 1); \\ T &\leq T' = k + \log_2 k + 1. \end{aligned}$$

(B) We noted previously that if the normal form of a (2,2) automaton has N_0 stages, then $\frac{1}{4}(N_0 + 1)$ of these stages are used for (repeated) input. An input stage cannot have index $> \frac{1}{4}(N_0 + 1)$ without introducing irredundancies. If an input stage A has index d, then it requires $\leq d - 1$ additional stages and an additional time $\leq \log_2 d$ to repeat stage A $d - 1$ additional stages and an additional time $\leq \log_2 d$ to repeat stage A d times, starting with one input and using stages with just two fan-outs. Therefore, the maximum possible additional time is $\leq \log_2[\frac{1}{4}(k2^{k+1})] = k + \log_2 k - 1$. An upper bound on the number of additional stages per output stage is

$$\begin{aligned} MAX \sum (d_i - 1) &= \tfrac{1}{4}(N_0 + 1) - 1 + \tfrac{1}{4}(N_0 + 1) - 1 \\ d_1 + \cdots + d_r &= \tfrac{1}{2}(N_0 + 1) \\ d_i &\leq \tfrac{1}{4}(N_0 + 1) \\ &\leq \tfrac{1}{2}\, 2^T - 2 \text{ (since } N_0 \leq 2^T - 1) \\ &= k2^k - 2, \text{ by part A.} \end{aligned}$$

Thus, for n output stages the maximum number of additional stages is $n(k2^k - 2)$. Combining with A we get

$$\begin{aligned} N^* &\leq n(k2^{k+1} - 1) + n(k2^k - 2) \\ &= 3n(k2^k - 1), \\ T^* &\leq (k + \log_2 k + 1) + (k + \log_2 k - 1) \\ &= 2(k + \log_2 k). \end{aligned}$$

Theorem 5 is proved.

C. Further Results on the Algorithmic Complexity of (p, q) Automata

R. P. Stanley

1. Introduction

In *SPS 37-34*, Vol. IV, the concept of the algorithmic complexity of a discrete function (a function with a finite domain and range), as defined by Ofman (Ref. 7) was investigated. (For the basic definitions and properties concerning algorithmic complexity, the reader is referred to *SPS 37-34*, Vol. IV, p. 298.) In this article we shall deal only with (2,2) automata. Since a 2-symbol alphabet corresponds to a Boolean function, we can apply the known theory of Boolean functions to give results on algorithmic complexity. The theory of Boolean functions is first used to improve bounds obtained in *SPS 37-34*, Vol. IV, p. 298, and in the previous article on the complexity of any function. Next, the concept of equivalence of functions under permutation groups on the domain and range is introduced, and it is shown that certain well-known groups preserve algorithmic complexity, while other groups do not have too large an effect. Finally, we discuss criteria which guarantee that a function has low complexity.

2. Bounds on the Complexity of Any Function

We shall denote the set {0, 1} by V; V^k denotes the set of all 2^k of the k-tuples of 0's and 1's. In *SPS 37-34*, Vol. IV, p. 298 we showed that every function $f: V^k \to V^n$ can be computed on a {2, 2} automation with simultaneous characteristics as follows (N is the number of stages; T is the computation time):

$$\left.\begin{aligned} N &\leq n(k2^{k-1} - 1) + k, \\ T &\leq k + \log_2 k + 1. \end{aligned}\right\} \qquad (1)$$

Henceforth we shall be concerned primarily with the case $n = 1$. For if N_1 and T_1 denote the maximum complexities of all functions $f_1: V^k \to V$, then it is easy to see that every $f: V^k \to V^n$ satisfies

$$\left\{\begin{aligned} N &\leq n N_1 - (n-1)k, \\ T &\leq T_1. \end{aligned}\right.$$

By considering the actual form of the 2^{2^k} distinct functions $f: V^k \to V$, it is possible to improve the results of Eq. (1). To do this, we use the well-known fact that

the 2^{2^k-1} distinct functions $f_1: V^k \to V$ satisfying $f_1(0,0,\cdots,0) = 0$ have the form

$$f_1(x_1, \cdots, x_k) = \sum_S x_{\sigma(1)} x_{\sigma(2)}, \cdots x_{\sigma(r)},$$

where S is some subset of $P\{1, 2, \cdots, k\} - \{\phi\}$, the power set (set of all subsets) of $\{1, 2, \cdots, k\}$, excluding the null set. The sum runs over all elements $\{\sigma(1), \cdots, \sigma(r)\}$ of S. All arithmetic is mod 2. Since $P\{1, 2, \cdots, k\} - \{\phi\}$ contains $2^k - 1$ elements, there are 2^{2^k-1} different subsets S. For example, if $k = 3$ and $S = \{\{1\}, \{12\}, \{23\}\}$, then $f_1(x_1, x_2, x_3) = x_1 + x_1 x_2 + x_2 x_3$ (mod 2). The remaining 2^{2^k-1} functions (those satisfying $f_1(0, \cdots, 0) = 1$) are obtained by adding 1 (mod 2) to the functions obtained above. We leave to the reader the easy verification that these 2^{2^k} functions are all distinct. This then accounts for all possible functions $f_1: V^k \to V$. The reader should be aware that we are using "+" here to denote ordinary mod 2 addition (the "exclusive OR") and not the "inclusive OR" as is often used in digital circuit terminology.

An immediate consequence of the above observations follows:

Theorem 1. If every function $f: V^k \to V$ can be computed in a time $\leq T(k)$ on a (2, 2) automaton, then every function $g: V^{k+1} \to V$ can be calculated in a time

$$T(k + 1) \leq T(k) + 2.$$

Proof. Assume

$$g(x_1, \cdots, x_{k+1}) = \sum_S x_{\sigma(1)} x_{\sigma(2)} \cdots x_{\sigma(r)},$$

where $S \subseteq P\{1, 2, \cdots, k + 1\} - \{\phi\}$. Then we can write

$$g = x_{k+1} \sum_{S_1} x_{\tau(1)} \cdots x_{\tau(2)} + \sum_{S_2} x_{w(1)} \cdots x_{w(t)}$$

$$= x_{k+1} g_1 + g_2,$$

where $S_1, S_2 \subseteq P\{1, 2, \cdots, k\} - \{\phi\}$. (We have simply factored x_{k+1} out of all terms in which it appears.) Consequently g_1 and g_2 are functions mapping V^k into V.

By assumption, we can compute g_1 and g_2 in a time $\leq T(k)$. It requires at most one additional time unit to multiply g_1 by x_{k+1} (mod 2) and one further time unit to add g_2 (mod 2). Therefore, $T(k + 1) \leq T(k) + 2$, and the proof is complete.

More generally, we can factor out k_0 variables x_{k+1}, x_{k+2}, $\cdots$, x_{k+k_0} and write

$$g = g_1 + x_{k+1} g_2 + x_{k+2} g_3 + \cdots + x_{k+k_0} g_{k_0+1} + x_{k+1} x_{k+2} g_{k_0+2} + \cdots + x_{k+1} \cdots x_{k+k_0} g_{2k_0}.$$

Then similarly to the above it can be proved that $T(k + k_0) \leq T(k) + 1 + k_0$, if $T(k) \geq 1 + \log_2 k_0$.

Theorem 2. If every function $f: V^k \to V$ requires $\leq N(k)$ stages to be computed on a (2, 2) automaton, then every function $g: V^{k+1} \to V$ requires

$$N(k + 1) \leq 2N(k) - k + 3$$

stages to be computed on a (2, 2) automaton.

Proof. As in the proof of the previous theorem we write

$$g = x_{k+1} g_1 + g_2.$$

By assumption g_1 and g_2 each require $\leq N(k)$ stages to be computed. Since they have the same k stages of input, both g_1 and g_2 can be computed with $\leq 2N(k) - k$ stages. It then requires at most two additional stages to compute $x_{k+1} g_1$ (one of the stages is used to input x_{k+1}) and one additional stage to compute $g = x_{k+1} g_1 + g_2$, so that

$$N(k + 1) \leq 2N(k) - k + 3,$$

as required.

Corollary 1. Every function $f: V^k \to V$ can be computed on a (2, 2) automaton satisfying

$$N(k) \leq 3 \cdot 2^{k-2} + k - 2.$$

Proof. Clearly $N(2) = 3$. Moreover, the function $2N(k) - k + 3$ is an increasing function of $N(k)$. Therefore by the previous theorem an upper bound on $N(k)$ will be given by the solution of the difference equation

$$N(k + 1) = 2N(k) - k + 3$$

with the initial condition $N(2) = 3$. The unique solution is then

$$N(k) = 3 \cdot 2^{k-2} + k - 2,$$

and the corollary follows.

This improves the result

$$N(k) \leq k2^{k-1} + k - 1$$

obtained in *SPS 37-34*, Vol. IV, p. 298.

3. Permutation Groups on the Domain and Range of Boolean Functions

In what follows we shall employ the concept of permutation groups on the domain V^k and range V^n of a Boolean function $f:V^k \rightarrow V^n$. Specifically, let G be a permutation group on the set of all k-tuples $(x_1, \cdots, x_k)$ and let H be a permutation group on the set of all n-tuples $(y_1, \cdots, y_n)$. Then we say that the functions $f:V^k \rightarrow V^n$ and $g:V^k \rightarrow V^n$ are equivalent with G on the domain and H on the range if for some fixed $\pi \in G$, $\sigma \in H$ we have

$$\sigma f(\pi(\mathbf{x})) = g(\mathbf{x}) \qquad (2)$$

for all $\mathbf{x} \in V^k$. We use the notation (G, H) to denote a group G acting on the domain and a group H acting on the range of a Boolean function. If Eq. (2) holds then we also say that f and g are equivalent under (G, H) [it is easily seen that the relation of equivalence under (G, H) is an equivalence relation]. We let $E(G, H)$ denote the number of equivalence classes induced by the equivalence relation (G, H).

We shall be primarily concerned with the following permutation groups G, H

(1) $G = C_2^k$, the group of all 2^k complementations of variables. This is the so-called elementary abelian group of order 2^k.

(2) $G = S_k$, the symmetric group on the k variables, i.e., the group of all permutations of input bits.

(3) $G = G_k$, the smallest group containing both C_2^k and S_k. G_k is the semidirect product of C_2^k and S_k (Ref. 8, Sect. 6, Theorem 4).

(4) $G = GL_k$, the full linear group on the input variables.

(5) $G = A_k$, the affine group on the input variables. It is the smallest group containing both C_2^k and GL_k.

(6) $H = 1$, the group consisting only of the identity. In other words, two elements $\mathbf{y}_1$, $\mathbf{y}_2$ of the range are considered equivalent if and only if $\mathbf{y}_1 = \mathbf{y}_2$.

(7) For $n = 1$, $H = N$, the group of complementations on the range. N is the cyclic group of order 2.

For the definitions and basic properties of these groups see Ref. 9 Chap. IX. For the general theory of permutation groups on the domain and range of a function see Ref. 10, Chap. 5. For some of their applications to Boolean functions the reader is referred to Refs. 8, 11–13.

4. Application of the Permutation Groups

An immediate application of the above remarks to the concept of algorithmic complexity is given by the following theorem.

Theorem 3. The algorithmic complexity M_f of $f:V^k \rightarrow V$ is invariant under (G_k, N), i.e., if f and g are equivalent under (G_k, N), then f and g have exactly the same algorithmic complexity on a (2, 2) automaton.

Proof. A permutation $\pi \in S_k$ of the variables merely corresponds to a relabeling of the stages of the automaton and consequently does not affect the algorithmic complexity. If a complementation $\sigma \in C_2^k$ is carried out on certain variables, then we need only appropriately alter the functions $f_2:V^2 \rightarrow V$ of pairs of input variables which correspond to the first step of the computation. Again the algorithmic complexity is not affected. Moreover, the complementation on the range can simply be incorporated into the last step of the computation, when necessary, so that the algorithmic complexity is not altered.

To illustrate, Fig. 10 shows an automaton for computing $f(x_1, x_2, x_3) = x_1 x_2 + x_3 + 1$ and the corresponding automaton for computing $f(x_2, x_3 + 1, x_1 + 1) + 1 = x_2 x_3 + x_1 + x_2$. Since every permutation in G_k is the product of permutations in C_2^k and S_k, the proof is complete.

Theorem 3 results in a substantial decrease in the number of functions $f:V^k \rightarrow V$ which have to be considered when dealing with algorithmic complexity. Rather than dealing with 2^{2^k} functions, we instead consider $E(G_k, N)$ classes of functions. For a method of calculating $E(G_k, N)$, based on a theorem of De Bruijn (Ref. 14), the reader is referred to Ref. 12. See also Ref. 10, Chap. 5, for a clear

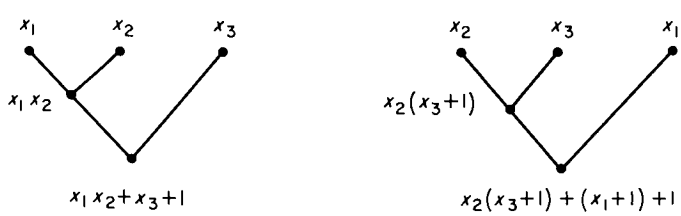


Fig. 10. Equivalent automata

Table 1. $E(G_k, N)$ compared with 2^{2^k}

k	$E(G_k,N)$	2^{2^k}
1	2	4
2	4	16
3	14	256
4	222	65,536
5	616,126(6.2×10^5)	4,294,967,296(4.3×10^9)
6	200,253,952,527,184(2.0×10^{14})	18,446,744,073,709,551,616(1.8×10^{19})

explanation of De Bruijn's theorem. Table 1 reproduces the values of $E(G_k, N)$ obtained in Ref. 12 for $1 \leq k \leq 6$, and compares these with 2^{2^k}.

Since the equivalence (G_k, N) results in such a great reduction in the number of possible cases, it is of interest to pose the following problem.

Problem. What is the largest permutation group acting on V^k such that (G, N) preserves algorithmic complexity? (If $(G, 1)$ preserves algorithmic complexity, so does (G, N).) In other words, for what complexity-preserving group G is $E(G, N)$ a minimum?

We shall not attempt to solve this problem here. It is easy to see that the group GL_k (and hence A_k) does not preserve algorithmic complexity. For instance, when $k > 1$ then the mapping $x_1 \to x_1 + x_2 + \cdots + x_k$ results in an increase of complexity in computing $f(x_1, \cdots, x_k) = x_1$. We do, however, have the following.

Theorem 4. If $f: V^k \to V$ and $g: V^k \to V$ are equivalent under (A_k, N), and if f can be computed with $\leq N_f$ stages in a time $\leq T_f$ on a (2, 2) automaton, then g can be computed with $\leq N_g$ stages in a time $\leq T_g$ on a (2, 2) automaton, where

$$N_g \leq N_f + (k-1)(k-2),$$
$$T_g \leq T_f + \{\log_2 k\}$$

(braces denote the "rounding upwards" function). The same results hold for (GL_k, N), since $GL_k \subset A_k$

Proof. By assumption there is a $\pi \in A_k$ such that $g(\mathbf{x}) = f(\pi(\mathbf{x}))$. Hence, g can be computed by first computing $\pi(\mathbf{x})$ and then $f(\pi(\mathbf{x}))$. Since the output of π corresponds to the input of f, we, therefore, have that

$$\left\{ \begin{array}{l} N_g \leq N_f + N_\pi - k, \\ T_g \leq T_f + T_\pi - 1, \end{array} \right\} \tag{3}$$

and it remains to find an upper bound for N_π, T_π.

Now π is determined by a nonsingular matrix R over $GF(2)$ and by a fixed column vector $\mathbf{v}$ of k components. Here $\pi(\mathbf{x}) = R\mathbf{x} + \mathbf{v}$. The addition of the vector $\mathbf{v}$ to $R\mathbf{x}$ does not affect the complexity, since it can be incorporated into the last step of the computation. If x_i and y_i denote the i-th component of $\mathbf{x}$ and $R\mathbf{x}$, respectively, then each y_j is the sum (mod 2) of certain of the x_i. There are not more than k such x_i, and the mod 2 sum of $\leq k$ variables has a complexity

$$N \leq 2k - 1,$$
$$T \leq 1 + \{\log_2 k\}.$$

Now at most one y_j is the sum of all k of the x_i, since otherwise R would have two rows of all 1's and be singular. Therefore, $k - 1$ of the y_j are the sum of $\leq k - 1$ of the x_i, and for these y_j we have

$$N \leq 2(k-1) - 1 = 2k - 3.$$

Since all the y_j can be computed during the same "time," we have $1 \leq T \leq 1 + \{\log_2 k\}$; since the k input variables need not be separated, we have

$$N_\pi \leq (2k-1) + (2k-3)(k-1) - k(k-1)$$
$$= k^2 - 2k + 2,$$
$$T_\pi \leqq 1 + \{\log_2 k\}.$$

Therefore, from Eq. (3) we get

$$N_g \leq N_f + k^2 - 3k + 2 = N_f + (k-2)(k-1),$$
$$T_g \leq T_f + \{\log_2 k\}.$$

This completes the proof of Theorem 4. (A more careful analysis would probably improve the result on N_π.)

It would in general be useful to know the algorithmic complexity of the operation $R\mathbf{x}$, R fixed; in particular, it would have applications to encoding and decoding error-correcting codes. This question appears to be related to the subject of projective planes, but we shall not go into it any further in this discussion. We mention only the conjecture that the best bound for N_π is of the order $Cx^{3/2}$, where C is a constant.

5. Criteria for Small Complexity

We shall now discuss certain properties of functions $f: V^k \to V$ which tend to make their complexities small. We start with two definitions.

Definition 1. Let $f: V^k \to V$. Then we define the length $|f|$ of f to be the number of terms appearing when f is written out as a sum of products.

For instance,

$$|0| = 0,$$
$$|1| = 1,$$
$$|x_1 x_2 + x_3| = 2,$$
$$|x_1 x_2 + x_1 + x_2 + 1| = 4.$$

Furthermore we define the length of f under (G, H), denoted by $|f|_{(G,H)}$, to be the minimum length of all functions in the equivalence class to which f belongs under (G, H).

Definition 2. Let $f:V^k \to V$, and let S_0 and S_1 denote the set of elements in V^k which are mapped into 0 and 1, respectively. Clearly, $S_0 \cup S_1 = V^k$, $S_0 \cap S_1 = \phi$. If s_i denotes the number of elements in S_i, then we define two "measures" μ and $|\mu|$ of f as follows:

$$\mu(f) = s_1 - s_0,$$
$$|\mu|(f) = |s_1 - s_0|.$$

The measure μ was defined and discussed by Hertzig and Dean in Ref. 13, but for our purposes $|\mu|$ will be more convenient.

A function f is called neutral (Dean and Hertzig use the term balanced) if $\mu(f) = |\mu|(f) = 0$. Some properties of neutral functions are given in Refs. 12 and 13. For instance, it is shown in Ref. 12 (as a special case of Lemma 11) that the number of neutral classes of functions $f:V^k \to V$ under $(G_k, 1)$ is greater than or equal to

$$\frac{1}{2^k k!}\binom{2^k}{2^{k-1}} \sim \left(\frac{2}{\pi}\right)^{1/2} \frac{2^{2^k - 3k/2}}{k!}.$$

It is easy to see that the measure $|\mu|(f)$ is invariant under (G_k, N). The length $|f|$, of course, is not invariant; this is the advantage of considering $|\mu|$.

The significance of the foregoing definitions to the theory of algorithmic complexity is given by the following theorems.

Theorem 5. Let $f:V^k \to V$ and let $\ell = |f|_{(G_k, N)}$. Then

$$N_f \leq k(\ell + 1) - 1,$$
$$T_f \leq \{\log_2 k\} + \{\log_2 \ell\} + 1.$$

Proof. Each of the ℓ terms is a product of $\leq k$ variables. Each term, therefore, requires $N \leq 2k - 1$, $T \leq 1 + \{\log_2 k\}$. Since input need not be repeated, all ℓ terms can be computed with $N \leq (2k - 1)\ell - k(\ell - 1) = k(\ell + 1) - \ell$, $T \leq 1 + \{\log_2 k\}$. To add the ℓ terms requires an additional $N \leq \ell - 1$, $T \leq \{\log_2 \ell\}$. Hence

$$N_f \leq k(\ell + 1) - \ell + (\ell - 1) = k(\ell + 1) - 1,$$
$$T_f \leq \{\log_2 k\} + \{\log_2 \ell\} + 1,$$

as required.

Loosely speaking, Theorem 5 states that "short" functions (functions with small ℓ) have low complexities. It consequently is of interest to ask: What is the maximum value $L = L(k)$ of $|f|_{(G_k, N)}$? Substituting this value into Theorem 5 will give an upper bound on the complexity of any function $f:V^k \to V$. It seems reasonable to conjecture that $L(k) = 2^{k-1}$, although this value does not improve previously-obtained bounds. The best we can prove in this article is the following.

Theorem 6. If

$$L(k) = \max_{f:V^k \to V} |f|_{(G_k, N)},$$

then

$$2^{k-3}/(k \log_2 k) < L(k) < 2^k.$$

As a consequence, $\log_2 L(k) \sim k$.

Proof. In the previous article, Theorem 4, we proved that there is a function $f:V^k \to V$ such that all (2, 2) automata computing f require

$$T < k - \log_2 \log_2 k.$$

Then by Theorem 5 we must have

$$\log_2 k + \log_2 L(k) + 3 > \{\log_2 k\} + \{\log_2 L(k)\} + 1 > k - \log_2 \log_2 k,$$

from which it follows that

$$L(k) > 2^{k-3}/(k \log_2 k).$$

To prove the other inequality, note that there are only 2^k possible products of the variables. If all 2^k products are used, then

$$f = 1 + x_1 + \cdots + x_k + x_1 x_2 + \cdots + x_{k-1} x_k + \cdots + x_1 x_2 \cdots x_k = (1 + x_1)(1 + x_2) \cdots (1 + x_k),$$

so that $|f|_{(G_k, N)} = 1$. Hence we cannot have $L(k) = 2^k$, so $L(k) < 2^k$.

Theorem 6 is proved.

The relationship between the measure $|\mu|$ and algorithmic complexity is given by Theorem 7.

Theorem 7. Let $V^k \to V$. If $|\mu|(f) = m$, then there exists a (2, 2) automaton realizing f which satisfies

$$N \leq \tfrac{1}{2}k(2^k - m) + k - 1,$$

$$T \leq \{\log_2 k\} + \{\log_2 (2^k - m)\}.$$

Proof. If $|\mu|(f) = m$, then one of S_0 or S_1 has $\frac{1}{2}(2^k - m)$ elements and the other one has $\frac{1}{2}(2^k + m)$ elements. Say $|S_0| = \frac{1}{2}(2^k - m)$ to be definite. We now use a procedure first used in *SPS 37-34*, Vol. IV, p. 298 to compute f. To determine whether the input is equal to a given fixed element of S_0 requires

$$N \leq 2k - 1,$$

$$T \leq \{\log_2 k\} + 1.$$

Hence to test all $\frac{1}{2}(2^k - m)$ elements of S_0, without repeating input, requires

$$N \leq k + \tfrac{1}{2}(2^k - m)(k - 1),$$

$$T \leq \{\log_2 k\} + 1.$$

At this point we have $\frac{1}{2}(2^k - m)$ stages, of which exactly one is equal to 1 if and only if the input was in S_0; otherwise all $\frac{1}{2}(2^k - m)$ stages are 0. To determine whether one of these $\frac{1}{2}(2^k - m)$ stages is 1 requires an additional N, T as follows:

$$N \leq \tfrac{1}{2}(2^k - m) - 1,$$

$$T \leq \{\log_2 \tfrac{1}{2}(2^k - m)\}$$

$$= \{\log_2 (2^k - m)\} - 1.$$

Therefore, N and T are bounded by

$$N \leq \tfrac{1}{2}(k - 1)(2^k - m) + \tfrac{1}{2}(2^k - m) + k - 1$$

$$= \tfrac{1}{2}k(2^k - m) + k - 1,$$

$$T \leq \{\log_2 k\} + 1 + \{\log (2^k - m)\} - 1$$

$$= \{\log_2 k\} + \{\log_2 (2^k - m)\},$$

which proves the theorem.

The bounds for N and T obtained above are maximized by $m = 0$, but this gives no new information.

Theorem 7 in effect states that highly non-neutral functions have low complexities, but this is not surprising. Theorems 5 and 7 combined suggest that the place to look for algorithmically complex functions is among the nearly neutral functions of long length.

It is useful in formulating and testing hypotheses to have a table describing the equivalence classes of the Boolean functions of k variables under (G_k, N), for small values of k. Table 2 gives this information for $k \leq 3$. Included are representative functions of minimal length in each equivalence class, the number of functions in each class, and the common measure $|\mu|$ of the functions in each class. Also given are the cycle index polynomials for the relevant groups, which are of combinatorial significance (see, for example, Ref. 10, (Chap. 5) and Ref. 15 (Chap. 4) for a discussion of cycle index polynomials, and Refs. 8, 11 and 12 for applications to Boolean functions). The writer is preparing a similar table for $k = 4$.

6. Miscellaneous Criteria

There are numerous other criteria which produce functions of low complexity, but lack of space prohibits going into detail. We shall mention only three such criteria without giving explicit bounds on the complexity, though it is quite easy to obtain low bounds.

(1) There exists an $r \times k$ matrix R of r rows and k columns, $r \leq k$, such that the column vector $\mathbf{x}$ is in S_0 (or S_1) if and only if $R\mathbf{x} = 0$. In particular, this includes the case where S_0 or S_1 is closed under addition. This example has applications to error-correcting encoding and decoding.

(2) The weight (i.e., the number of 1's) of each vector in S_0 (or S_1) is small, e.g., is $o(k)$.

(3) The function f can be expressed as the composition of $r \leq k - 1$ functions $f_1, \cdots, f_r$, each of $< k$ variables, such that no two f_i depend on a common variable. For example, we could have $f(x_1, \cdots, x_k) = g(f_1(x_1, \cdots, x_q), f_2(x_{q+1}, \cdots, x_k))$.

In conclusion, we mention a simple method for constructing a function f if S_0 is given. If $S_0 = \{\mathbf{a}_1, \cdots, \mathbf{a}_r\}$, and if a_{ij} is the j^{th} component (0 or 1) of $\mathbf{a}_i$, then it is easy to see that

$$f(x_1, \cdots, x_k) = \prod_{i=1}^{r} \left(1 + \prod_{j=1}^{k} (1 + a_{ij} + x_j)\right).$$

For example, if $S_0 = \{(0111), (1010)\}$, then $f(x_1, \cdots, x_4) = ((x_1 + 1)\, x_2\, x_3\, x_4 + 1)\, (x_1\, (x_2 + 1)\, x_3\, (x_4 + 1) + 1)$. At times this is a useful canonical form for f.

Table 2. Boolean functions of $k \leq 3$ variables under (G_k, N)

	Representative function of minimal length	Number in class	$\lvert\mu\rvert$
I. $k = 1 : P_{G_1}(x_1, x_2) = \frac{1}{2}(x_1^2 + x_2)$; $P_N(z_1, z_2) = \frac{1}{2}(z_1^2 + z_2)$			
1	0	2	2
2	a	2	0
		4	
II. $k = 2 : P_{G_2}(x_1, \ldots, x_4) = \frac{1}{8}(x_1^4 + 3x_2^2 + 2x_1^2 x_2 + 2x_4)$; $P_N(z_1, z_2) = \frac{1}{2}(z_1^2 + z_2)$			
1	0	2	4
2	a	4	0
3	$a + b$	2	0
4	ab	8	2
		16	
III. $k = 3 : P_{G_3}(x_1, \ldots, x_8) = \frac{1}{48}(x_1^8 + 13x_2^4 + 6x_1^4 x_2^2 + 8x_1^2 x_3^2 + 12x_4^2 + 8x_2 x_6)$; $P_N(z_1, z_2) = \frac{1}{2}(z_1^2 + z_2)$			
1	0	2	8
2	a	6	0
3	$a + b$	6	0
4	ab	24	4
5	$a + b + c$	2	0
6	$ab + c$	24	0
7	$ab + ac$	24	4
8	$ab + bc + a$	24	0
9	$ab + ac + bc$	8	0
10	$ab + ac + bc + a$	8	4
11	abc	16	6
12	$abc + a$	48	2
13	$abc + a + b$	48	2
14	$abc + a + b + c$	16	2
		256	

D. A Useful Partition for Rook Domains

H. Taylor

1. Summary

The background for questions treated in this article is found in Golomb and Posner (Ref. 16). The lemma proved here is dubbed Lemma K because it extends a consequence of Konig's Lemma (Ref. 17) to dimensions $k > 2$. It is the result of an attempt to find a simple proof of the theorem (4) proposed by Posner and proved separately by Rodemich. It is hoped that the method of proof, using a partition, will be useful for further explorations of similar questions. The problem for one-dimensional rook domains is still open (Ref. 18).

2. Background

The notation V_n^K denotes the set of all K-tuples from an n-set of letters. An n-set is a set with n elements. A hyperplane is a $(K-1)$ dimensional subspace of V_n^K or; equivalently, given a letter Z from the n-set of letters and a place i between 1 and K, the set of all those K-tuples of V_n^K which have Z in the ith place is a hyperplane. The union of all the hyperplanes to which a given K-tuple belongs is called the hyper rook domain of the given K-tuple, and we will use the abbreviation HRDα for the hyper rook domain of a K-tuple α. A special abbreviation is $\textcircled{Z}$ for the K-tuple which has the letter Z in every place. Thus, for example, if $K = 5$, then $\textcircled{Z} = (Z,Z,Z,Z,Z)$.

3. The Partition L

Without loss of generality, the set of n letters can be $\{1, \cdots, n\}$ so that the set of letters used in any K-tuple of V_n^K has a natural order from smallest to largest. A given K-tuple α will have, say, ℓ different letters in it.

The ℓ-cycle set to which α belongs is found thus: replace each letter in α by the next larger of the letters in α and replace the largest by the smallest to obtain a K-tuple α_1. Repeat the replacement process with α_1 and continue until $\alpha_l = \alpha$. The ℓ-set of K-tuples thus obtained is called the ℓ-cycle set of α.

The relation $\alpha\epsilon$, the ℓ-cycle set of β, is clearly reflexive, symmetric, and transitive. Hence, it defines a partition of V_n^K, here called the partition L.

The partition L has two useful properties. Suppose α and β are two different K-tuples which belong to the same ℓ-cycle set ϵL. Then:

(1) The ℓ-set of letters in α is the same as the ℓ-set of letters in β. Hence, for any letter Z either α and β both belong to HRD $\textcircled{Z}$ or else neither do, and they both belong if and only if Z belongs to the ℓ-set of letters of α.

(2) For each i from 1 to K the letter in the ith place of α is different from the letter in the ith place of β, so, in consequence, both α and β never belong to the same hyperplane.

4. Lemma K

Let all the K-tuples belonging to V_n^K be designated as θ-cells or χ-cells in an arbitrary way. Choose a maximum m-set M of θ-cells. Choose in such a way that no two belong to the same hyperplane. Put $n = m + r$. Let α be a cell whose HRD is disjoint from M. Then the number of χ-cells in HRDα is not less than $rn^{K-1} - (r-1)(n-1)^{K-1} = n^{K-1} + (r-1)\left[n^{K-1} - (n-1)^{K-1}\right]$.

Equality is obtained in case m parallel hyperplanes are filled with θ-cells, and the r remaining parallel hyperplanes are filled with χ-cells. Also, the number of χ-cells in V_n^K is not less than rn^{K-1}.

Proof. Without loss of generality, suppose $M = \{\textcircled{r+1}, \cdots, \textcircled{r+m}\}$, and suppose $\textcircled{r}$ is a typical cell whose HRD is disjoint from M. For each t from 1 to n put $H_t = \{$all K-tuples ϵV_n^K with t in the first place$\}$. Thus $H_1, \cdots, H_n$ are parallel hyperplanes and $H_{r+1}, \cdots, H_{r+m}$ each have one θ-cell ϵM.

Now suppose we are given an ℓ-cycle set $u \epsilon L$, and suppose that the ℓ-set of letters which appear in each K-tuple ϵu consists of s letters from $\{1, \cdots, r\}$ and ℓ-s letters from $\{r+1, \cdots, r+m\}$.

By property II of L we know that no two cells of u belong to any one hyperplane, and, hence, u has the property required of M. By property I we see that the cells of u belong to the HRD of exactly ℓ-s cells in M. Thus, if more than ℓ-s of thes cells of u were θ-cells, we could use them instead of ℓ-s of the θ-cells in M to form a new set with more than m of the θ-cells, no two in hyperplane, in contradiction to the maximality of m. Hence, it must be true that not more than ℓ-s of the cells in u are θ-cells, so the number of χ-cells in u is not less than s.

By property II we also know that exactly s of the K-tuples in u have a letter from $\{1, \cdots, r\}$ in the first place, so s is equal to the number of cells in $u \cap (H_1 \cup \ldots \cup H_r)$.

Now we know that for each $u \epsilon L$ the number of χ-cells in u is not less than the number of cells in $u \cap (H_1 \cup \ldots \cup H_r)$. By property I, each $u \epsilon L$ is either a subset of HRD $\textcircled{r}$ or else disjoint from it. Taking the union of those $u \epsilon L$ which are subsets of HRD $\textcircled{r}$ we have: the number of χ-cells in HRD $\textcircled{r}$ is greater than or equal to the number of cells in HRD $\textcircled{r} \cap (H_1 \cup \cdots \cup H_r)$.

The number of cells in HRD $\textcircled{r} \cap (H_1 \cup \cdots \cup H_r)$ is easily found to equal $rn^{K-1} - (r-1)(n-1)^{K-1}$, so the proof of lemma K is complete.

5. Theorem

Let each cell in V_n^K be given either the value 0 or else a value greater than or equal to 1, and give each hyperplane a value equal to the sum of its cells' values. For each cell α let $S\alpha$ denote the sum of the values of the hyperplanes to which α belongs. Let S be the sum of the cells' values over all the cells of V_n^K. If the condition that $S\alpha \geq n^{K-1}$ for every cell α whose value is 0 is satisfied, then $S \geq n_K^K$.

Proof. Designate any cell α as a θ-cell, if $S\alpha \geq n^{K-1}$; and designate the remaining cells as χ-cells. Any χ-cell must, thus, have a value no less than 1. Therefore, if β is an χ-cell, then

$$n^{K-1} > S\beta \geq \text{the number of } \chi\text{-cells } \epsilon \text{ HRD}\beta.$$

Consequently, Lemma K tells us that a maximum set of θ-cells with no two in a hyperplane will have no fewer than n elements. Because any cell belongs to exactly K hyperplanes, we can choose a set $N = \{\alpha_1, \cdots, \alpha_n\}$ of θ-cells, which has the property that each hyperplane contained in V_n^K contains exactly one θ-cell of N. This property of N means that $K \cdot S = \sum_{i=1}^{n} S\alpha_i$. But the designation of θ-cells requires that $\sum_{i=1}^{n} S\alpha_i \geq n \cdot n^{K-1}$. Thus, we have $k \cdot S \geq n^K$, which gives the desired conclusion.

References

1. Golomb, S. W., et. al., *Digital Communications With Space Applications,* Prentice Hall, Englewood Cliffs, N.J., 1965.

2. Ryser, H. J., *Combinatorial Mathematics,* Mathematical Association of America (Carus Math. Monograph No. 14), John Wiley and Sons, New York, 1963.

3. Hananai, H., The Existence and Construction of Balanced Incomplete Block Designs, *Ann. Math. Stat. 32,* pp. 361–386, 1961.

4. Bose, R. C., On construction of balanced incomplete block designs, *Ann. Eugen. 9,* pp. 353–399, 1939.

5. Titsworth, R., *Correlation Properties of Random-Like Periodic Sequences,* Progress Report 20-391, Jet Propulsion Laboratory, Pasadena, October 1, 1959.

6. Private communication.

7. Ofman, Yu, "On the Algorithmic Complexity of Discrete Functions," *Soviet Phys.-Doklady, Cyb. Cont. Theor.,* Vol. 7, pp. 589–591 (in translation), 1963.

8. Harrison, M. A., "The Number of Transitivity Sets of Boolean Functions," *Society for Industrial and Applied Mathematics Journal,* Vol. II, No. 3, September, 1963.

9. Birkhoff, G., and MacLane, S., *A Survey of Modern Algebra,* Revised Edition, The Macmillan Company, New York, 1953.

10. Beckenbach, E. F., Editor, *Applied Combinatorial Mathematics,* John Wiley and Sons, Inc., New York, 1964.

11. Harrison, M. A., *On the Classification of Boolean Functions by the General Linear and Affine Groups,* Technical Note, The University of Michigan College of Engineering, September, 1962.

12. Harrison, M. A., "The Number of Equivalence Classes of Boolean Functions under Groups Containing Negation," *IEEE Transactions on Electronic Computers,* Vol. EC-12, Number 5, October, 1963.

13. Hertzig, D., and Dean, R., "Decomposition and Equivalences of Numerical Functions on a Vector Space over GF(2), private communication from Prof. Dean of the California Institute of Technology, Pasadena, California.

14. DeBruijn, N. G., "Generalization of Polya's Fundamental Theorem in Enumerative Combinatorial Analysis," *Koninklijke Nederlandse Akademe van Wetenschappen,* Vol. 62, pp. 59–69, 1959.

15. Riordan, J., *An Introduction to Combinatorial Analysis,* John Wiley and Sons, Inc., New York, 1958.

16. Golomb, S. W., and Posner, E. C., "Rook Domains, Latin Squares, Affine Planes, and Error-Distributing Codes," *Trans. IEEE, Information Theory,* Vol. IT-10, pp. 196–208, 1964.

17. Golomb, S. W., Solomon, W., and Posner, E. C., "Hypercubes of Non-negative Integers," *Bull. Amer. Math. Soc.* (Research Problem No. 10, 1965), p. 587.

18. Hall, M., Jr., "Combinatorial Analysis," Theorem 5.4, California Institute of Technology Bookstore, 1964.

XXII. Communications Systems Research: Communication Statistics

A. An Integral Formula for Limiter Suppression Factor

R. C. Tausworthe

1. Introduction

Davenport's (Ref. 1) evaluation of the signal power appearing at the output of a bandpass-limiting amplifier yields an answer in terms of a hypergeometric function which, although affording a straightforward evaluation on a computer, is somewhat unfamiliar to most engineers. A very simple integral formula for this so-called limiter suppression factor is derived in what follows. Except in limiting cases, the integral still requires computer evaluation, but there is, perhaps, some advantage in the simplicity of its integral form. This work also presents a particularly simple (but rigorous) proof that the total output power in any particular output harmonic zone is independent of the input signal-to-noise ratio (SNR). This fact permits the computation of output SNR's from the suppression factors alone.

2. Calculation of Limiter Suppression

Assume that the input to a bandpass limiter is the waveform

$$x(t) = (2)^{1/2} \sin \left[\omega_0 t + \theta(t) + \phi\right] + n(t)$$

where

ϕ is a uniformly distributed random variable on the interval $(-\pi, \pi)$

$\theta(t)$ is a narrow-band process

$n(t)$ is a narrow-band, zero-mean, stationary Gaussian process with variance σ^2. (Thus σ^2 is then the input SNR.)

The uniform phase distribution of the input sinusoid induces a probability density on its amplitude, viz.

$$p_s(s) = \frac{d}{ds} Pr\left\{(2)^{1/2} \sin \left[\omega_0 t + \theta(t) + \phi\right] \leq s\right\}$$

$$= \begin{cases} \dfrac{1}{\pi (2 - s^2)^{1/2}}; & s^2 < 2 \\ 0 \quad\quad ; & s^2 > 2 \end{cases}.$$

The limiter output $y(t) = l \operatorname{sgn} x(t)$ prior to bandpass filtering can be expressed in the form

$$y(t) = (2)^{1/2} \mu_k \sin\left\{k \left[\omega_0 t + \theta(t) + \phi\right]\right\} + N_k(t)$$

where μ_k is the rms amplitude of some kth signal harmonic coming out of the limiter, and $N_k(t)$ is the remainder

of the output, consisting of noise, distortion, etc., uncorrelated with the signal harmonic. That is,

$$\mu_k = E\{y(t)\, w(t)\}$$

where $w(t)$ is a unit-variance kth signal harmonic

$$w(t) = (2)^{1/2} \sin\left\{k\left[\omega_0 t + \theta(t) + \phi\right]\right\}.$$

In terms of the input signal $s(t)$,

$$s(t) = (2)^{1/2} \sin\left[\omega_0 t + \theta(t) + \phi\right],$$

the harmonic can be expressed as

$$w(t) = (2)^{1/2} \sin\left[k \sin^{-1}\left(\frac{s(t)}{2}\right)\right].$$

The expectation yielding μ_k can thus be written

$$\mu_k = \int_{-2^{1/2}}^{+2^{1/2}} \int_{n=-\infty}^{\infty} l w\, sgn(s+n)\, p_s(s)\, p_n(n) ds dn$$

$$= 4l \int_0^{2^{1/2}} \int_0^{s} w p_s(s)\, p_n(n) dn ds$$

$$= 8l \int_0^1 \int_0^{(2x)^{1/2}} \sin(k \sin^{-1} x)\, p_s[x(2)^{1/2}]\, p_n(n) dn dx .$$

Upon inserting densities for s and n, the result obtained is

$$\mu_k = \frac{2l(2)^{1/2}}{\pi} \int_0^1 \frac{\sin k(\sin^{-1} x)\, erf\left(\frac{x}{\sigma}\right) dx}{(1-x^2)^{1/2}}$$

$$= \frac{2l(2)^{1/2}}{\pi} \int_0^{\pi/2} \sin(ky)\, erf\left(\frac{1}{\sigma} \sin y\right) dy$$

which can be expressed in terms of a confluent hypergeometric function, if desired.

3. Zonal Output Power

Following the method in Ref. 2, we define

$$X_s = (2)^{1/2} \cos\left[\theta(t) + \phi\right] + n_s(t)$$

$$X_c = (2)^{1/2} \sin\left[\theta(t) + \phi\right] + n_c(t)$$

where $n_s(t)$ and $n_c(t)$ are narrowband Gaussian processes such that

$$n(t) = n_c(t) \cos\omega_0 t + n_s(t) \sin\omega_0 t .$$

In this way, the limiter input waveform $x(t)$ is

$$x(t) = X_c \cos\omega_0 t + X_s \sin\omega_0 t$$

$$= (X_s^2 + X_c^2)^{1/2} \sin\left[\omega_0 t + \psi(t)\right]$$

in which $\psi(t)$ is a narrow-band process given by

$$\psi(t) = \tan^{-1} \frac{X_c}{X_s} .$$

Now the limiter output can be expressed as a sum of zonal harmonic components by using the Fourier series form of $sgn(\sin x)$

$$y(t) = l\, sgn\left[\sin(\omega_0 t + \psi)\right]$$

$$= \frac{4l}{\pi} \sum_{m=0}^{\infty} \frac{\sin\left[(2m+1)(\omega_0 t + \psi)\right]}{(2m+1)}$$

Since ψ is a narrow-band process, it follows that the total power in the kth harmonic zone is zero for even values of k, and

$$P_k = \frac{16l^2}{\pi^2 k^2} E\left[\sin^2 k(\omega_0 t + \psi)\right]$$

$$= \frac{8l^2}{\pi^2 k^2}$$

for odd k. The value of P_k is thus independent of σ.

The kth zonal harmonic suppression factor α_k is defined as the relative rms output signal harmonic amplitude in the kth zone. For $k = 1$, $\alpha_1 = \alpha$ is known as the signal suppression factor.

$$\alpha_k = \left(\frac{k\pi}{2l\, 2^{1/2}}\right) \mu_k$$

$$= k \int_0^1 \frac{\sin\left[k(\sin^{-1} x)\right] erf\left(\frac{x}{\sigma}\right) dx}{(1-x^2)^{1/2}}$$

$$= k \int_0^{\pi/2} \sin(ky)\, erf\left(\frac{1}{\sigma} \sin y\right) dy .$$

This is the integral formula which was to be found.

4. Some Limiting Values

For very small values of σ, we replace $erf\,(\sigma^{-1} \sin y)$ by $1 - erfc\,(\sigma^{-1} \sin y)$; but for small σ the only contribution

to the integral made by the *erfc* term occurs at very small values of y

$$\alpha_k = k\int_0^{\pi/2} \sin(ky)\left[1 - erfc\left(\frac{\sin y}{\sigma}\right)\right] dy$$

$$= 1 - k\int_0^{\pi/2} \sin(ky)\, erfc\left[\frac{1}{\sigma}\sin y\right] dy$$

$$\approx 1 - k^2\int_0^{\infty} y\, erfc\,(y/\sigma)\, dy\,.$$

A Taylor series then yields

$$\alpha_k \approx 1 - \frac{k^2}{4}\,.$$

When σ is very large, we replace the error function by its asymptotic small-argument value expansion

$$\alpha_k = k\int_0^{\pi/2} \sin ky\left(\frac{2}{(\pi)^{1/2}}\sum_{n=0}^{\infty}\frac{(-1)^n \sin^{2n+1} y}{n!\,(2n+1)\,\sigma^{2n+1}}\right) dy\,.$$

The least power of $\sin y$, which has a $\sin ky$ component, is the first to contribute to the integral, and since σ is large, for asymptotic results, we need keep only this first term. The first such term is, of course, that containing $\sin^k y$, so only the $2n + 1 = k$ term above is kept.

$$\alpha_k \approx \frac{(-1)^{\frac{k-1}{2}}\, 2}{\left(\frac{k-1}{2}\right)!\,\sigma^k\,(\pi)^{1/2}}\int_0^{\pi/2} \sin^k y \sin ky\, dy$$

$$\approx \frac{(\pi)^{1/2}}{\left(\frac{k-1}{2}\right)!\,2^k\,\sigma^k}\,.$$

In both asymptotic results above, the magnitude of α_k is monotonically decreasing as a function of k.

5. Zonal SNR's

The total output power P_k in the kth zone (k odd), according to the definition above, is related to the harmonic suppression factor as follows:

$$S_k = \mu_k^2 = P_k\,\alpha_k^2\,.$$

Furthermore, μ_k is so defined that noise and distortion at the limiter output are uncorrelated with the kth signal harmonic. Since a zonal filter passes all signals in the kth harmonic zone without loss or distortion, eliminating all other harmonic zones, the output noise power is

$$N_k = P_k(1 - \alpha_2^k)\,.$$

The output SNR is thus a function of α_k only.

$$\frac{S_k}{N_k} = \frac{\alpha_k^2}{1-\alpha_k^2}\,.$$

In the two limiting cases considered, we have

$$\left(\frac{S}{N}\right)_k = \begin{cases} \dfrac{\pi}{\left[\left(\frac{k-1}{2}\right)!\right]^2 2^{2k}}\left(\dfrac{S}{N}\right)_{in}^k & \text{for } \left(\dfrac{S}{N}\right)_{in} \text{small} \\ \dfrac{2}{k^2}\left(\dfrac{S}{N}\right)_{in} & \text{for } \left(\dfrac{S}{N}\right)_{in} \text{large} \end{cases}$$

which agree with previous results inferred from Davenport's paper.

B. Prefix Coding for Minimizing Storage Requirements

I. Jacobs[1] *and L. Kleinrock*[2]

1. Abstract

In many experimental situations, it is a useful data reduction and compression operation to construct an N-sample, K-cell histogram directly from the data at the experimental site (e.g., in a spacecraft). In this article a near-optimum, easily implemented method for assembling, storing, and transmitting such histograms in the form of binary coded sequences is presented. The optimality refers to minimizing the maximum storage in bits and/or transmission rate in bits/sec required to accommodate every possible (N,K) histogram.

The proposed method codes the integer $n(n = 0,1,2, \cdots, N)$ into a binary codeword of length $f(n)$ by use of a prefix code. The code is chosen such that $f(n)$ is an integer approximation to a certain linear function of n.

[1]Resident research fellow on leave from the Electrical Engineering Department of the Massachusetts Institute of Technology, Cambridge.

[2]Consultant from the Engineering Department University of California, Los Angeles.

The optimum code from this class of linear prefix codes is found as a function of N and K and is shown to be asymptotically optimum, compared to the theoretically best code; i.e.,

$$\lim_{N\to\infty} \frac{S}{S_{min}} = 1,$$

where S and S_{min} are the storage requirements for the best linear prefix code and the best of all possible coding methods, respectively. Further, for $N \gg K$, it is shown that $S - S_{min} < K/2$. Figures are included which compare the performance of the optimum linear prefix codes with the best possible code. These codes are near optimum, and their extremely simple implementation is well suited to dynamic (delay line) storage methods. Consequently, they become interesting not only for storage and transmission of histograms but also for other applications such as list storage of variable-length words in digital computers.

2. Introduction

The construction of an N-sample, K-cell histogram is a useful data reduction and compression operation for a wide variety of experiments (Ref. 3). An N-sample, K-cell histogram, which will be referred to as an (N,K) histogram, is defined as a set of K non-negative numbers $n_1, n_2, \cdots, n_K$ such that $n_1 + n_2 + \cdots + n_K = N$. In this article, we present an efficient, easily implemented, method for assembling, storing, and transmitting histograms in the form of binary-coded sequences. The method is asymptotically optimum in the sense of minimizing the maximum storage in bits and/or transmission rate (in bits/sec) required to accommodate every possible (N,K) histogram. It is well suited for dynamic (delay line) storage and has interesting implications to list storage of variable-length words in computers.

One obvious method of storing (N,K) histograms utilizes K blocks of storage, each block containing $\{\log_2 N\}$[3] bits, for a total storage of $K\{\log_2 N\}$ bits. [Since the total number of samples in the histogram is known to be N, one cell may be omitted and $(K-1)$ $\{\log_2 N\}$ bits will suffice.] This method gives constant-length binary representations of each n_k; since some n_k may equal N, we must provide each with $\{\log_2 N\}$ bits. Since at most one of the n_k can be as large (or almost as large) as N, this method is rather inefficient in terms of storage. A second method of storage incorporated into the Quantiler, a machine built at JPL for the purpose of calculating quantiles (Ref. 3 and *SPS 37-27*, Vol. III, pp. 103–112), uses a binary storage register (actually, a delay line) $N + K - 1$ bits long and allows variable-length representations of the n_k. The number of samples in the kth cell is indicated by the number of zeros contained between the $(k-1)^{st}$ and the kth one, $k = 1,2, \cdots, K$, except that the 0th and Kth ones are deleted. The beginning and end of the register are fixed and specified by other means. For small values of N, this second method has an advantage over the first; whereas, for large N, the second method is very inefficient. For example, if $N = 1024$ and $K - 1 = 256$,

$$(K-1)\log_2 N = 2560, N + K - 1 = 1280;$$

whereas, for $N = 16{,}384$, $K-1 = 256$

$$(K-1)\log_2 N = 3584, N + K - 1 = 16{,}640.$$

W. Tveitan (Ref. 4) has suggested a scheme which combines the variable length storage block of the second method and the logarithmic growth of storage with sample size of the first method. In particular, $k(\geq 2)$ ones (rather than the single one of the second method) are used to partition the storage register; a coded binary representation in which k adjacent 1's are not permitted gives the number of samples in each cell. Tveitan has shown that the required storage for the most difficult histogram is approximately $(K-1)[k + \log_a(N/K)]$, in which a is always less than 2 and approaches 2 as k becomes large. With the best choice of k, we have

$N=1024$, $K-1=256$, storage $=1255$ with $k=2$, $a=1.62$;

$N=16{,}384$, $K-1=256$, storage $=2515$ with $k=3$, $a=1.84$.

3. Optimum Storage

The minimum storage size required for (N,K) histograms is easily determined. The number of distinct histograms (partitions of N into K parts) is the binomial coefficient $\binom{N+K-1}{K-1}$. By Stirling's approximation, we have

$$\binom{N+K-1}{K-1} \approx \left(\frac{N+K-1}{2\pi(K-1)N}\right)^{1/2} \times \left(1+\frac{N}{K-1}\right)^{K-1}\left(1+\frac{K-1}{N}\right)^{N}. \quad (1)$$

The minimum binary storage for L distinct items is $\{\log_2 L\}$ bits, which is achievable if a "code book" is

[3] In this article for any real number k, $[k]$ is the integer part of k, while $\{k\}$ denotes the smallest integer greater than or equal to k.

used. Thus, the necessary and sufficient storage, say S_{min}, is approximately

$$S_{min} \approx S^* = (K-1)\log_2\left(1+\frac{N}{K-1}\right) + N\log_2\left(1+\frac{K-1}{N}\right) + \frac{1}{2}\log_2\left(\frac{N+K-1}{2\pi(K-1)N}\right). \quad (2)$$

Assuming $N \gg K \gg 1$, the usual case, yields

$$S^* \approx S^{**} = (K-1)\left(\log_2\left(\frac{N}{K-1}\right) + \log_2 e\right) - \frac{1}{2}\log_2(K-1). \quad (3)$$

For example,

$$N = 1024, K-1 = 256, \quad S^* = 919;$$
$$N = 16{,}384, K-1 = 256, \quad S^* = 1903.$$

We see that the three previously suggested methods fall significantly short of the optimum. Unfortunately, there does not appear to be any simple method, excluding code book lookup (which itself requires extremely large storage), for achieving the exact optimum.

4. Linear Prefix Code

A close approach to the optimum, however, is obtained by the use of prefix code theory (Ref. 5). These are variable-length codes. The prefix condition states that no codeword contains another codeword as its initial portion. As a result, the codewords are uniquely and instantaneously decipherable. A binary codeword satisfying the prefix condition is assigned to each number between 0 and N. The codewords representing the contents of the K cells can then be placed one after the other without the use of additional separators. Note, however, that the same prefix code is used for each cell. This precludes the possibility of achieving the optimum, since a dynamic code based on the remaining samples only might well be superior. However, we show in Sect. 6 that the best of the prefix codes is very nearly optimum.

The problems of choosing a good prefix code remains. It appears reasonable that a code for which all histograms have the same storage requirement is best.[4] This equistorage property is achieved if the length, say $f(n)$, of the code word assigned to the number n is a linear function of n, that is

$$f(n) = an + b. \quad (4)$$

[4]This statement is based on the conjecture that if some histograms have larger storage requirements than others then "nature" will pick the worst of these and cause a larger storage than in the equistorage case.

In this case the storage S required for the histogram $(n_1, n_2, \cdots, n_K)$ is (again observing that the Kth cell need not be sent),

$$\begin{aligned} S &= f(n_1) + f(n_2) + \cdots + f(n_{K-1}) \\ &= a(n_1 + n_2 + \cdots + n_{K-1}) + b(K-1) \\ &\leq aN + b(K-1), \end{aligned}$$

since it is always true that $\sum_{k=1}^{K} n_k = N$. The equality sign is achieved when $n_K = 0$.

Unfortunately, diophantine restrictions preclude the realization of Eq. (4). We approximate this linear relation in Eq. (4) by use of the following assignment when the integers m,b, and l are optimized, as shown below.

$$\begin{aligned} f(0) &= f(1) = \cdots = f(m-1) = b, \\ f(m) &= f(m+1) = \cdots = f(2m-1) = b + l, \\ f\left(\left[\frac{N}{m}\right]m\right) &= \cdots = f(N) = b + \left[\frac{N}{m}\right]l. \end{aligned} \quad (5)$$

It can be seen that the maximum required storage S for $K-1$ cells is then

$$S = \left[\frac{N}{m}\right]l + (K-1)b, \quad (6)$$

and is achieved, for example, when all but one of the cells contain zero. Note that the assignment $f(n)$ as a function of n has a "slope" equal to l/m.

Now it is necessary to select the integers l, m, and b such that two conditions are met: (1) the code is realizable, and (2) S is minimized. The realizability (prefix) condition requires only that the Kraft inequality (Ref. 6) for unique decipherability,

$$\sum_{n=0}^{N} 2^{-f(n)} \leq 1, \quad (7)$$

be satisfied. We weaken this condition by setting the upper limit equal to ∞, thereby assigning codewords to arbitrarily large numbers, and increasing the code word lengths.

We have, from Eqs. (5) and (7) (with N set equal to ∞)

$$m\left[2^{-b} + 2^{-(b+l)} + 2^{-(b+2l)} + \cdots\right] \leq 1$$

or

$$m2^{-b} \leq 1 - 2^{-l}. \quad (8)$$

Whenever m, b, and l are such that Eq. (8) is satisfied, the code can be realized.

Next we consider choosing m or b to minimize S, the maximum required storage, while holding ℓ fixed and satisfying Eq. (8). Since for fixed N,ℓ,K and b, we see from Eq. (6) that S is nonincreasing with increasing m, we take Eq. (8) to be satisfied with the equal sign.

$$m = 2^b (1 - 2^{-\ell}) = 2^b - 2^{b-\ell} . \tag{9}$$

From Eq. (9), as long as $b \geq \ell$, we find that m is an (nonnegative) integer whenever b is an integer. Since, on the other hand, Eq. (9) shows that b is not necessarily an integer whenever m is an integer, for purposes of optimization we hereafter consider b to be the independent variable with m determined by Eq. (9).

Next, observe from Eq. (6) that

$$\left(\frac{N}{m} - 1\right)\ell + (K-1)b < S \leq \frac{N\ell}{m} + (K-1)b . \tag{10}$$

The lower bound is useful as long as $N \geq m$. Therefore, we select b to minimize $N\ell/m + (K-1)b$. Treating b as a continuous variable and holding ℓ, N, and K fixed, we have

$$\frac{d}{db}\left(\frac{N\ell}{m} + (K-1)b\right) = -\frac{N\ell}{m^2}\frac{dm}{db} + (K-1) . \tag{11}$$

However, from Eq. (9),

$$\frac{dm}{db} = (\ell n\, 2)\, 2^b (1 - 2^{-\ell}) = m/\log_2 e . \tag{12}$$

Thus, from Eqs. (11) and (12), the derivative is zero whenever

$$-\frac{N\ell}{m \log_2 e} + (K-1) = 0 . \tag{13}$$

Furthermore, the second derivative is

$$\frac{d^2}{db^2}\left(\frac{N\ell}{m} + (K-1)b\right) = \frac{N\ell}{m (\log_2 e)^2} > 0 ,$$

so that the function is concave.

We conclude from Eqs. (9) and (13) that $\left(\frac{N\ell}{m} + (K-1)b\right)$ is minimized by setting $b = b_0$ where

$$b_0 = \log_2\left(\frac{N}{K-1}\right) + \log_2\left(\frac{\ell}{1-2^{-\ell}}\right) - \log_2 \log_2 e . \tag{14}$$

Of course, b_0 may not be an integer. Thus, use of b_0 provides a further lower bound to S

$$S \geq S_\ell ,$$

in which we define [see Eqs. (9), (10), (13), and (14)]

$$S_\ell = (K-1)\left(\log_2 e + \log_2\left(\frac{N}{K-1}\right) + \log_2\left(\frac{\ell}{1-2^{-\ell}}\right) - \log_2 \log_2 e\right) - \ell . \tag{15}$$

Recall that this lower bound is useful as long as $N \geq m$; in other words,

$$\frac{N}{m} = \frac{(K-1)\log_2 e}{\ell} \geq 1 . \tag{16}$$

Thus, for $\ell = 1$ the bound is useful whenever $K > 1$. When this condition is violated, we see from the assignment in Eq. (5) that $S = (K-1)b$, and we are using a constant-length code (of length b). In this case, the best choice of b is $\{\log_2 N\}$ and so

$$S \geq (K-1)\{\log_2 N\} . \tag{17}$$

Next we determine upper bounds for the case $\ell = 1$ and show that this choice of ℓ is optimum in all cases of interest. For $\ell = 1$,

$$b_0 = \log_2 \frac{N}{K-1} + 1 - \log_2 \log_2 e . \tag{18}$$

For our upper bound argument, we must use integer values of b. Suppose we round off b_0 to the nearest integer, say b_1. Thus,

$$b_0 - \frac{1}{2} < b_1 \leq b_0 + \frac{1}{2} .$$

Because of the concavity of $N/m + (K-1)b$ as a function of b, the use of b_1 must produce a smaller value of S than the larger of S_+ and S_-, where S_+ is in the upper bound for $\ell = 1$ resulting from the use of $b_0 + \frac{1}{2}$ and S_- is the upper bound resulting from the use of $b_0 - \frac{1}{2}$. From Eq. (10),

$$S_+ = \frac{N}{m_+} + (K-1)(b_0 + \tfrac{1}{2}) ,$$

in which [from Eq. (9) with $\ell = 1$ and $b = b_0 + \frac{1}{2}$ and Eq. (18)] ,

$$m_+ = 2^{b_0 + \frac{1}{2} - 1} = (2)^{\frac{1}{2}} \frac{N}{K-1}\frac{1}{\log_2 e} .$$

It follows that

$$S_+ = (K-1)\left[(2)^{-\frac{1}{2}} \log_2 e + b_0 + \tfrac{1}{2}\right]$$
$$= (K-1)(b_0 + c_1), c_1 = 1.52 \cdots .$$

Similarly,

$$S_- = (K-1)[(2)^{1/2}\log_2 e + b_0 - 1/2]$$
$$= (K-1)(b_0 + c_2),\ c_2 = 1.54\cdots$$

Thus, $S_- > S_+$. Hence, for $\ell = 1$,

$$S \leq S_- = (K-1)(\log_2\left(\frac{N}{K-1}\right) + (1 + c_2) - \log_2\log_2 e)\,. \tag{19}$$

We now determine, for each ℓ, the values of K for which the lower bound S_ℓ, $\ell > 1$, exceeds the upper bound S_- for $\ell = 1$. For such ℓ and K, the choice $\ell = 1$ will therefore be optimum. Consider the difference [see Eqs. (15) and(19)]

$$S_\ell - S_- = (K-1)\left(\log_2 e - (1 + c_2) + \log_2\frac{\ell}{1-2^{-\ell}}\right) - \ell$$
$$= (K-1)\left(\log_2\left(\frac{\ell}{1-2^{-\ell}}\right) - c_3 - \frac{\ell}{K-1}\right),$$
$$c_3 = 1 + c_2 - \log_2 e = 1.1\cdots \tag{20}$$

We wish to find those K (for ℓ, a fixed positive integer) which makes the above difference positive. From Eq. (16), we have

$$\frac{\ell}{K-1} \leq \log_2 e = 1.44\cdots.$$

[for $[\ell/(K-1)] > \log_2 e$ we get a constant-length code]. Using this last inequality in Eq. (20), we see that for $\ell \geq 6$, the difference $S_\ell - S_-$ goes positive for any value of $K(>1)$, and so $\ell = 1$ is preferable to all $\ell \geq 6$.

For values of $1 < \ell < 6$ we solve for those K and ℓ which make $S_\ell - S_- < 0$. Carrying out this calculation, we get the results shown in Table 1. This table lists those values of K (for a given ℓ) for which $\ell = 1$ might not be optimum.

Table 1. Values of K for which $\ell = 1$ might not be optimum

ℓ	2	3	4	5	6,7,8,···
K	$K<8$	$K<6$	$K<6$	$K<5$	none

We have already assumed that $K \gg 1$ in the analysis; consequently, since for all $K \geq 6$, we have that $\ell = 1$ is optimum, we have essentially shown that $\ell = 1$ is optimum for all ranges of interest.

We now address ourselves to the question as to whether or not b_0 should be rounded up or down to an integer. Accordingly, define

$$b_U = \{b_0\};$$
$$b_L = [b_0].$$

Clearly

$$b_U = b_L + 1\,. \tag{21}$$

Defining [cf Eq. (6)] the storage when using b_U and b_L, we get

$$S_U = \left[\frac{N}{m_U}\right] + (K-1)b_U\,;$$
$$S_L = \left[\frac{N}{m_L}\right] + (K-1)b_L\,,$$

where $m_U = 2^{b_U-1}$ and $m_L = 2^{b_L-1}$. Forming the difference between these two equalities, and using Eq. (21), we obtain

$$S_L - S_U = \left[\frac{N}{m_L}\right] - \left[\frac{N}{m_U}\right] - (K-1)$$
$$= \left[\frac{N}{2^{b_L-1}}\right] - \left[\frac{N}{2^{b_L}}\right] - (K-1)\,.$$

As an approximation, we remove the integral constraints represented by square brackets and obtain

$$S_L - S_U \cong N\,2^{-b_L} - (K-1)$$
$$= 2^{\left\{\log_2 N - \left[\log_2\left(\frac{N}{K-1}\right) + 1 - \log_2\log_2 e\right]\right\}} - (K-1)\,.$$

When this difference is positive, we should round up to b_U; conversely for b_L, and so we inquire as to when the following inequality holds.

$$2^{\left\{\log_2 N - \left[\log_2\left(\frac{N}{K-1}\right) + c_4\right]\right\}}$$
$$\geq K-1;\ c_4 = 1 - \log_2\log_2 e = 0.47\cdots.$$

Taking logs,

$$\log_2 N - \left[\log_2\left(\frac{N}{K-1}\right) + c_4\right] \geq \log_2(K-1)\,,$$

or

$$\log_2\left(\frac{N}{K-1}\right) \geq \left[\log_2\frac{N}{K-1} + c_4\right]. \tag{22}$$

Ineq. (22) gives the (approximate) condition for rounding up to $\{b_0\}$. If the fractional part of $\log_2(N/K-1)$ is less than $1-c_4$, round down to $[b_0]$; otherwise, round up. Note that whenever $N/(K-1)$ is a power of two, it makes almost no difference if we round up or down. We can always decide which way to round off either by using Eq. (22) or by checking both possibilities as to their storage requirements.

5. The Optimum Codes

In the previous section, we showed that $\ell = 1$ is optimum (for all $K \geq 6$). Here we repeat the equations which define the optimum codes, describe their construction, and give an example.

Given the numbers N and K, we find the quantity b_0 from Eq. (18)

$$b_0 = \log_2 \left(\frac{N}{K-1}\right) + 1 - \log_2 \log_2 e \, .$$

We then choose b as follows:

$$b = \begin{cases} \{b_0\} & \text{if } \log_2 \left(\frac{N}{K-1}\right) \geq \left[\log \frac{N}{K-1} + c_4\right]; \\ [b_0] & \text{otherwise.} \end{cases} \tag{23}$$

Thus, having found b, we then obtain the integer m from Eq. (9) with $\ell = 1$, viz,

$$m = 2^{b-1} \, . \tag{24}$$

The lengths of all code words are now defined in terms b, m (and $\ell = 1$) from Eq. (5) as

$$f(0) = f(1) = \cdots = f(m-1) = b$$

$$f(m) = f(m+1) = \cdots = f(2m-1) = b+1$$

$$\vdots$$

$$f\left(\left[\frac{N}{m}\right] m\right) = \cdots = f(N) = b + \left[\frac{N}{m}\right] \tag{25}$$

Having determined the lengths of all code words, we now describe the method of their construction. (They exist due to the prefix condition.) A prefix code is a tree code (Ref. 7), and can thus be drawn as a tree. Terminal nodes in the tree represent code words. Branches between nodes represent binary digits, a branch of positive slope representing a zero and a branch of negative slope representing one. Our codes have groups of m code words where the word length of each group increases linearly, as shown in Fig. 1. The binary representation for a code word is found by recording the binary digits found on the path leading from the root of the tree to the terminal in question.

In Fig. 1 we choose to assign the integers $n (n = 0,1,2, \cdots, N)$ such that 0 corresponds to the uppermost code word, 1 to the next uppermost, etc., until N corresponds to the lowermost code word. (This assignment is shown in Fig. 1 as the integer in parenthesis.) For easier implementation, we choose to move the bth digit (counting from right to left) to the rightmost position for each of these code words; this does not destroy the prefix property, since this digit is always a zero. Henceforth, we discuss the code words in this altered form. If we list the coded representation of the integer n, we obtain a list such as that shown in Table 2.

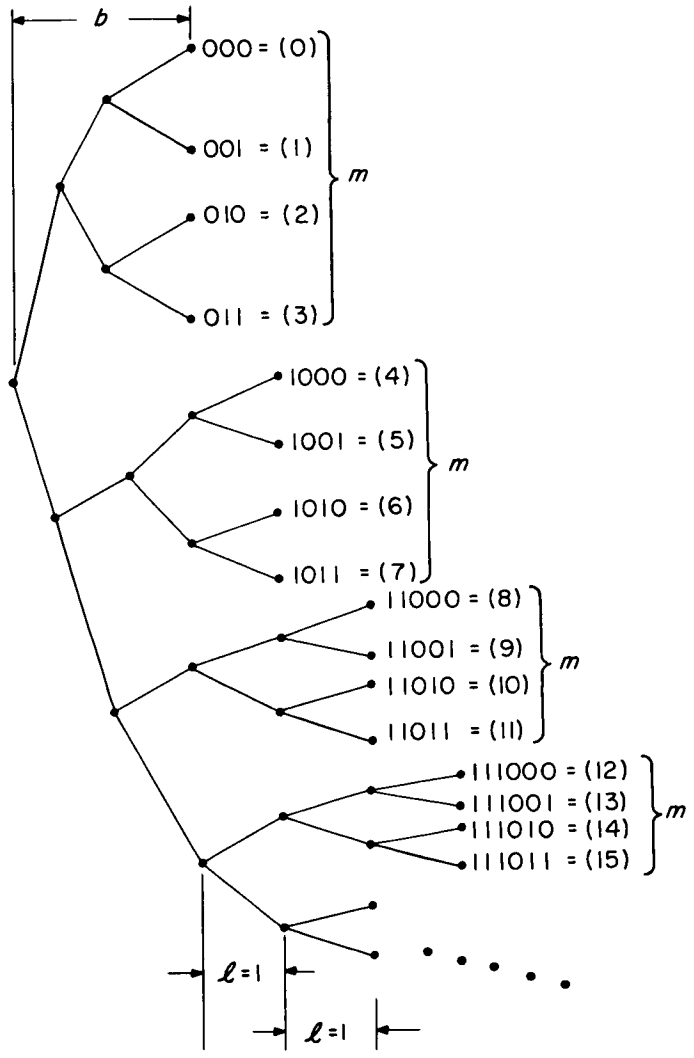


Fig. 1. The linear prefix codes represented as a tree

In general, as shown by Eq. (25), there are m words of length b, then m words of length $b+1$, $\cdots$, then m words of length $b + \frac{N}{m} - 1$, and finally some (up to a maximum of m) words of length $b + \frac{N}{m}$ Counting from right to left, we see that the first digit is always a zero. The $b-1$ digits to the left of this zero proceed successively from the $(b-1)$-bit representation of zero to the $b-1$ bit representation of $m - 1 = 2^{b-1} - 1$ as n grows

Table 2. The linear prefix code

n	Coded version of n	Length f(n)	
	000	3 = b	
1	010	3 = b	m
2	100	3 = b	
3	110	3 = b	
4	1000	4 = b+1	
5	1010	4 = b+1	m
6	1100	4 = b+1	
7	1110	4 = b+1	
8	11000	5 = b+2	
9	11010	5 = b+2	m
.	.	.	
.	.	.	
.	.	.	

from zero to $m-1$. For $\eta = m$, the $b-1$ bits "overflow," producing a 1 in the $(b+1)^{st}$ bit, and resetting the first b bits to zero. The count on these bits then climbs to $n = 2m$, at which point a new overflow is generated; this second overflow produces an additional 1 in the $(b+2)^{nd}$ position and resets the first b bits to zero, etc. Thus, in the coded version of the integer n, the number, say β_n, of 1's to the left of the b^{th} digit reveals how many overflows have occurred (thereby giving the quotient $[n/m]$). The first bit is always zero, and the $b-1$ bits to the left of the zero give the usual binary representation of the remainder, say γ_n. That is, in general,

$$n = m\beta_n + \gamma_n.$$

The first digit is always zero and serves as a natural way to separate code words which are written side by side. For example, given that $b=4$, the code word 1111010 decodes into the integer $3(8)+5 = 29$ (i.e., $m = 2^{b-1} = 8$).

Let us carry out the construction of a particular code. In one of the examples of Sect. 2, we considered $N = 1024$, $K-1 = 256$. For those values, we get, from Eq. (18),

$$b_0 = \log_2\left(\frac{1024}{256}\right) + 0.47 = 2.47\,.$$

Thus, since $\log_2(N/K-1) = [(\log_2(N/K-1) + 0.47]$, we get $b = \{b_0\} = 3$. Also, $m = 2^{b-1} = 4$. For $n = N = 1024$, we get the code word $\underbrace{111\cdots\cdots11}_{256\text{ "ones"}}000$. Bounds on the maximum storage S can be obtained from Eq. (10), which gives

$$1023 < S \leq 1024$$

In this case, we easily find histograms which require the storage given by the upper bound, e.g., the histogram containing one entry at $n = N = 1024$, and $K-1 = 256$ entries at $n = 0$, giving $S = 1 \cdot f(1024) + (255)f(0) = 259 + 255(3) = 1024$.[5] Comparing this storage size to the ideal minimum size of $\left\{\log_2\binom{N+K-1}{K-1}\right\} \cong S^* = 919$ and also to the other known schemes (as mentioned in Sect. 2) which gave values of 2560, 1280, and 1255, we see that the method suggested here is a considerable improvement over other previously considered schemes, but it still falls somewhat short of being ideal.

The JPL code described earlier in Sect. 2 (in which the number of samples in the k^{th} cell is indicated by the number of zeros contained between $(k-1)^{st}$ and k^{th} one i.e. k^{th} one) is a special case of the codes described here for which $b = 1$; such codes are optimum when $N/(K-1) < 2$.

6. Performance

In Sect. 3 we showed that the minimum storage requirement for an (N,K) histogram was

$$S_{min} = \left\{\log_2\binom{N+K-1}{K-1}\right\}, \qquad (26)$$

and approximations to this were given in Eqs. (2) and (3). In Sect. 4, we showed that for the optimum fixed prefix code, the required storage S, was[6]

$$S = \left[\frac{N}{m}\right] + (K-1)b\,.$$

For the moment, we take the liberty of approximating S by a quantity $\bar{S}$ in which we replace b by b_0 and $\left[\frac{N}{m}\right]$ by $\frac{N}{m}$, giving

$$\bar{S} \cong S = \frac{N}{2^{b_0-1}} + (K-1)b_0$$

$$= \frac{N}{2^{\log_2\left(\frac{N}{K-1}\right)-0.53}} + (K-1)\left(\log_2\left(\frac{N}{K-1}\right) + 0.47\right),$$

or

$$\bar{S} = (K-1)\log_2\left(\frac{N}{K-1}\right) + 1.92\,(K-1)\,. \qquad (27)$$

[5]Recall that we agreed not to store the contents of the K^{th} cell in the histogram, since $n_K = N - \sum_{k=1}^{K-1} n_k$.

[6]$\ell = 1$, as discussed in Sect. 4.

Comparing $\bar{S}$ and S** (see Eq. 3) we see that the prefix codes here described have a storage requirement which differs from the minimum by approximately $(1.92 - \log_2 e)(K - 1) = 0.48(K - 1) < K/2$ bits. Asymptotically as $N \to \infty$, we have, for any fixed K,

$$\lim_{N \to \infty} \frac{\bar{S}}{S^{**}} = 1 . \tag{28}$$

In this asymptotic sense, our family of linear prefix codes is optimum.

In Figs. 2 and 3 below, the minimum storage S* is compared with the storage S obtainable by our codes.[7] Fig. 2 shows S and S* as a function of N when $K-1 = 2^k$ (k an integer). For this form, we get (when $N=2^j$, j an integer),

$$S = (K-1)\log_2 N + (K-1)\left[2 - \log_2 (K-1)\right] ,$$

giving S as a simple linear function of $\log_2 N$. We see that S and S* are extremely close, as predicted; and their difference grows approximately as $K/2$.

In Fig. 3 S and S* are shown as a function of K with $N = 2^j$ (j an integer). Again we observe that the prefix codes are near optimum.

[7] We consider $N \geq 2(K-1)$ as the only range of interest.

7. Implementation

In order to achieve the absolutely minimum storage S_{min}, we would require an extremely complex coding

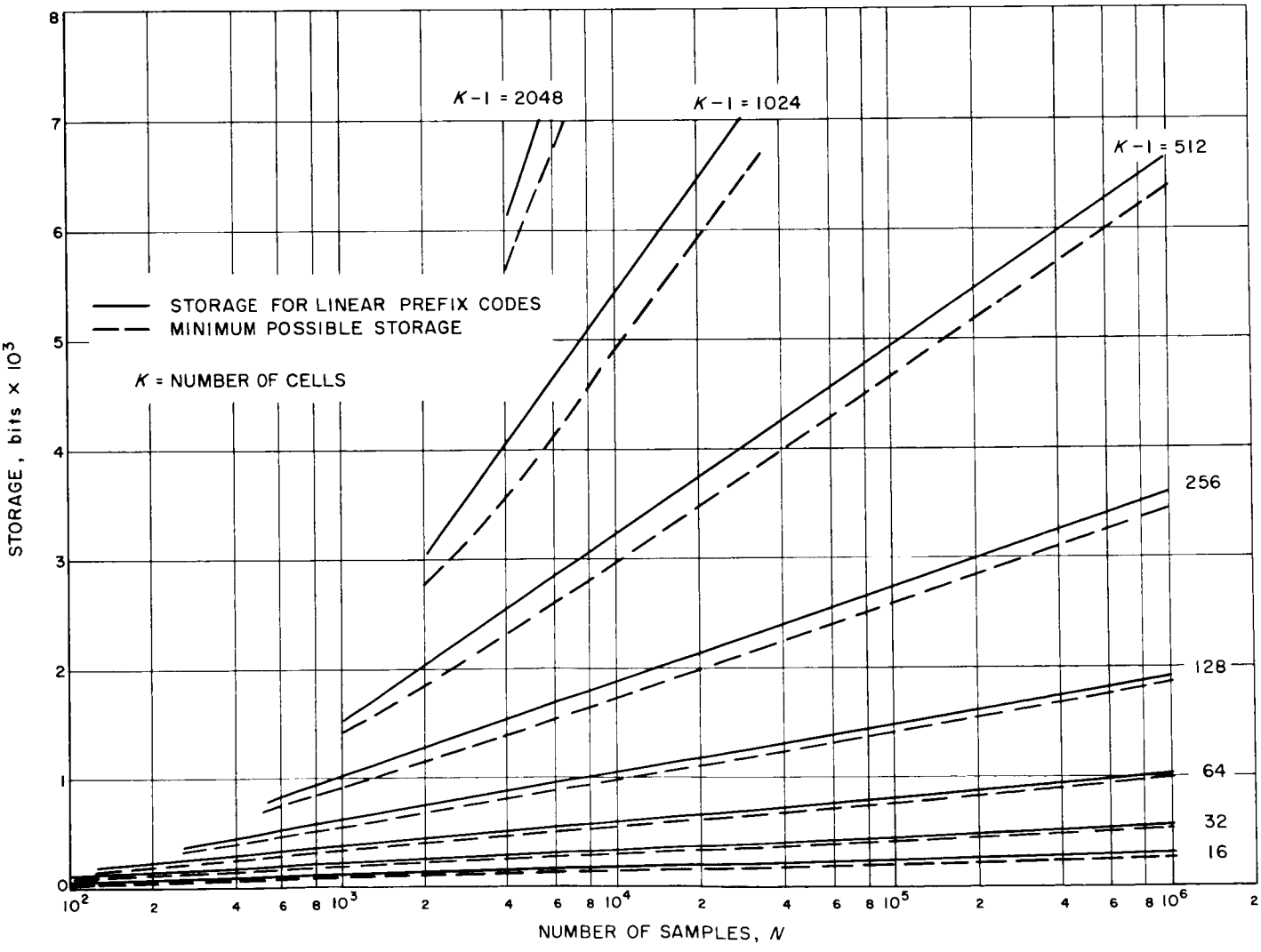


Fig. 2. Storage requirements for linear prefix codes compared to theoretical minimum

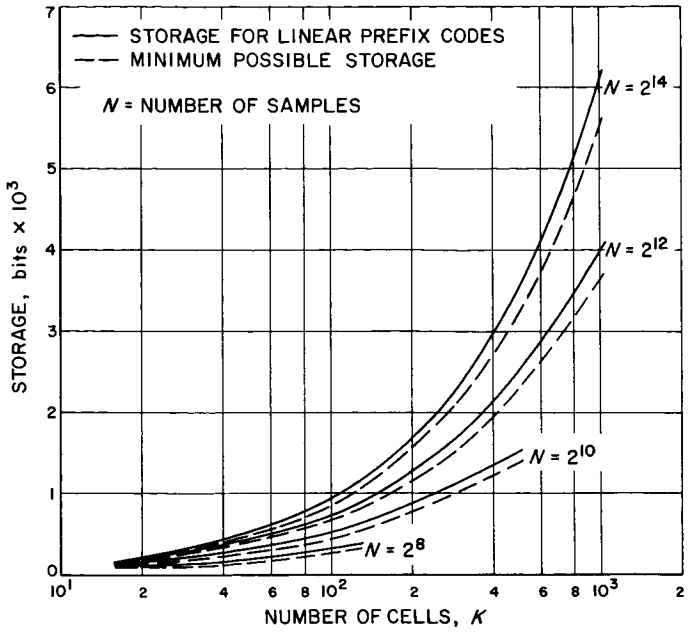


Fig. 3. Storage requirements for linear prefix codes compared to theoretical minimum

method. By considering fixed prefix coding, we have shown that the storage penalty is on the order of $K/2$ bits. This penalty is rather insignificant compared to S_{min} when $N \gg K$. In accepting this suboptimality, we generate codes which are extremely simple to implement, and thus they become most interesting for practical applications as well as for theoretical interest. In this section one method of carrying out the implementation is discussed.

Given N and K, we calculate the number b and m as discussed in Sect. 5. In general, we need not store the value n_K since $\sum_{k=1}^{K} n_k = N$. However, since it simplifies the implementation somewhat, we choose to store all values of n_k. (The storage cost is thereby obtained by replacing $K-1$ by K in our equations.)

In Fig. 4 we show a diagram which describes an implementation of the method for generating the coded histogram. The basic unit is an S-bit delay line buffer that is initially set to all zeros. Each time a new sample is taken in the experiment generating the histogram, the cell into which it falls is stored as the new cell number. The delay line then begins to shift bits out to the right; these bits bypass the unit delay box and are fed into (and through) the $(b-1)$-bit adder. As soon as the "zero-tester" detects a zero, it disables itself for $b-1$ bits since the next $b-1$ bits will be the remainder γ_n. At the same time, a one is added to the cell counter. If the contents of the cell counter do not match that of the new cell number, then the shifting continues as before; after $b-1$ bits, the zero tester begins testing again. For example, a bit stream might look as follows (for $b = 4$):

```
                                                  ↓ start
···0001111100101110100111110010011011 0 → direction of
    *       *       *       * * *               bit flow.
```

The asterisks indicate locations at which the zero tester will detect zeros. The coded numbers represented in this example are

$$\cdots, 41, 26, 23, 2, 19 \quad \downarrow \text{start}$$

Eventually, the cell counter and the new cell number will agree, at which time a 1 will be added to the $b-1$ bits in the adder,[8] since we must increase the count in this cell by 1. If no overflow occurs from this addition, the shifting continues as usual. If an overflow occurs we must insert a 1 directly after the remainder γ_n; this insertion is accomplished by switching the bit stream through a unit delay whose contents have been set to 1 initially. (It is this 1 which is inserted.) The rest of the bit stream is now shifted through the unit delay until all S bits have come through. We have thus increased by 1 the "score" of the cell in which the most recent experimental result occurred. We now wait for a new experimental result and repeat. After N such cycles, the histogram is complete, and the coded form is ready for any processing or transmission.

8. Conclusion

In this article we have considered the problem of finding an efficient method of storing N-sample K-cell histograms. In Sect. 2, we derived an expression for the theoretically minimum storage S_{min}. We pointed out that previously considered codes fell far above this minimum, and also, it appeared that to attain the minimum storage, one would have to go to extremely complex codes or else to code book lookup (which involves large storage). In Sect. 4 we showed that linear prefix codes[9] were capable

[8]This adder can be a $(b-1)$-bit parallel adder as shown, or it can be a one bit serial adder which carries out the addition for $b-1$ shifts and then tests for overflow.

[9]The $N+1$ code words in these codes are used to designate the integers $n = 0,1,2, \cdots, N$. The length $f(n)$ of the nth word is a simple linear function of n (to an integer approximation). Further, no code word is a prefix for any other code word, hence the name linear prefix code.

of coming close to the optimum. Indeed, for K fixed, we have shown that

$$\lim_{N\to\infty} \frac{\bar{S}}{S_{min}} = 1,$$

where S is an approximation asymptotically correct to second order to the storage requirement of the linear prefix codes. Further, for $N \gg K$, the difference in storage between the optimum scheme and the method proposed here is less than $K/2$ bits. Figs. 2 and 3 illustrate this suboptimality of the linear prefix codes.

The key point in considering these codes is that they are almost trivial to implement. An example of the code

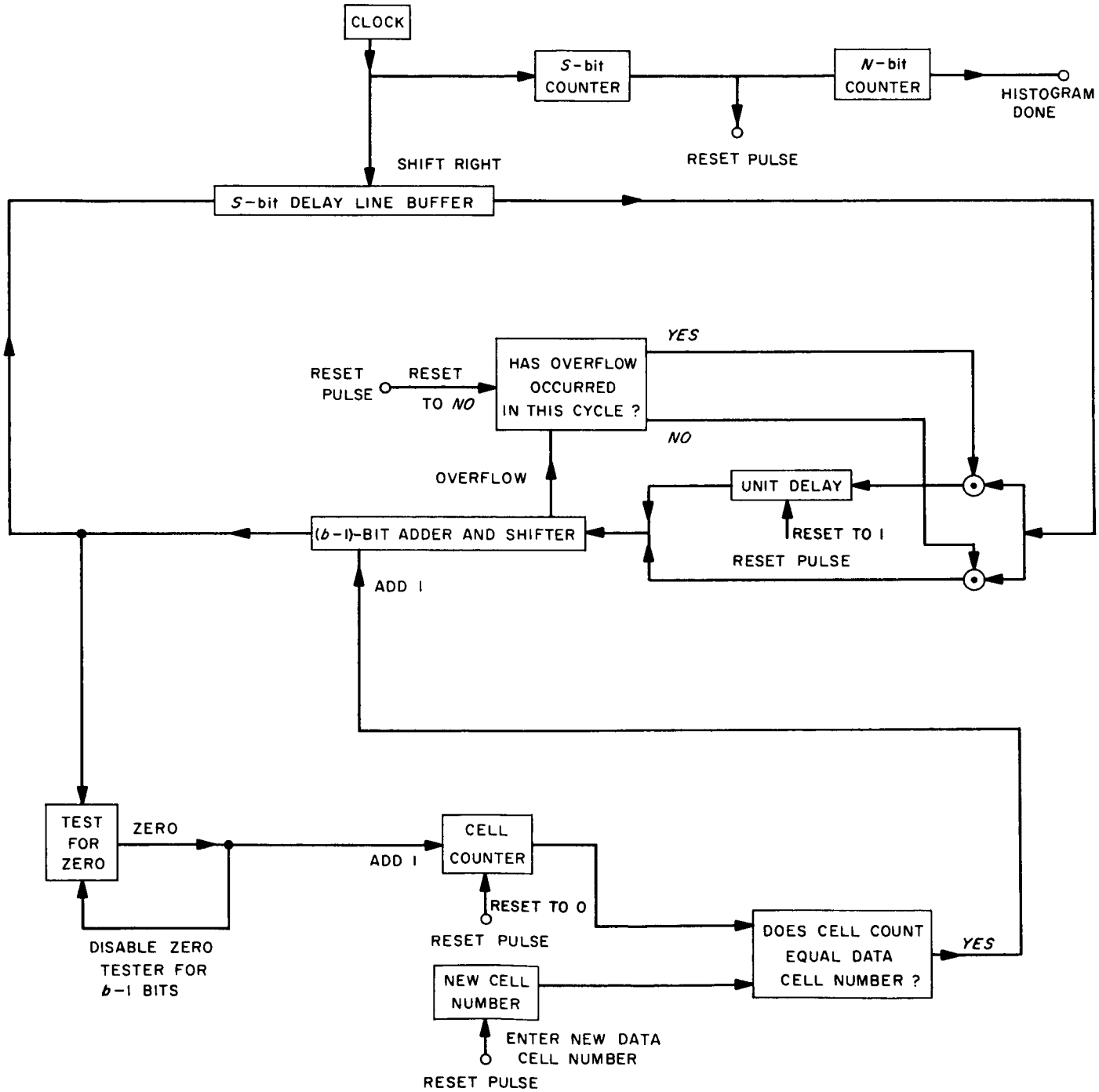


Fig. 4. Implementation of the coded histogram

tree is given in Fig. 1, and the list of code words is shown in Table 1. A block diagram illustrating the implementation of these codes is given in Fig. 4. The coded histogram is constructed "on the fly" as the new data are brought in.

Thus, we have shown how to construct an easily implemented, near-optimum code for storing any (N,K) histogram. If the histogram is to be transmitted over some communication channel, then these codes come close to minimizing the transmission of data on such a channel needed for transmitting histograms.

C. Lunar Radar Maps in Polar Coordinates

T. Pitcher

1. Summary

This article concerns a study of the error term in a lunar radar mapping technique suggested by the aperture-synthesis techniques of the radio astronomers. The purpose is to take advantage of the diurnal rotation of the Moon's apparent axis of rotation relative to an Earth-bound radar. Occasionally, one can obtain the Fourier transform of the radar brightness map of the Moon, evaluated at a discrete set of points on spokes emanating from the Moon's sub-Earth point; these measurements are corrupted by noise. Estimates for the error in mapping the Moon by this process are found; a sum of an approximation error due to the discrete data and due to noise is obtained. One of the two estimation techniques proposed is suitable for use with JPL's monostatic Venus station radar. A method is given for finding the best choice of experimental parameters to use in order to minimize the bound for the expected maximum error in the estimate of the brightness map.

2. Background

This material follows a suggestion by John H. Thompson of Jodrell Bank, University of Manchester, England (Ref. 8), and is an adaptation of the aperture-synthesis technique of the British school of Radio Astronomers (Ref. 9).

Consider the dish of the Moon for radius r. Let the (apparent) axis of rotation of the Moon relative to the Earth-based radar be at angle $\theta(0 \leq \theta \leq 2\pi)$ with respect to some fixed direction (the vertical of Fig. 5) at a certain time. Let the angular rotation vector be Ω and suppose, for convenience, units so chosen that a round-trip doppler shift of $2f$ occurs at distance f from the sub-radio point origin. A line of constant doppler shift $2f$ is shown in Fig. 5.

In the actual situation, the function of doppler shift versus distance from subradar point changes during the day due to the fact that the magnitude of Ω changes. This effect is ignored here; it can be compensated for by changing the sampling rate.

Let the brightness map of the Moon be denoted by $g(P)$, at point P on the Moon's surface; a coordinate free notation is used to avoid confusion later on. This brightness map is the fraction of power received from point P using a radar uniformly illuminating the Moon's disk. This map takes into account both the reflectivity and the scattering law, but we are not trying to separate the two effects in these experiments.

We have a received waveform $s(t)$, and have a correlator (Ref. 10) which measures the correlation $R(\theta,\tau)$, where $R(\theta,\tau)$ is the long term average of $s(t)s(t+\tau)$; $R(\theta,0)$ is the total received power; and, hence, it is the integral of the brightness map over the entire lunar disk, corrupted, however, by receiver noise.

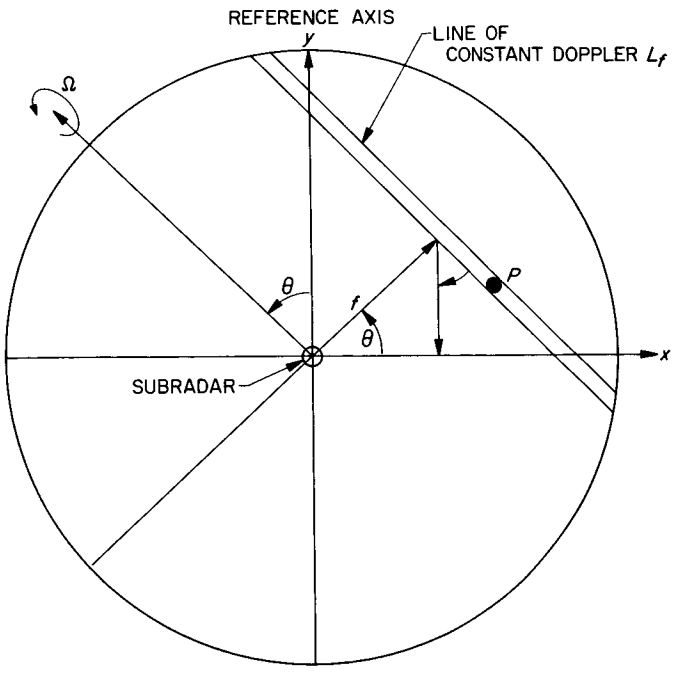


Fig. 5. Explanation of Fourier transform

The power $P_\theta(f)$ at shifted frequency f in the received signal yields the covariance function $R(\theta,\tau)$ as a Fourier transform. That is,

$$R(\theta,\tau) = \int_f P_\theta(f)\, e^{2\pi i f\tau}\, df\,.$$

On the other hand, $P_\theta(f)$ is also the received power from the line f at distance f from the subradar point

$$P_\theta(f) = \int_{l_f} g(P)ds\,.$$

Hence, reverting to $g(x,y)$ notation, we have

$$R(\theta,\tau) = \int_f \int_s g(x,y)\, e^{2\pi i f\tau} df ds\,.$$

But $f = x\cos\theta + y\sin\theta$; also $dfds = dxdy$, since $(x,y)\rightarrow(f,s)$ is an orthogonal transformation. Consequently,

$$R(\theta,\tau) = \int_{x,y} \int g(x,y)\, e^{2\pi i(x\cos\theta + y\sin\theta)\tau} dxdy\,;$$

is the required expression. This is the basis of the experimental technique.

Measurements are taken at a density proportional to the angular rotation rate of the apparent axis of rotation of the Moon, which makes a full revolution on these special days. Although this rotation rate is not uniform, we shall assume that it is, to get spokes uniformly spaced around the circle of θ. We then have $R(\theta,\tau)$ evaluated on spokes uniformly spaced around the circle; on each spoke, the values are known at uniformly spaced points (Fig. 6). The spoke for $R(\theta,\tau)$ makes angle θ with the coordinate direction corresponding to $\tau\cos\theta$ in transform space.

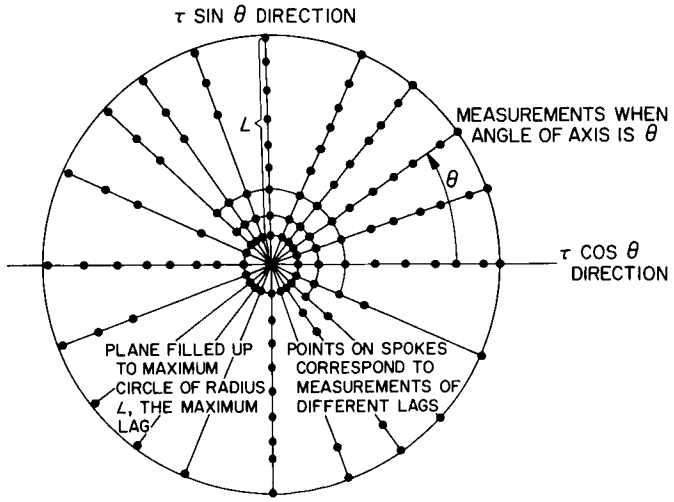


Fig. 6. Spokes of measurements

Another problem worth considering is the one of using this technique on days when the apparent axis does not make a complete rotation. The question would then be, how should estimates of the brightness map on different days with different partial revolutions of the apparent axis be put together?

3. L_2-Estimate

This section estimates the square error in estimates obtainable as above for the brightness map of the Moon. In the noise free case, the estimate gives the minimum square error obtainable. In the presence of noise, the use of windows would reduce the error, since this type of estimate is no longer optimum. In any case, the method of Sect. 5 is the practical one. Those who are interested in these experiments can turn immediately to Sect. 5.

Let g be a function defined on the (x,y) plane and vanishing outside a circle of radius r, with center at the origin; r is the radius of the Moon; g is the brightness map. Let R be the two dimensional Fourier transform of g in a (u,v) cartesian plane. In the (x,y) plane, the radius is λ say, angle ϕ; the transformed plane has radius τ, angle θ; $u = \tau\cos\theta$, $v = \tau\sin\theta$.

$$R(\tau,\theta) = \int_{\lambda=0}^{\tau} \int_{\theta=0}^{1} g(\lambda,\phi)\, e^{i\tau\lambda\cos 2\pi(\theta-\phi)} \lambda d\lambda d\phi\,,$$

in polar coordinates (λ,ϕ) and (ρ,ϕ).

Values of R at the points (τ_j, θ_k) are regarded as known exactly in the noiseless case. Let L be the "cutoff lag," K the number of spokes, i.e., $1/K$ the spacing between the θ angles. Then

$$\begin{cases} \tau_j = \dfrac{j}{T} L & j = 0, 1, \cdots, T\,; \\ \theta_k = \dfrac{k}{K} & k = 0, 1, \cdots, K-1\,. \end{cases}$$

The function g is to be estimated to minimize the mean square error between g and its estimate g.

Set

$$g_{(j-1)K+2+k}(\tau,\theta) = \begin{cases} e^{i\tau\tau_j \cos 2\pi(\theta-\theta_k)} & \text{if } \tau \le L \\ 0 & \text{if } \tau > L\,, \end{cases}$$

and let P_ℓ be the point with coordinates (τ_j,θ_k), where $(j-1)K + 2 + k = \ell$.

For arbitrary h_1, h_2, we will write (h_1, h_2) for the inner product of h_1 and h_2 and $G(h_1, \cdots, h_n)$ for the Grammian of $h_1, \cdots, h_n$, i.e.,

$$(h_1,h_2) = \int_0^\infty \int_0^1 h_1(\rho,\theta)h_2(\rho,\theta)\tau d\tau d\theta$$

and

$$G(h_1, \cdots, h_n) = \begin{vmatrix} (h_1,h_1) & (h_2,h_1) \cdots (h_n,h_1) \\ (h_1,h_2) & (h_2,h_2) \cdots (h_n,h_1) \\ \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \quad \cdot \\ (h_1,h_n) & (h_2,h_n) \cdots (h_n,h_n) \end{vmatrix}$$

The functions $e_1, \cdots, e_{TK+1}$ given by the determinant (the last row having entries which are functions, not numbers)

$$e_l(\rho,\theta) = \frac{1}{\sqrt{G(g_1, \cdots, g_l)}\ \sqrt{G(g_1, \cdots, g_{l-1})}} \begin{vmatrix} (g_1,g_1)(g_2,g_1) & \cdots (g_l,g_1) \\ (g_1,g_2)(g_2,g_2) & \cdots,(g_l,g_2) \\ \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \\ (g_1,g_{l-1})(g_2,g_{l-1}) & \cdots (g_l,g_{l-1}) \\ g_1 \quad g_2 & \cdots \quad g_l \end{vmatrix} \tag{1}$$

form an orthonormal set spanning the same subspace as the set $(g_1, \cdots, g_{TK+1})$ (Ref. 11, p. 24). Hence, an estimate $\hat{g}$ for g is given by

$$\hat{g} = \sum_{l=1}^{TK+1} \beta_l e_l + X, \tag{2}$$

where X is orthogonal to all the e_l and

$$\beta_l = (g, e_l) =$$

$$= \frac{1}{\sqrt{G(g_1,\cdots,g_l)}\ \sqrt{G(g_1,\cdots,g_{l-1})}} \begin{vmatrix} (g_1,g_1) & \cdots & (g_l,g_1) \\ \cdot \quad \cdot & \cdot & \cdot \quad \cdot \\ (g_1,g_{l-1}) & \cdots & (g_l,g_{l-1}) \\ (g,g_1) & \cdots & (g,g_l) \end{vmatrix}$$

$$= \frac{1}{\sqrt{G(g_1,\cdots,g_l)}\ \sqrt{G(g_1,\cdots,g_{l-1})}} \begin{vmatrix} (g_1,g_1) & \cdots & (g_l,g_1) \\ \cdot \quad \cdot & \cdot \quad 1 & \cdot \\ (g_1,g_{l-1}) & \cdots & (g_l,g_{l-1}) \\ R(P_1) & \cdots & R(P_l) \end{vmatrix} \tag{3}$$

Eqs. (1)–(3) express g as a linear combination of the g_m with coefficients which are linear combinations of the known values of R plus the error term X. The

L_2-magnitude of X, $\| X \|_2^2$, is given by (Ref. 11, p. 25),

$$\int |X|^2 = \frac{G(g, g_1, g_2, \cdots, g_{TK+1})}{G(g_1, g_2, \cdots, g_{TK+1})} \tag{4}$$

$$= \frac{\begin{vmatrix} \| g \|_2^2 & F(P_1) & F(P_2) & \cdots & F(P_{TK+1}) \\ F(P_1) & (g_1,g_1) & (g_2,g_1) & \cdots & (g_{TK+1},g_1) \\ F(P_2) & (g_1,g_2) & (g_2,g_2) & \cdots & (g_{TK+1},g_2) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ F(P_{TK+1}) & (g_1,g_{TK+1}) & (g_2,g_{TK+1}) & \cdots & (g_{TK+1},g_{TK+1}) \end{vmatrix}}{G(g_1, g_2, \cdots, g_{TK+1})}.$$

The only unknown in Eq. (4) is $\| g \|_2^2 = \| R \|_2^2$ which has to be estimated separately as the sum of the received R^2. It is easily shown, using standard determinant manipulations, that the error caused by replacing $\| R \|_2^2$ by $\| R \|_2^2 + \epsilon$ is exactly ϵ, so that estimating the sum of the $R^2(\tau_j, \theta_k)$ is fortunately not so critical.

4. L_2 Estimate with Noise

If we introduce noise into the measurements by replacing $R(P_j)$ with $\hat{R}(P_j) = R(P_j) + n_j$, where the n_j are random variables, but still use the same computational procedures as in Sect. 3, then the β_j are replaced by say $\beta_j + \gamma_j$. A simple linearity argument, then, shows that

$$\gamma_j = \frac{1}{\sqrt{G(g_1,\cdots,g_j)}\sqrt{G(g_1,\cdots,g_{j-1})}} \begin{vmatrix} (g_1,g_1) & \cdots & (g_j,g_1) \\ \cdot & \cdot & \cdot \\ (g_1,g_{j-1}) & \cdots & (g_j,g_{j-1}) \\ n_1 & \cdots & n_j \end{vmatrix},$$

and we write

$$\hat{g} = \sum_{j=1}^{TK+1} (\beta_j + \gamma_j) e_j = \sum_{j=1}^{TK+1} \gamma_j e_j + X.$$

Thus, the error is the sum of the last two terms above. However, in the presence of noise, there is a better procedure than imitating the optimum procedure in the noise free case.

If the n_j are Gaussian so are the γ_j; furthermore, if the n_j are orthogonal with mean zero and variance σ^2 then

$$E\left(\int_0^L \int_0^1 \left(\sum_{j=1}^{TK+1} \gamma_j e_j \right)^2 \tau d\tau d\theta \right) = \sum_{j=1}^{TK+1} E(|\gamma_j|^2)$$

$$= \sigma^2 \sum_{j=1}^{TK+1} \frac{1}{G(g_1,\cdots,g_j) G(g_1,\cdots,g_{j-1})} \sum_{l=1}^{j} \hat{G}_{jl},$$

where

$$\hat{G}_{jl} = \begin{vmatrix} (g_1,g_1) & \cdots (g_{l-1},g_1) & (g_{l+1},g_1) & \cdots (g_j,g_1) \\ (g_1,g_2) & \cdots (g_{l-1},g_2) & (g_{l+1},g_2) & \cdots (g_j,g_2) \\ \cdot & \cdot & \cdot & \cdot \\ (g_1,g_{j-1}) & \cdots (g_{l-1},g_{j-1}) & (g_{l+1},g_{j-1}) & \cdots (g_j,g_{j-1}) \end{vmatrix}.$$

The assumption of Gaussian noise in the receiver yields identically distributed, independent, hence uncorrelated, Gaussian errors n_j, so that the above formula is indeed appropriate. Since $E(\gamma_j) = 0$, all j and (because the γ_j are linear combinations of the n_i which have mean 0), the cross terms $X\ \gamma_j$ have expected value zero. Thus the expected square error in g is a sum of two terms, one the sum $\Sigma E|\gamma_j|^2$ and the other $||X||_2^2$.

5. The Practical Approach

A different and practical approximation yielding a different kind of estimate of error can be obtained through the use of Bernstein polynomials. The nth Bernstein polynomial approximation to a function h defined on (0,1) is defined as

$$B_n(x) = \sum_{\nu=0}^{n} \binom{n}{\nu} h\left(\frac{\nu}{n}\right) x^{\nu}(1-x)^{n-\nu}.$$

The basic approximation theorem is

$$|h(x) - B_n(x)| \leq 5/4 \max_x |h'(x)|/\sqrt{n}$$

(Ref. 12, Theorem 1.6.1, p. 20).

We will first approximate $R(\tau_j,\theta)$ as a function of θ on the circle of radius τ_j, using the values $R(\tau_j, \theta_k)$, $k = 0, 1, \cdots, K-1$. We have

$$|R(\tau_j,\theta) - B_k(\tau_j,\theta)| \leq \frac{5}{4} \max_{\theta} \left| \frac{\partial}{\partial\theta} R(\tau_j,\theta) \right| / \sqrt{K}$$

$$= \frac{5}{4} A_j/\sqrt{K} \text{ say for } j \geq 1;\ A_0 \text{ can be defined as } 0.$$

Let

$$\tilde{R}(\tau,\theta) = \begin{cases} R(0,\theta) = F(0) & \text{if } 0 \leq \tau \leq L/2T\ ; \\ B_K(\tau_j,\theta) & \text{if } \frac{j-\frac{1}{2}}{T} L < \tau < \frac{j+\frac{1}{2}}{T} L\ ; \\ 0 & \text{if } \left(1 + \frac{1}{2T}\right) L \leq \tau\ , \end{cases}$$

and let $\tilde{g}$ be the inverse Fourier transform of $\tilde{R}$. We will estimate $||R - \tilde{R}||_1$ which, as is seen by reference to the definition of Fourier transform, dominates $||g-\tilde{g}||_\infty$, the supremum of $|g-\tilde{g}|$ over the entire lunar disk.

For $|\tau-\tau_j| \leq L/2T$ (half a spoke subdivision), we have

$$|R(\tau,\theta) - \tilde{R}(\tau,\theta)| \leq |R(\tau,\theta| - |R(\tau_j,\theta)| + |R(\tau_j,\theta) - \tilde{R}(\tau_j,\theta)|$$

$$\leq |\tau - \tau_j| C_j + 5/4\ A_j/\sqrt{K}$$

where for each j, $0 \leq j \leq T$,

$$C_j = \max_{\substack{\tau-\tau_j \leq L/2T \\ 0 \leq \theta \leq 1}} \left| \frac{\partial R}{\partial \tau}(\tau,\theta) \right|.$$

thus

$$||R-\tilde{R}||_1 = \int_0^1 d\theta \left\{ \int_0^{L/2T} + \sum_{j=1}^{T} \int_{\frac{j-\frac{1}{2}}{T}\,L}^{\frac{j+\frac{1}{2}}{T}\,L} + \int_{\frac{T+\frac{1}{2}}{T}\,L}^{\infty} \right\} |R-R|\tau d\tau$$

$$\leq \int_0^1 d\theta \left\{ \int_0^{L/2T} C_0\tau^2 + \sum_{j=1}^{T} \int_{\frac{j-\frac{1}{2}}{T}\,L}^{\frac{j+\frac{1}{2}}{T}\,L} \left(|\tau-\tau_j|C_j + \frac{5}{4} A_j/\sqrt{K} \right) \tau d\tau \right\}$$

$$+ X',$$

where

$$X' = \int_0^{\theta} \int_{\frac{T+\frac{1}{2}}{T}\,L}^{1} |R|\tau d\tau d\theta .$$

This gives

$$||R-R||_1 \leq \left(\frac{L}{T}\right)^3 \left\{ \frac{C_0}{24} + \sum_{j=1}^{T} \frac{1}{4} jC_j \right\} + \frac{5}{4} \frac{1}{\sqrt{K}} \left(\frac{L}{T}\right)^2 \sum_{j=1}^{T} jA_j + X'. \tag{5}$$

Now

$$\frac{\partial R(\tau,\theta)}{\partial \tau} = \frac{\partial}{\partial \tau} \int_0^1 d\phi \int_0^r g(\lambda,\phi) \left(e^{i\lambda\tau \cos 2\pi(\theta-\phi)}\right) \lambda d\lambda$$

$$= \int_0^1 d\phi \int_0^r \frac{\lambda}{\tau} g(\lambda,\phi) \frac{\partial}{\partial \lambda} \left(e^{i\lambda\tau \cos 2\pi(\theta-\phi)}\right) \lambda d\lambda,$$

by simple algebra. Integrate by parts to obtain

$$\frac{\partial R(\tau,\theta)}{\partial \tau} = \frac{1}{\tau} \int_0^1 d\phi \left\{ e^{i\tau \cos 2\pi(\theta-\phi)}\, r^2 g(r,\phi) - \int_0^N e^{i\lambda\tau \cos 2\pi(\theta-\phi)}\, r \left[2\lambda g(\lambda,\phi) + \lambda^2 \frac{\partial}{\partial \lambda} g(\lambda,\phi) \right] d\lambda \right\}.$$

We finally obtain the desired estimate for c_j; $\delta > 0$

$$C_j \leq \frac{T}{(j-\frac{1}{2})L} \int_0^1 d\phi \left\{ r^2 |g(\lambda,\phi)| + \int_0^r \left[2\lambda |g(\lambda,\phi)| + \lambda^2 \left| \frac{\partial}{\partial \lambda} g(\lambda,\phi) \right| \right] d\lambda \right\}$$

$= D \frac{T}{L} \frac{1}{j}$, where D is a constant assumed from some rough *a priori* knowledge of g and of $\frac{\partial g}{\partial \lambda}$.

Thus, setting $A = \max_j A_j$, which again depends upon *a priori* knowledge of R or rather of $\frac{\partial R}{\partial \theta}$, one has

$$||g-\tilde{g}||_\infty \leq ||R-R||_1 \leq \left(\frac{L}{T}\right)^3 \left(\frac{C_0}{24} + \frac{1}{4} \frac{DT^2}{L} \right) + \frac{5}{4} \frac{1}{\sqrt{K}} L^2 A \left(\frac{T+1}{2T} \right) + X'.$$

Qualitatively speaking, then, there exist constants E_1 and E_2 which are to be estimated roughly *a priori* such that

$$||g-g||_\infty \leq E_1\frac{L^2}{T} + E_2\frac{L^2}{\sqrt{K}} + \int_0^1\int_L^\infty |R|\,\tau d\tau d\theta, \quad (6)$$

for large T.

The estimate for the supremum error in the estimate $\tilde{g}$ of g just obtained involves parameters T, L, K. We have assumed not only that the apparent rate of rotation is a constant independent of the apparent rotation axis but also that the apparent rotation axis itself rotates uniformly. Neither assumption is true. Failure of the first assumption gives different spacings of the τ_j on different spokes; failure of the second assumption results in different angular spacings between adjacent spokes. Nevertheless, let us speak as if both of the assumptions held.

The parameter L is the cutoff lag, at the experimenter's disposal. The parameter T is the number of lags in the correlation function at which data are taken on each spoke and are fixed by the equipment. (T is typically between 50 and 450.) Finally, K is the number of spokes used to cover one complete revolution. That is, K is the number of integrations of signal per day, where each spectrum is the result of a single integration. The experimenter can choose K arbitrarily large up to the maximum switch rate of the monostatic radar. Since the apparent axis is rotating, the integration time should be short enough so that resolution is not lost by averaging over arcs on the circles at a fixed radius from the sub-Earth point. On the other hand, a law of diminishing returns sets in if the integration time is too short. For then each measurement becomes noisier, and the effect of our numerical procedure would be to average adjacent measurements. Thus, there would be essentially no gain with too large K; furthermore, there would be more complicated data handling.

The experimenter must decide which T,L to choose. One consideration, however, is the following. For a fixed number of lags in $R(\theta,\tau)$, it may be better to decrease L, that is to "throw away" data far out on the spokes. Examining formula (6), we find that the sum $E_1\ L^2/T + E_2\ L^2/\sqrt{K}$ increases with L, whereas $\int_{\tau=L}^\infty |R|\,\tau d\tau$ decreases with L. Hence, there is an optimum L with fixed T; T may as well be as large as the available correlator allows. The fact that a term $E_2\ L^2/\sqrt{K}$ occurs does not imply that K should be taken large, since the error due to the noise increases linearly with K.

Noise is taken into account rather easily, since if R is perturbed by noise so as to increase the L_2-error in R by η, then the L_1-error in R, and hence the supremum error in $\tilde{g}$, is increased by at most η. When this error is taken into account, the experimenter can then choose his L and K so as to minimize the total expected supremum error in $\tilde{g}$. The procedure would evolve as the experiment evolves.

To be specific, consider the equation for $\Delta_{j,\theta}$ where we define

$$\Delta_{j,\theta} = |\,R(\tau_j,\theta) - B_k(\tau_j,\theta)\,|\,. \quad (7)$$

If $R(\tau_j;\theta)$ is wrong by noise $n_{j,\theta}$, then $E(\Delta_{j,\theta},)$ increases by at most σ, where σ is the standard deviation of $n_{j,\theta}$, assumed independent of j,θ. (For the origin $j=0$, the noise variance is less, since every spoke contributes. The variance for that point is actually σ^2/K.)

Then $E(||\,R-\tilde{R}\,||_1)$ increases by at most σ times the area of the circle of radius L, which is L^2 in these units. Consequently,

$$E(||\,g-\tilde{g}\,||_\infty) \leq E_1\frac{L^2}{T} + E_2\frac{L^2}{\sqrt{K}} + \int_{\theta=0}^1\int_{\rho=L}^\infty |\,R\,|\tau d\tau d\theta + \sigma L^2 \quad (8)$$

If the σ-term is standardized so that σ^2 represents the variance of the noise for $K=1$, then the variance for K spokes is $\sigma^2 K$. Finally,

$$E(||\,g-\tilde{g}\,||_\infty) \leq \sigma L^2\sqrt{K} + E_1\frac{L^2}{T} + E_2\frac{L^2}{\sqrt{K}} + \int_{\theta=1}^1\int_{\tau=L}^\infty |\,R\,|\,\tau d\tau d\theta\,. \quad (9)$$

Eq. (9) is the actual equation the experimenter would use to minimize the expected supremum error in the estimate $\tilde{g}$ of the brightness map g. He would try to choose the cutoff lag L and integration time $1/K$ so as to minimize $E(||\,g-\tilde{g}\,||_\infty)$, subject to the equipment being buildable with these parameters. Having chosen T, choose L and K to minimize Eq. (9).

For example, suppose that the actual $R(\tau,\theta)$ obeys $R(\tau,\theta)\sim e^{-\beta\tau^2}$ for τ large, which is implied by similar estimates for $g(x,y)$. Then

$$\int_{\theta=0}^1\int_{\tau=L}^\infty |\,R\,|\,\tau d\tau d\theta \sim \frac{1}{2\beta}\,e^{-\beta L^2}. \quad (10)$$

Then, if $E_0 = E_1/T$, then

$$E(||\,g-\tilde{g}\,||_\infty) \leq \sigma L^2\sqrt{K} + E_1\frac{L^2}{\sqrt{K}} + E_0L^2 + \frac{1}{2\beta}\,e^{-\beta L^2}. \quad (11)$$

We shall minimize the right hand side of Eq. (11) by the proper choice of L and K. Call the right hand side of Eq. (11) $f(L,K)$. Then

$$\begin{cases} \dfrac{\partial f}{\partial K} = \dfrac{1}{2\sqrt{K}}\sigma L^2 - \dfrac{E_1 L^2}{2K\sqrt{K}}\,; \\ \dfrac{\partial f}{\partial L} = 2\sigma L\sqrt{K} - 2E_1\dfrac{L}{\sqrt{K}} + 2E_0 L - Le^{-\beta L^2}\,. \end{cases} \tag{12}$$

Setting these derivative equal to 0 and solving, we find for the optimum value K_0

$$K_0 = \frac{E_1}{\sigma}\,. \tag{13}$$

Then, ignoring the trivial root

$$L_0 = 0 \text{ of } \frac{\partial f}{\partial L}\Big)_{L_0,K_0} = 0,$$

we have

$$L_0 = \sqrt{\frac{1}{\beta}\log\frac{1}{2E_0}}\,, \tag{14}$$

since the first two terms in

$$\frac{\partial f}{\partial L}\Big)_{L_0,K_0} = 0$$

cancel, identically in L_0, at $K = K_0$.

We see that K_0 and L_0 are determined independently, i.e., by disjoint sets of parameters. The fact that L_0 is finite is the "data rejection" phenomena mentioned before. The fact that K_0 is finite means that it pays to do some averaging over areas before fitting Bernstein polynomials. Both of these facts are reasonable to expect in advance.

The main difficulty with applying this method would be to find E's by some *a priori* choices. These estimates for the E's would presumably change as the experiment progresses, and more knowledge of g and R is obtained.

References

1. Davenport, W. B., Jr., "Signal-to-Noise Ratios in Band-Pass Limiters," *J. Appl. Phys.*, Vol. 29, No. 6, pp. 720–727, June 1955.
2. Davenport, W. B., Jr., and Root, W. L., *Random Signals and Noise*, McGraw-Hill Book Co., New York, pp. 158–159, 1958.
3. Posner, E. C., "The Use of Quantiles for Space Telemetry Data Compression," Proceedings of the National Telemetery Conference, Sect. 1–3, 1964.
4. Tveitan, W., private communication.
5. Gilbert, E. N., "Synchronization of Binary Messages," *IRE Transactions on Information Theory*, Vol. IT-6, pp. 470–477, 1960.
6. Kraft, L. G., *A Device for Quantizing, Grouping, and Coding Amplitude Modulated Pulses*, Masters Thesis, Electrical Engineering Department, MIT, March, 1949.
7. Fano, R. M., *Transmission of Information*, John Wiley and Sons, Inc., New York, 1961.
8. Thompson, J., private communication to R. Goldstein, Jet Propulsion Laboratory, Pasadena, 1965.
9. Ryle, M., and Hewish, A., "The Synthesis of Large Radio Telescopes," *Monthly Notices Roy. Astronom. Soc.*, Vol. 120, pp. 220–230, 1960.

References (Cont'd)

10. Victor, W. K., Goldstein, R. M., et. al., *Radar Exploration of Venus,* Technical Report 32-396, Jet Propulsion Laboratory, Pasadena, 1965.

11. Achieser, N. I., *Theory of Approximation,* translated, Ungar, New York, 1959.

12. Lorentz, G. G., *Bernstein Polynomials,* Toronto University Press, 1963.

XXIII. Communications Systems Research: Data Compression

A. Probabilistic Metric Spaces of Infinite Entropy

E. C. Posner and H. C. Rumsey, Jr.

1. Summary

This article continues the study of probabilistic metric spaces (PMS), a study begun in *SPS 37-34*, Vol. IV, p. 292 in order to make precise the notion of data compression. The question asked is: Which PMS have bounded ϵ; δ entropy as $\epsilon \to 0$, with δ fixed? The answer is that such PMS are the disjoint union of two spaces, one having probability δ, and the other a countable space having finite (Shannon) entropy. As a result of this theorem, we can conclude that all the spaces arising in data compression studies have unbounded $\epsilon;\delta$ entropies as $\epsilon \to 0$ and δ is fixed (and < 1).

2. Introduction

The notion of PMS was introduced (*SPS 37-34*, Vol. IV, p. 292) in order to have a precise framework for discussing and comparing data compression systems and, at the same time, to give a mechanism for generating data compression systems that achieve an efficiency near the theoretically optimum limit.

Basically, a PMS is a complete metric space (X,d), together with a probability measure μ on the Borel sets of X. (The Borel sets are the sets of the σ-algebra generated by the open sets.) This measure μ is required to have the key property that the probability of every nonempty open set is positive. In the sense of the needs of the ultimate data user, the metric d represents a "fidelity criterion": the distance between two points of X tells the user how good a representation of the first point the second point is. These points of X are outcomes of an experiment distributed according to probability μ. In one of the main applications, for example, the points of X are sample functions of a stochastic process defined on the unit interval.

A PMS is said to be probabilistically compact (PC) if, for every $\delta > 0$ and every open covering U of X, there is a finite subset V of U that covers all of X except possibly a set of probability δ or less. We proved in *SPS 37-34*, Vol. IV, p. 292 that PC is equivalent to separability of the underlying metric space.

Partitions U of X by disjoint open sets that miss at most δ in probability and in which the diameter of every open set is at most ϵ are called $\epsilon;\delta$ partitions. When X is separable, U has a countable subpartition which covers every point that U covers; hence we may as well assume

U itself is countable. We associate with each U a discrete probability space whose ith probability is the conditional probability that a point be in the ith set of the partition, given that the point is covered by U. This discrete probability space has an associated entropy (possibly infinite). The infimum $H_{\epsilon;\delta}(X)$ of all entropies for $\epsilon;\delta$ partitions of X is called the $\epsilon;\delta$ entropy of X. It was proved in *SPS 37-34*, Vol. IV, p. 292 that $H_{\epsilon;\delta}(X)$ is finite if X is separable and $\epsilon;\delta$ are both positive.

We were led to the definition of $H_{\epsilon;\delta}(X)$ by the following considerations: The maximum diameter ϵ represents the precision with which experimental outcomes must be known, in a practical sense, by someone not witnessing the actual experiment. Thus, knowing merely the index of that open set containing the true outcome x one knows the outcome to within ϵ.

With probability at most δ, however, x is not covered by the open set whose index is given so that the data transmission scheme fails with probability at most δ. The average number of bits required to specify outcomes within ϵ and with probability at least 1-δ is, thus, at least $H_{\epsilon;\delta}(X)$, and can be made arbitrarily close to $H_{\epsilon;\delta}(X)$ by suitable choice of an $\epsilon;\delta$ partition.

This note is devoted to the study of the behaviour of $H_{\epsilon;\delta}(X)$ as $\epsilon \to 0$ with δ fixed. A particular consequence is that if X is nonatomic (has no points of positive probability), then for every $\delta < 1$, $H_{\epsilon;\delta}(X)$ is unbounded as $\epsilon \to 0$. For the data sources occurring in space experiments, the conclusion is that arbitrarily many bits are needed to transmit experimental outcomes to within arbitrarily high precision—certainly not a surprising result but one which, nevertheless, needs rigorous mathematical verification.

3. Statement of Theorem

The precise result to be proved is the following two part theorem, whose converse is trivial.

Theorem: Let X be a separable PMS, and let $0 \leq \delta < 1$ be given. Then

1. If $H_{\epsilon;\delta}(X)$ remains bounded as $\epsilon \to 0$, X can be written as $X = Y \cup Z$, where Y and Z are disjoint Borel subsets of X with $\mu(Y) \leq \delta$, and where Z is the union of all the atoms of X. Furthermore, if $H(Z)$ is the (conditional) entropy of the discrete probability space Z, then $H(Z) \geq \lim_{\epsilon \to 0} H_{\epsilon;\delta}(X) \triangleq H_\delta(X)$. (Note that $H_\delta(X)$ is allowed to be infinite.)

2. If $H_{\epsilon;\delta}(X)$ remains bounded as $\epsilon \to 0$ for all $\delta > \delta_0$ but not necessarily for $\delta = \delta_0$, and is definitely unbounded for every $\delta < \delta_0$, then $X = Y \cup Z$, where Y and Z are disjoint Borel subsets of X, $\mu(Y) = \delta_0$, Z is the set of atoms of X, and $H(Z) = H_{\delta_0}(X)$.

4. Proof of Part 1

A first observation: $H_{\epsilon;\delta}(X)$ is nondecreasing for fixed δ as ϵ decreases, since the definition of an $\epsilon;\delta$ partition allows diameters equal to or less than ϵ. Thus, if $\epsilon_1 \leq \epsilon_2$, every $\epsilon_1;\delta$ partition is an $\epsilon_2;\delta$ partition. Consequently, the infimum over the class of $\epsilon_1;\delta$ partitions is equal to or greater than the infimum over the class of $\epsilon_2;\delta$ partitions. This proves that $\lim_{\epsilon \to 0} H_{\epsilon;\delta}(X)$ always exists (infinite limits allowed).

Let U be an $\epsilon;\delta$ partition of X with open sets $U_i, i = 1,2,\ldots$; let $p_i = \mu(U_i)/(\sum_j \mu(U_j))$, $i = 1, 2, \cdots$, where the indexing is arranged such that $p_1 \geq p_2 \geq \ldots$. Then the entropy $H(U)$ is defined as $H(U) = -\sum_i p_i \log_2 p_i$. Since $H_{\epsilon;\delta}(X)$ is finite, $H(U)$ is finite for certain U. Let us consider one of those.

First we note that $p_1 \geq 2^{-H(U)}$. This crucial inequality follows from the following considerations

$$2^{-H(U)} = \prod_i (p_i^{p_i}) \leq \prod_i (p_1^{p_i}) \text{ (since } p_1 \geq p_i, \text{ all } i)$$

$$\leq p_1^{\sum_i p_i} = p_1 .$$

Since $H(U)$ can be made arbitrarily close to $H_{\epsilon;\delta}(X)$ by the very definition of the latter quantity, for certain U we have $H(U) < (3/2)H_{\epsilon;\delta}(X)$, and we need consider only such U. Furthermore, for ϵ sufficiently small, $H_{\epsilon;\delta}(X)$ is close to $H_\delta(X)$ so that $H(U) < 2H(X)$ for ϵ small.

We find that for the $\epsilon;\delta$ partitions U under consideration, the largest probability p_1 satisfies $p_1 > 2^{-2H\delta}$. Since $\mu(U_1) = p_1 \sum_j \mu(U_j)$ and since $\sum_j \mu(U_j) \geq 1 - \delta$, we see that $\mu(U_1) > (1 - \delta)2^{-2H\delta}$. Since, by assumption, H_δ is finite, we conclude:

There exists a constant $B > 0$ such that for arbitrarily small $\epsilon > 0$, there is an open set $U_1 = U_1(\epsilon)$ such that $\mu(U_1) \geq B$ and such that U_1 has diameter at most ϵ.

This last conclusion enables atoms to be isolated in X. We recall the definition of the limit superior (lim sup) of a family of sets $\{V_n\}$

$$V = \limsup V_n = \bigcap_{n>1} \bigcup_{k>n} V_k, \text{ a measurable set.}$$

That is, V is the set of points contained in infinitely many of the V_n. We shall use lim sups to construct atoms. The Arzela-Young Theorem (Ref. 1, p. 53, Theorem 2) states that if each $\mu(V_n) \geq B$, then $\mu(V) \geq B$.

We use the Arzela-Young Theorem as follows: Let V_1 be a set in a $1;\delta$ partition of X of largest probability among sets in the partition. Let V_2 be similarly defined, but for $1/2;\delta$ partitions. Similarly, for each n, let V_n be defined for $1/2^{n-1};\delta$ coverings. As we proved above, it follows that for some positive B, $\mu(V_n) \geq B$, for all n. Thus, if $V = \limsup V_n$, then $\mu(V) \geq B > 0$.

The above result on $\mu(V)$ is to be used to isolate an atom of X. Since $\mu(V) > 0$, V is nonempty. Let x be in V, and consider those V_n containing x. Call the union of these V_n the set W. The diameter of W is at most 2, since two points in W are at distance at most 1 from x. Let the sets from among the V_n in the natural order that contain x be denoted by $V_{(i)}$, $i = 1, 2, \ldots$. Define $W_1 = W = \bigcup_{i=1}^{\infty} V_{(i)}$, $W_2 = \bigcup_{i=2}^{\infty} V_{(i)}, \ldots, W_n = \bigcup_{i=n}^{\infty} V_{(i)}, \ldots$.

The sets W_n produce our atom. Each W_n has a diameter at most $1/2^{n-2}$ since each W_n still contains the point x, and diameter $V_{(i)} \leq 1/2^{i-1}$. Also, each W_n contains some V_j, and so $\mu(W_n) \geq B$.

Now we consider $\bigcap_{n=1}^{\infty} W_n$. This set has diameter 0, since diameter $W_n \to 0$ as $n \to \infty$. Hence, $\bigcap_{n=1}^{\infty} W_n = \{x\}$. Also, $\mu\left(\bigcap_{n=1}^{\infty} W_n\right) \geq$ B, since $\mu(W_n) \geq B$ for all n. Consequently, $\mu(\{x\}) \geq B$, so that x is an atom.

To sum up what we have just proved: If a value of δ, $0 \leq \delta < 1$, is given, and $H_{\epsilon;\delta}(X)$ remains bounded as $\epsilon \to 0$, then X contains an atom. A transfinite induction completes the proof of Part 1, as follows:

Let the atoms of X be x_i, $i \geq 1$. Let $\sum_i \mu(\{x_i\}) = M$. We observe that for every $\eta > 0$, open sets U_i can be put around each x_i such that $\mu(U_i) < \mu(\{x_i\}) + \eta/2^i$. We conclude that $\sum_i \mu(U_i) < M + \eta$.

Suppose it were true that $M < 1 - \delta$; we could then choose $\eta > 0$ so that $M + \eta < 1 - \delta$. From the U_i as above, let $U = \bigcup_i U_i$, and define $X' = X - U$. Define a probability measure μ' on X' as follows: if $A \subset X'$; $\mu'(A) = \mu(A)/\{1-\mu(U)\}$. Then it could readily be verified that X' is a separable PMS under the same metric d as before, but with new probability measure μ', that is, such that $\mu'(A) > \mu(A)/(1 - (M + \eta))$. Consequently, the PMS X' satisfies the hypothesis of Part 1 of the theorem, but with a δ' equal to $\delta/(1-(M+\eta)) < \delta/\delta = 1$.

Since the hypothesis of Part 1 holds for X', we conclude that X' contains an atom y. But y is also an atom of X, hence y is in U. Therefore, y is not in X', and we have produced a contradiction. The previous assumption that $M < 1 - \delta$ must be retracted, so, in fact, $M \geq 1 - \delta$.

Now define $Z = \bigcup_i \{x_i\}$. We have proved that $\mu(Z) \geq 1 - \delta$, so it follows that if $Y = X - Z$, then $\mu(Y) \leq \delta$.

To prove Part 1 of the theorem, it remains only to show that $H(Z) \geq H_{\delta_0}$. To do this, let ϵ be given such that $H_{\delta;\epsilon}(X)$ is close to $H_\delta(X)$ (i.e., is large if H_δ is infinite). As above, cover each x_i by U_i where $\mu(U_i) < \mu(\{x_0\})\ (1 + \eta)$, for a non-negative η. We can also require that diameter $U_i < \epsilon$, for all i. Such U_i may not be disjoint, but we can choose a subset of the x_i whose U_i are disjoint and such that all the x_j are in some U_i. We can furthermore require that the U_i corresponding to the first, say, N of the x_i are disjoint from all other U_j by taking ϵ sufficiently small, depending on N. Each of these partitions U chosen according to these criteria are $\epsilon;\delta$ partitions of X, since

$$\sum_i \mu(\{x_i\}) \geq 1 - \delta.$$

More importantly, $H(U)$ is as close to $H(Z)$ as we care to make it, provided only that η is chosen small and N large. This follows from direct substitution into the formula defining $H(U)$, with N chosen so large that the first N terms of the series for $H(Z)$ have a sum close to $H(Z)$ (a large sum suffices if $H(Z)$ is infinite). Then we may choose η small to reach the desired conclusion about $H(U)$.

On the other hand, since U is an $\epsilon;\delta$ partition, we must have $H(U) \geq H_{\epsilon;\delta}(X)$ for ϵ the largest diameter of any U_i. If $H(Z)$ is infinite, then of course $H(Z) \geq H_\delta$; but if $H(Z)$ is finite, we can obtain partitions U such that $H(U) \to H(Z)$. Furthermore, each $H(U)$ satisfies $H(U) \geq H_{\epsilon;\delta}(X)$ for a sequence of ϵ approaching 0. Since $H_{\epsilon;\delta}(X)$ then approaches $H_\delta(X)$ for the sequence of partitions U, we finally conclude that $H(Z) \geq H_\delta$ by taking limits on both sides of the inequality $H(U) \geq H_{\epsilon;\delta}(X)$. Part 1 of the theorem is proved at last.

5. Proof of Part 2

Observe that for fixed ϵ, $H_{\epsilon;\delta_1}(X) \geq H_{\epsilon;\delta_2}(X)$ if $\delta_1 \leq \delta_2$, since every $\epsilon;\delta_1$ covering is an $\epsilon;\delta_2$ covering. Hence, finite $H_{\delta_1}(X)$ implies finite $H_{\delta_2}(X)$ whenever $\delta_1 \leq \delta_2$. Consequently, the limit inferior (lim inf) of all δ such that $H_\delta(X)$ is finite exists, and this lim inf is precisely δ_0.

Let $\delta > \delta_0$, so that $H_\delta(X)$ is finite. As Part 1 applies, the union Z of the atoms of X has probability at least $1 - \delta$. Since δ is any number greater than δ_0, we have proved that $\mu(Z) \geq 1 - \delta_0$. But just as in the final argument of Part 1, we can contradict $\mu(Z) > 1 - \delta_0$. Thus, let $Z_1 \subseteq Z$ with $H(Z_1)$ finite. Also let us require $\mu(Z_1) > 1 - \delta_0$. The set of atoms Z_1 can be covered by an $\epsilon;\delta_1$ partition for arbitrarily small ϵ, for some $\delta_1 \leq 1 - \mu(Z_1)$. But then $\delta_1 < \delta_0$. On the other hand, as in the end of the proof of part 1, the entropy of these $\epsilon;\delta_1$ partitions can be made arbitrarily close to $H(Z_1)$. Hence $H_\delta(X)$ must be finite, a fact which contradicts the definition of δ_0. The only remaining alternative must, in fact be the case: $\mu(Z) = 1 - \delta_0$.

The final point to be settled is that $H(Z) = H_{\delta_0}(X)$. From Part 1, we know that $H(Z) \geq H_\delta(X)$, all $\delta > \delta_0$. We can prove (although we shall not do so here) that $H_{\delta_0}(X) = \limsup_{\delta > \delta_0} H_\delta(X)$ for any δ_0 (a right continuity property). This right continuity enables us to conclude that $H(Z) \geq H_{\delta_0}(X)$. It remains to prove that $H_{\delta_0} \geq H(Z)$.

First observe that if $\delta = \delta_0$, if U is an $\epsilon;\delta$, partition, and if W is the set covered by U, then for ϵ sufficiently small, $\mu(U \cap Y)$ is arbitrarily small. Otherwise, as in Part 1, Y would contain an atom of X, contradicting the fact that Y is the complement of the set Z of atoms of X. Due essentially to the right continuity of $H_\delta(X)$ invoked above, we can conclude also for $\epsilon;\delta_0$ partitions that $\mu(U \cap Y)$ is arbitrarily small as $\epsilon \to 0$.

A slight modification of the argument used to complete the proof of Part 1 can now be used to show that for ϵ sufficiently small, any $\epsilon;\delta_0$ partition U has entropy $H(U)$ such that $H(U) \geq C(\epsilon)H(Z_\epsilon)$, where the constant $C(\epsilon)$ approaches 1 as $\epsilon \to 0$, and Z_ϵ is a subset of Z such that $H(Z_\epsilon) \sim H(Z)$. Hence, the infimum $H_{\epsilon;\delta_0}(X)$ of $H(U)$ for $\epsilon;\delta_0$ partitions U satisfies $H_{\epsilon;\delta_0}(x) \geq C(\epsilon)H(Z)$, if $H(Z)$ is infinite.

If $H(Z)$ is infinite, then we conclude that $H(U)$ must get arbitrarily large for any $\epsilon;\delta_0$ partition as $\epsilon \to 0$. Hence the infimum $H_{\epsilon;\delta_0}(X)$ of these $H(U)$ also gets arbitrarily large as $\epsilon \to 0$. Consequently, $H_{\delta_0}(X)$ is infinite, and the theorem is proved in case $H(Z)$ is infinite.

If $H(Z)$ is finite, then, by taking the lim sup as $\epsilon \to 0$ of both sides of the inequality $H_{\epsilon;\delta_0}(X) \geq C(\epsilon)H(Z)$, we finally conclude that $H_{\delta_0}(X) \geq H(Z)$. Since we already know that $H(Z) \geq H_{\delta_0}(X)$, Part 2 is proved, and the proof of the theorem is now complete.

B. Detection of Mixtures of Data Sources

I. Eisenberger

1. Summary of Results

In this article, we consider the distribution of a mixture of two normal distributions and investigate some conditions for which the distribution is parapolymodal (unimodal with more than two points of inflection).

If a random variable is such that, with probability p, it comes from a normal population with density $N(\mu_1, \sigma_1)$ and, with probability $1 - p$ it comes from a normal population with density $N(\mu_2, \sigma_2)$, its density function is given by

$$f(x;p) = pN(x; \mu_1,\sigma_1) + (1-p)N(x; \mu_2,\sigma_2)$$

The following results are derived.

(1) If $\mu_1 = \mu_2$, a necessary and sufficient condition that there exist values of p, $0 < p < 1$, for which $f(x;p)$ is parapolymodal is that

$$\sigma_1^2 \sigma_2^2 < (\sigma_2^2 - \sigma_1^2)^2/8$$

alternatively,

$$\sigma_1/\sigma_2 < (3)^{1/2} - (2)^{1/2} \text{ or } \sigma_1/\sigma_2 > (3)^{1/2} + (2)^{1/2}$$

(2) If $\mu_1 \neq \mu_2$, but $f(x;p)$ is unimodal for all p, then the above condition is sufficient that there exist values of p for which $f(x;p)$ is parapolymodal.

(3) For fixed values of $\mu_1, \sigma_1, \mu_2, \sigma_2$, if there exist values of p such that $f(x;p)$ is bimodal, then there exist values of p such that $f(x)$ is parapolymodal.

(4) For every set of values $(\mu_1, \sigma_1, \mu_2, \sigma_2)$ values of p exist for which $f(x;p)$ is neither bimodal nor parapolymodal.

2. Introduction

The result of a typical statistical spacecraft experiment is subject to serious misinterpretation if the experimenter assumes that the samples were taken from a single population when, in fact, they were drawn from a mixture of two or more populations. For example, the distribution of particle energies when all the particles come from the same source is very likely to differ markedly from the distribution in effect when the particles are mixed with others from a different source. In this instance, since the distribution of particle energies can reasonably be expected to be unimodal when all the samples come from a single source, bimodality suggests strongly the presence of more than one kind of particle in the population. As a consequence, the modal properties of a mixture of two normal distributions were investigated (Ref. 2).

It should be noted, however, that it may not be necessary to demand bimodality before suspecting that a mixture of two populations is present. A milder condition, and a more fundamental one, involves the number of points of inflection that the density function possesses. A positive continuous density function, ranging from $-\infty$ to $+\infty$, has at least two points of inflection. If it has only two, the distribution is unimodal. However, since the converse is not true, it may be possible to detect a mixture of two distributions when the number of points of inflections exceeds two, even when the distribution is unimodal. Therefore, it becomes important to consider ways in which this type of distribution can arise. To simplify the discussion, if a distribution is unimodal but has four or more points of inflection, it will be called parapolymodal.

3. Theoretical Discussion

Since a point of inflection of $f(x;p)$ corresponds to an extreme value of $f'(x;p)$, the zeros of $f''(x;p)$ will be investigated. Putting $f''(x;p) = 0$, one obtains

$$f''(x;p) = \frac{p}{\sigma_1^5 (2\pi)^{1/2}} \left[(x-\mu_1)^2 - \sigma_1^2\right] \exp\left[\frac{-\frac{1}{2}(x-\mu_1)^2}{\sigma_1^2}\right] + \frac{(1-p)}{\sigma_2^5 (2\pi)^{1/2}} \left[(x-\mu_2)^2 - \sigma_2^2\right] \exp\left[\frac{-\frac{1}{2}(x-\mu_2)^2}{\sigma_2^2}\right] = 0 \tag{1}$$

Now let $\mu_1 = \mu_2 = \mu$. If $\sigma_1 = \sigma_2$, $f(x;p)$ is a normal density function for all p and is neither bimodal nor polyparamodal. Hence, we can assume without loss of generality that $\sigma_2 > \sigma_1$. Since $x = \mu \pm \sigma_1$ is not a root of $f''(x;p)$, Eq. (1) can be divided by the first term of $f''(x;p)$ and obtain, after rearranging,

$$g_1(y) = \frac{y^2-\sigma_2^2}{y^2-\sigma_1^2} h_1(y) = -\frac{\sigma_2^5 p}{\sigma_1^5(1-p)} \tag{2}$$

where $y = x - \mu$

and

$$h_1(y) = \exp\left[\tfrac{1}{2}y^2\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right)\right]$$

Since

$$-\frac{\sigma_2^5 p}{\sigma_1^5(1-p)} < 0$$

and this term takes on all finite negative values exactly once on the interval $0 < p < 1$ for all fixed values of σ_1 and σ_2, each value of y for which $g_1(y) < 0$ is a root of the equation

$$g_1(y) = -\frac{\sigma_2^5 p}{\sigma_1^5(1-p)}$$

for some unique p, and, hence, a root of $f''(x;p) = 0$ for exactly one value of p. For $y \gtrless \sigma_2$, $y \lessgtr -\sigma_2$ and $-\sigma_1 < y < \sigma_1$, $g_1(y) \gtrless 0$, one is interested only in values of y on the intervals $\sigma_1 < y < \sigma_2$ and $-\sigma_2 < y < -\sigma_1$. For $\sigma_1 < y < \sigma_2$, $g_1(y) < 0$, $g_1(y) \to -\infty$ as $y \to \sigma_1^+$ and $g_1(\sigma_2) = 0$. Therefore, since $g_1(y)$ is continuous on $\sigma_1 < y < \sigma_2$, $g_1(y)$ takes on all finite negative values at least once on this interval. Moreover, if $g_1(y)$ is monotone increasing on this interval, all negative values will be attained exactly once. Since there will then exist a one-to-one correspondence between the values of $g_1(y)$, for $\sigma_1 < y < \sigma_2$ and $-\sigma_2^5 p/\sigma_1^5 (1-p)$ for $0 < p < 1$, $f'(x;p)$ will, therefore, have a single extreme value for all p. Hence, $f(x;p)$ will have only one point of inflection on this interval.

Similarly, for $-\sigma_2 < y < -\sigma_1$, $g_1(y) < 0$, $g_1(y) \to -\infty$ as $y \to -\sigma_1^-$ and $g_1(-\sigma_2) = 0$. If $g_1(y)$ is monotone decreasing on this interval, $f(p)$ will have only one point of inflection on $-\sigma_2 < y < -\sigma_1$. Since increasing monotonicity is implied by $g_1'(y) > 0$ on $\sigma_1 < y < \sigma_2$, conditions satisfying this relation will now be investigated. Due to the symmetry of $g_1(y)$ about $y = 0$, the results obtained on $\sigma_1 < y < \sigma_2$ will also hold for $-\sigma_2 < y < -\sigma_1$.

For $\sigma_1 < y < \sigma_2$,

$$g'(y) = h_1(y)\left[\frac{2y(y^2-\sigma_1^2) - 2y(y^2-\sigma_2^2)}{(y^2-\sigma_1^2)^2} + \frac{y(y^2-\sigma_2^2)}{y^2-\sigma_1^2}\left(\frac{\sigma_2^2-\sigma_1^2}{\sigma_1^2\sigma_2^2}\right)\right]$$

$$= \frac{yh_1(y)(\sigma_2^2-\sigma_1^2)}{\sigma_1^2\sigma_2^2(y^2-\sigma_1^2)^2}\left[2\sigma_1^2\sigma_2^2 + (y^2-\sigma_2^2)(y^2-\sigma_1^2)\right]$$

$$\geqslant \frac{yh_1(y)(\sigma_2^2-\sigma_1^2)}{\sigma_1^2\sigma_2^2(y^2-\sigma_1^2)^2}\left[2\sigma_1^2\sigma_2^2 - \frac{(\sigma_2^2-\sigma_1^2)^2}{4}\right] \geqslant 0$$

$$\text{if } \sigma_1^2\sigma_2^2 \geqslant \frac{(\sigma_2^2-\sigma_1^2)^2}{8} \tag{3}$$

Thus, for values of σ_1, σ_2 satisfying inequality (3), $g_1(y)$ increases monotonically on $\sigma_1 < y < \sigma_2$ and for each value of p, $0<p<1$, there is only one value of y on the interval $\sigma_1 < y < \sigma_2$ and only one value of y on the interval $-\sigma_2 < y < -\sigma_1$, for which $f''(x;p) = 0$. Hence, $f(x;p)$ has only two points of inflection and is not parapolymodal.

However, for $y = \left(\frac{\sigma_1^2+\sigma_2^2}{2}\right)^{1/2}$

$$g_1'\left(\frac{\sigma_1^2+\sigma_2^2}{2}\right)^{1/2} = \frac{\left(\frac{\sigma_1^2+\sigma_2^2}{2}\right)^{1/2} h_1\left(\frac{(\sigma_1^2+\sigma_2^2)}{2}\right)^{1/2}(\sigma_1^2-\sigma_2^2)}{\sigma_1^2\sigma_2^2\left(\frac{\sigma_2^2-\sigma_1^2}{2}\right)^2}\left[2\sigma_1^2\sigma_2^2 - \frac{(\sigma_2^2-\sigma_1^2)}{4}\right] < 0 \tag{4}$$

if $$\sigma_1^2\sigma_2^2 < \frac{(\sigma_2^2-\sigma_1^2)^2}{8}$$

This result shows that for values of σ_1 and σ_2 satisfying inequality (4) there exists an interval about $y = [(\sigma_1^2+\sigma_2^2)/2]^{1/2}$ for which $g_1(y)$ is decreasing. For each y_0 in this interval there is a $y_1 < y_0$, a $y_2 > y_0$ and a value of p such that $g_1(y_1) = g_1(y_0) = g_1(y_2) = -\sigma_2^5 p/\sigma_1^5(1-p)$. Consequently, for this value of p, $f''(x_1;p) = f''(x_0;p) = f''(x_2;p) = 0$, where $x_i = y_i + \mu$, $i = 0, 1, 2$. Now suppose that for two of the values y_0, y_1, y_2, say y_1 and y_2, $f'(x_1;p)$ and $f'(x_2;p)$ are not extreme values of $f'(x;p)$. Then $f'''(x_1;p) = f'''(x_2;p) = 0$ and $f'''(x_0) > 0$.

However, from $f'''(x_1;p)$ one always has

$$1 + \frac{\sigma_1^7(1-p)(y_1^2-3\sigma_2^2)}{\sigma_2^7 p(y_1^2-3\sigma_1^2)}h_1(y_1) = 0, \text{ (assuming } y_1^2 \neq 3\sigma_1^2) \tag{5}$$

From $f''(x_1;p) = 0$, one also has

$$1 + \frac{\sigma_1^5(1-p)(y_1^2-\sigma_2^2)}{\sigma_2^5 p(y_1^2-\sigma_1^2)}h_1(y_1) = 0 \tag{6}$$

Equating Eqs. (5) and (6), one finds that a necessary condition for $f''(x_1;p) = f'''(x_1;p) = 0$ is that

$$\frac{\sigma_1^2(y_1^2-3\sigma_2^2)}{\sigma_2^2(y_1^2-3\sigma_1^2)} - \frac{y_1^2-\sigma_2^2}{y_1^2-\sigma_2^2} = 0 \tag{7}$$

Since the left side of Eq. (7) is a rational function of y_1, independent of p, there exists an interval about y_1 and an interval of values of p (depending on y) such that for all values y_α in this interval $f''(y_\alpha; p) = 0 \neq f'''(y_\alpha; p)$. And, if these values of y_α are chosen sufficiently close to y_1, for the corresponding values y_β about y_2 and y_γ about y_0 such that $f''(y_\beta;p) = f''(y_\gamma;p) = 0$, one has, due to continuity, $f'''(y_\beta;p) \neq 0$, $f'''(y_\gamma;p) \neq 0$. Thus, for these values of p, $f'(x;p)$ has at least three extreme values on $\sigma_1 < y < \sigma_2$.

Since $f'(x;p)$ can easily be shown to have at most three extreme values on this interval [since $g_1(y)$ can have at most one local minimum on $\sigma_1 < y < \sigma_2$], $f'(x;p)$ has exactly three extreme values on $\sigma_1 < y < \sigma_2$.

Thus, one can conclude that when $\mu_1 = \mu_2$, $f(x;p)$ has either two or six points of inflection depending on whether σ_1 and σ_2 satisfy inequality (3) or inequality (4). Since it is shown in Ref. 2 that when $\mu_1 = \mu_2$, $f(x;p)$ is unimodal for all p, $f(x;p)$ is parapolymodal when it has six points of inflection. This proves Result (1) for $\sigma_2 > \sigma_1$, and it is easy to show that the same result is obtained for $\sigma_1 > \sigma_2$ by dividing Eq. (1) by the second term of $f''(x;p)$.

It is also shown in Ref. 2 that if $(\mu_2 - \mu_1)^2 < (27\sigma_1^2\sigma_2^2)/4(\sigma_1^2+\sigma_2^2)$ $f(x;p)$ is unimodal for all p and if $(\mu_2-\mu_1)^2 > 8\sigma_1^2\sigma_2^2/(\sigma_1^2+\sigma_2^2)$, the bimodality of $f(x;p)$ then depends on p. However, it is also obvious (for fixed values of

p, σ_1, σ_2), that, as $(\mu_2 - \mu_1)^2$ increases, $f(x;p)$ tends continually toward bimodality since each normal distribution of the mixture tends to reproduce itself multiplied by a factor of p or $1 - p$. When $f(x;p)$ is parapolymodal for $(\mu_2 - \mu_1)^2 = 0$, there are six points of inflection. Hence, $f'(x;p) \to 0$ as $x \to \pm \infty$, $f'(x;p)$ is positive and has two maxima and one minimum when x is less than the mode and negative with two minima and one maximum for x greater than the mode. As $f(x;p)$ tends to bimodality as $(\mu_2 - \mu_1)^2$ increases with p, σ_1, σ_2 fixed, for some critical value of $(\mu_2 - \mu_1)^2 > 0$, either the abscissa of the positive minimum or the abscissa of the negative maximum changes sign. (Both of these events cannot occur simultaneously since $f(x;p)$ can have at most three extreme values.) When this sign change occurs, $f(x;p)$ has three extreme values and is bimodal. However, as $(\mu_2 - \mu_1)^2$ increases from zero to the critical value, there are always at least four extreme values and one zero of $f'(x;p)$, which means that if $f(x;p)$ is parapolymodal when $(\mu_2 - \mu_1)^2 = 0$, $f(x;p)$ remains parapolymodal as $(\mu_2 - \mu_1)^2$ increases until it becomes bimodal. This proves Result (2).

Now suppose that for fixed values of μ_1, σ_1, μ_2, σ_2, $(\mu_2 - \mu_1)^2 > 8\sigma_1^2\sigma_2^2/(\sigma_1^2 + \sigma_2^2)$. This means that there are two critical values of p, say p_1 and p_2, such that for $p_1 \leqq p < 1$ or $0 < p \leqq p_2$, $f(x;p)$ is unimodal; and for $p_2 < p < p_1$, $f(x;p)$ is bimodal. When $f(x;p)$ is bimodal, $f'(x;p) \to 0$ as $x \to \pm \infty$, and $f(x;p)$ has three distinct zeros and has two positive maxima and two negative minima. As p increases to p_1 or decreases to p_2, two adjacent zeros of $f'(x;p)$ move toward each other and coincide when p equals one of the critical values. When this occurs, $f'(x;p)$ has one positive maximum, two negative minima and a maximum between the two negative minima which coincides with a zero. But this means that for x_0 equal to the maximum coinciding with a zero, $f(x_0,p)$ is not an extreme value, but it is a point of inflection with zero slope. Thus, $f(x;p)$ has a single extreme value which is a maximum and four points

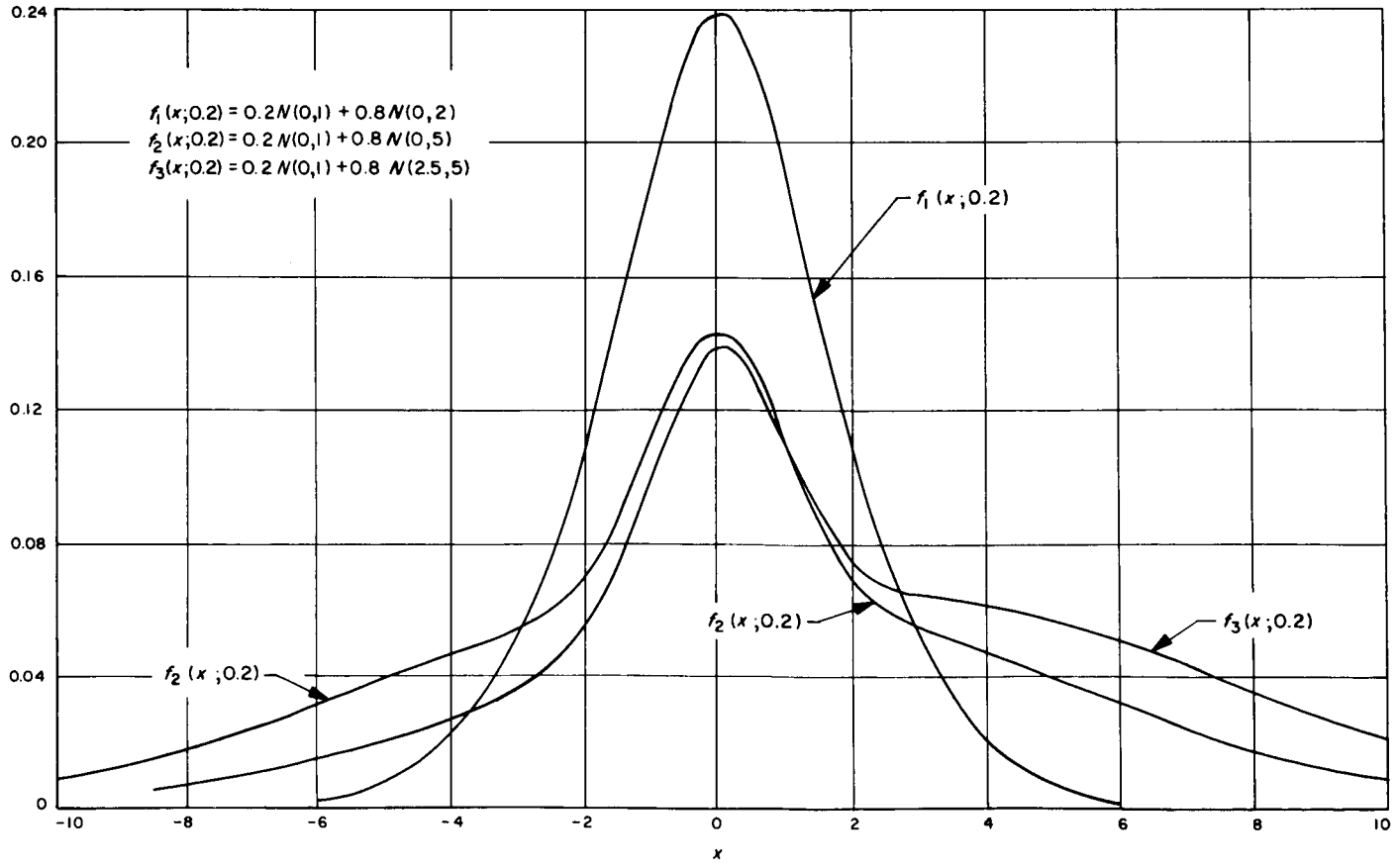


Fig. 1. Parapolymodal and nonparapolymodal distributions for $\mu_1 = \mu_2$, and the effect on paramodality of increasing $(\mu_2 - \mu_1)^2$ when bimodality is impossible

of inflection, and, hence, it is parapolymodal. Therefore, there is an interval of values of p equal to or greater than p_1 and an interval of values of p equal to or less than p_2 for which $f(x;p)$ is parapolymodal. This proves Result (3).

The general form of Eq. (2) is given by

$$g(x) = \frac{(x-\mu_2)^2 - \sigma_2^2}{(x-\mu_1)^2 - \sigma_1^2} h(x) = \frac{-\sigma_2^5 p}{\sigma_1^5(1-p)}$$

where

$$h(x) = \exp\left[-\frac{1}{2}\frac{(x-\mu_2)^2}{\sigma_2^2} + \frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2} \right]$$

In general, when $\mu_1+\sigma_1 \neq \mu_2 \pm \sigma_2$ or $\mu_2-\sigma_1 \neq \mu_2 \pm \sigma_2$, there are two intervals for which $g(x) < 0$ and continuous on each open interval. At one endpoint of each interval $g(x) = 0$, and $g(x) \to -\infty$ as x approaches the other endpoint from within the interval. Thus, regardless of the behavior of $g(x)$ in the interior of each interval, there is a half interval of values of x about each endpoint within each of the two intervals, corresponding to intervals of values of p near zero and one, for which the inverse function $g^{-1}(x)$ is double-valued. For each value of p in these intervals, $f(x;p)$ has only two points of inflection; and, hence, it is neither bimodal nor parapolymodal. For the special case where $\mu_1+\sigma_1 = \mu_2 \pm \sigma_2$ or $\mu_1-\sigma_1 = \mu_2 \pm \sigma_2$, one root of $f''(x;p)$ for all p is $x = \mu \pm \sigma_2$, and since there is now only one interval for which $g(x) < 0$, which has the same properties as the two intervals in the general case, values of p exist sufficiently close to zero and one for which

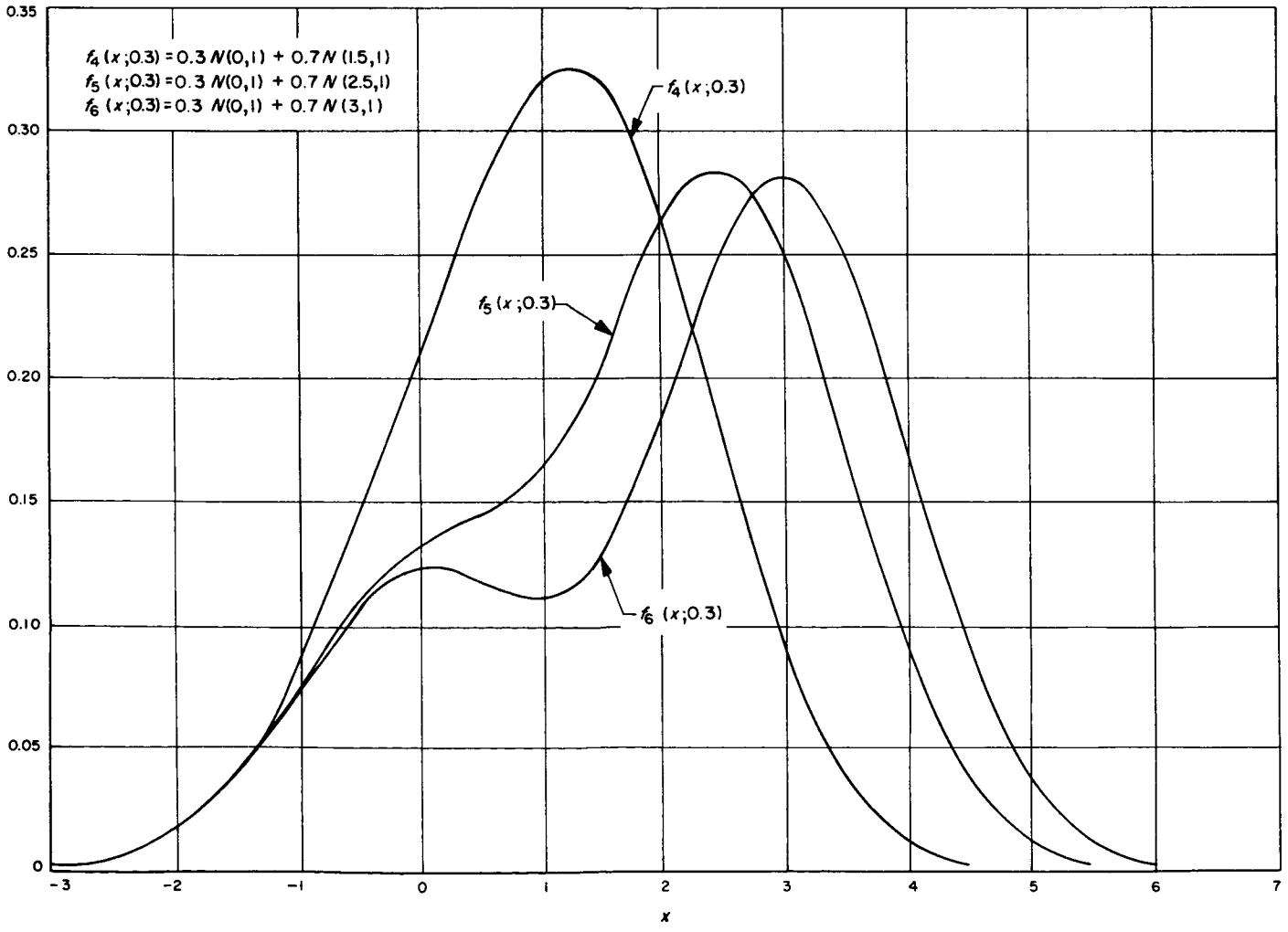


Fig. 2. The effect on modality and paramodality of increasing $(\mu_2-\mu_1)^2$

$f(x;p)$ has only one additional point of inflection. Thus, in all cases, values of p exist for which $f(x;p)$ is neither bimodal nor parapolymodal. This proves Result (4).

4. Examples

In Fig. 1, $f_1(x;.2)$ and $f_2(x;.2)$ are unimodal for all p since in each case $\mu_1 = \mu_2$. $f_3(x;3)$ is unimodal for all p since in this case, $(\mu_2 - \mu_1)^2 = 6.25 < 27\sigma_1^2\sigma_2^2/4(\sigma_1^2 + \sigma_2^2) = 6.49$. However, $f_2(x;.2)$ and $f_3(x;.2)$ are parapolymodal for $p = 0.2$, while $f_1(x;.2)$ is not, since in the case of $f_1(x;.2)$, $\sigma_1^2\sigma_2^2 = 4 > (\sigma_2^2 - \sigma_1^2)^2/8 = 9/8$. Because for $f_3(x;.2)$, $(\mu_2 - \mu_1)^2 > 0$; the paramodality property of $f_3(x;.2)$ is seen to be more pronounced than that for $f_2(x;.2)$.

Fig. 2 illustrates the effect of increasing $(\mu_2 - \mu_1)^2$. For fixed values of p, σ_1 and σ_2, as $(\mu_2 - \mu_1)^2$ increases the distribution progresses from nonparapolymodal to parapolymodal and then finally to bimodal.

The density functions in Fig. 3 are mixtures of the same two normal distributions with different values of p. Since in each case $(\mu_2 - \mu_1)^2 = 9 > 8\sigma_2^2\sigma_1^2/(\sigma_1^2 + \sigma_2^2) = 4$, the bimodality property depends on the value of p. As p increases, it can readily be seen that the density function becomes first parapolymodal and then bimodal. As p continues to increase, the process would reverse; the function would again become parapolymodal and then nonparapolymodal.

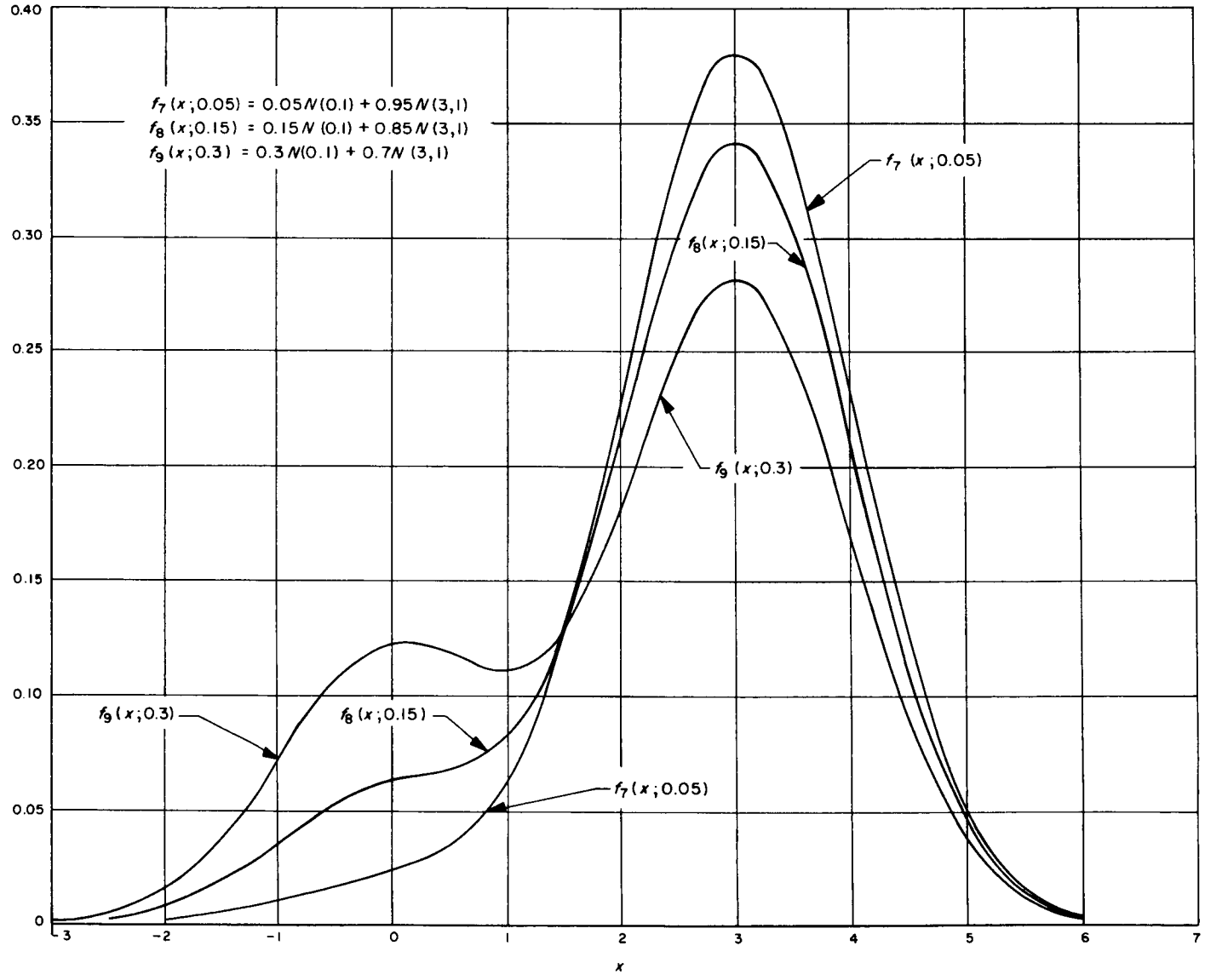


Fig. 3. The effect of p on modality and paramodality when bimodality is possible

The density functions in Fig. 4 are again mixtures of the same normal distributions, but here, since $(\mu_2-\mu_1)^2 = 4 < 27\sigma_1^2\sigma_2^2/4(\sigma_1^2+\sigma_2^2) = 6.35$, the functions are unimodal for all p. However, since $\sigma_1^2\sigma_2^2 = 16 < (\sigma_2^2-\sigma_1^2)^2/8 = 28.125$, Result (2) applies. For a sufficiently large value of p, say $p_0(>0.6)$, the density function is nonparapolymodal for all $p > p_0$. For a sufficiently small value of p say $p_1(<0.1)$, the function is again nonparapolymodal for all $p < p_1$. For $p_1 < p < p_0$, the function is parapolymodal.

It is intuitively reasonable to conclude that the paramodality property of a mixture of two normal distributions, for which the parameters are such that $f(x;p)$ is unimodal for all p, can never be as pronounced as that which is possible for a mixture whose modality depends on p. A comparison between Figs. 3 and 4 illustrates this. Nevertheless, in many cases where $f(x;p)$ is unimodal for all p, particularly when $(\mu_2-\mu_1)^2 > 0$, the paramodality property is sufficiently pronounced for an interval of values of p so that a mixture of two distributions could be suspected. Thus, the effect of requiring parapolymodality for the detection of a mixture of two normal distributions, instead of the stricter condition of bimodality, is to enlarge the set of values of p for which detection is possible.

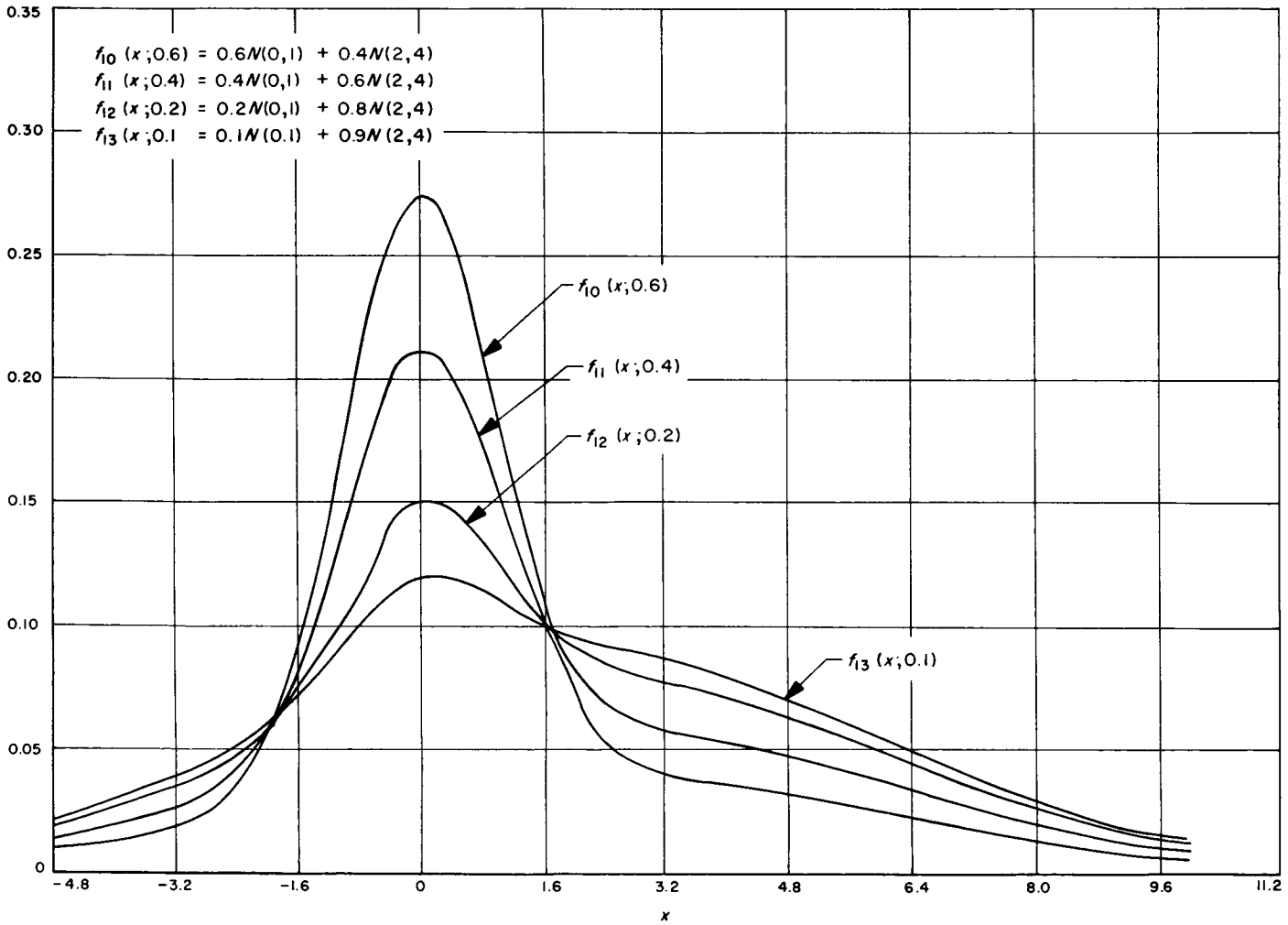


Fig. 4. The effect of p on paramodality when bimodality is impossible

References

1. Barberian, S. K., "Measure and Integration," MacMillian, New York, 1965.

2. Eisenberger, I., "Genesis of Bimodal Distributions," *SPS 37-18*, Vol. IV, Jet Propulsion Laboratory, Pasadena, pp. 178–180, and *Technometrics*, Vol. 6, No. 4, pp. 357–363, November 1964.

XXIV. Communications Systems Research: Coded Communications

A. Effects on the Reliability of Telemetry Data Caused by the Earth-to-Spacecraft Link

W. C. Lindsey

1. Introduction

In the design of two-way communications links (*SPS 37-31*, Vol. IV, pp. 374–379; *SPS 37-32*, Vol. IV, pp. 284–288; *SPS 37-33*, Vol. IV, pp. 290–296; and *SPS 37-34*, Vol. IV, pp. 242–247), it is of interest to understand how the performance of the Earth-to-spacecraft link affects the performance of the spacecraft-to-Earth telemetry link. Consequently, results are presented here which illustrate the performance of the Earth-based telemetry demodulator as a function of the performance of the spacecraft carrier-tracking loop. It is shown that, for two-way communications links, the performance of the spacecraft-to-Earth telemetry channel is essentially unaffected by that of the Earth-to-spacecraft link if the signal-to-noise ratio in the spacecraft carrier-tracking loop is greater than 16 db.

2. Model

In the following discussion, the notation and results reported in the above issues of the *SPS* shall be utilized. Briefly, a two-way coherent communications link may be defined as one which radiates a phase-modulated RF carrier from the Earth to the spacecraft; the spacecraft coherently tracks the noise-corrupted carrier component by means of a phase-locked loop. The phase-locked-loop estimate of the frequency and phase of the transmitted carrier is used (after appropriate frequency translation) as a carrier for the transmission of telemetry data back to Earth. Since carrier tracking is noisy, the phase-locked-loop estimate of the phase possesses a random perturbation which affects the performance of the ground-based telemetry demodulator. Thus, the observed data on Earth is perturbed by both the uplink and downlink additive

noise. It is, however, clear that the effects of the uplink noise can be made arbitrarily small by increasing the power in the uplink carrier component.

It was shown (*SPS 37-33*) that P_E, the bit error probability in an uncoded biphase-modulated telemetry system for the ground system telemetry demodulator, is given by

$$P_E = \frac{1}{\pi}\int_0^{\infty} \frac{I_0|\alpha_1 + \alpha_2 \exp(j\phi)|}{I_0(\alpha_1)\, I_0(\alpha_2)}\, d\phi \times \int_{(2R_D\cos\phi)^{1/2}}^{\infty} (2\pi)^{-1/2} \exp(-y^2/2)dy,$$

where $R_D = S_2T_b/N_{02}$ is the signal-to-noise ratio in the data; α_1, the signal-to-noise ratio in the spacecraft carrier-tracking loop; α_2, the signal-to-noise ratio in the ground receiver carrier-tracking loop; ϕ, the ground system phase error; and $I_0(x)$, the imaginary Bessel function of zero order and of argument x.

3. Results

By fixing R_D in the previous formula, one can easily study the performance of the two-way link as a function of the parameters α_1 and α_2. Plotted in Fig. 1 is the probability of error P_E versus the signal-to-noise ratio α_1 in the spacecraft carrier-tracking loop for various values of the signal-to-noise ratio α_2 in the ground receiver carrier-tracking loop. In particular, we have set $R_D = 4.8$, which, in the absence of any spacecraft and ground receiver timing errors, produces a probability of error of $P_E = 10^{-3}$. For all practical purposes, it can be concluded that, for $\alpha_1 > 40$ (16 db), the performance of the ground-based telemetry demodulator is not affected by the additive noise on the uplink. Hence, development of transmitting systems which could establish (at the greatest range expected) this signal-to-noise ratio in the spacecraft carrier-tracking loop would largely alleviate the deleterious effects produced by the uplink noise in the detection of telemetry data.

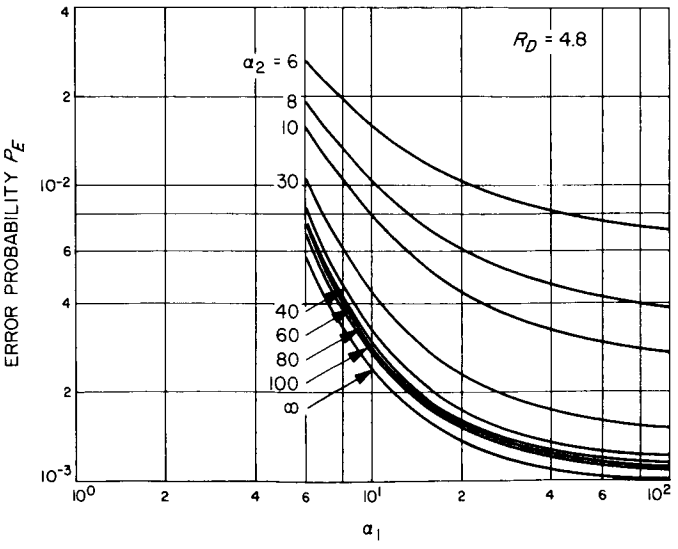


Fig. 1. Error probability P_E vs signal-to-noise ratio α_1 in the spacecraft's carrier-tracking loop, with α_2 as a parameter

B. New Techniques for Determining Weights of Code Words

G. Solomon and R. McEliece

A general technique for determining the weights of code vectors in cyclic codes was presented in *SPS 37-33*, Vol. IV, pp. 278–281. Since that time, much has been learned about this technique. This article, in addition to giving a more lucid explanation of the technique than that in *SPS 37-33*, presents general formulas for certain coefficients, an explicit formula for one of them, certain important applications, and a generalization to cyclic codes over any finite prime field.

1. Introduction

In the past few years, coding theory has witnessed the discovery of new classes of practical codes of a highly algebraic nature. These codes are easily encodable, and their error-correcting procedure is a sequence of easily mechanizable algebraic operations. These codes—e.g., Bose-Chaudhuri, cyclic, polynomial, and shift register (Ref. 1)—use the shift register device for part of the systematic encoding as well as for the multiplication operations in the associated finite fields during decoding. While certain group codes can be characterized by specifying the encoding and decoding procedures in terms of certain polynomials and field parameters, there are, however, many cases for which the precise error-correcting properties of these codes (i.e., the minimum weight of all non-zero vectors of a code) are not known. Even less is known, in general, about the exact distribution of weights in a code; such information is needed to compute error probabilities in various applications of the code.

To be specific, for n odd, let A be any (n,k) group code, i.e., a k-dimensional subspace of $V_n(F)$, the vector space of dimension n over the field F of two elements. The weight $w(a)$ of a vector $a = (a_i)$ in A is defined to be the number of i such that $a_i = 1$. A classical coding theory problem is to determine the error-correcting properties of the code A, i.e., the minimum weight of all non-zero vectors in A. In general, however, the values of all weights actually occurring for words in the code and sometimes the numbers of vectors having these weights are needed. The problem of finding a formula or analytic expression for the weight of a vector in terms of the parameters which characterize the code is made more interesting by the fact that our code operations are all modulo (mod) 2, whereas the weight function is integer-valued.

Presented here is an explicit expression for the weight of a vector:

$$w(a) = \sum_{i=0}^{N} \Gamma_{2^i}(a) 2^i .$$

In the case of cyclic codes, the coefficients Γ_i are expressed as functions of a number of parameters. These parameters arise naturally in the general representation of cyclic codes given in Ref. 2. The coefficients in the formula above are obtained by computing symmetric sums of the nth roots of unity. (Γ_1 has been known for some time.) Using the explicit expressions for Γ_2 and Γ_4, we obtain general results for weights occurring in certain cyclic codes.

2. Binary Representation of the Weight Vector

Let $a = (a_0, a_1, \cdots, a_{n-1})$ be a binary n-tuple. If w is the number of ones in a, then

$$\Gamma_1 = \sum_{i=0}^{n-1} a_i \equiv \binom{w}{1} \pmod 2$$

$$\Gamma_2 = \sum_{i<j} a_i a_j \equiv \binom{w}{2} \pmod 2$$

$$\vdots$$

$$\Gamma_w = \sum_{i_1<i_2<i_3\cdots<i_w} a_{i_1} a_{i_2} \cdots a_{i_w} \equiv \binom{w}{w} \equiv 1 \pmod 2$$

$$\Gamma_l = 0, \quad l > w .$$

The weight $w(a)$ of a is given in binary form uniquely by

$$\sum_{i=1}^{N} w_i 2^i ,$$

where $N = \{\log_2 n\}$ and, as usual, $w_i = 0$ or 1. The following lemma is crucial.

Lemma: If $w = \sum w_i 2^i$ is the binary expansion of w, then $w_i \equiv \Gamma_{2^i} \pmod 2$. More generally, if

$$w = \sum_{i=0}^{N} w_i p^i$$

is the p-ary expansion of the integer w (i.e., with each w_i satisfying $0 \leq w_i \leq p-1$, and $N = \{\log_p w\}$), then

$$w_i \equiv \binom{w}{p^i} \equiv \frac{w!}{p^i!\,(w-p^i)!} \pmod p .$$

Proof: Expanding the polynomial $(1+x)^w$ over the field $GF(p)$, we obtain

$$(1+x)^w = (1+x)^{\sum w_i p^i} = \prod_{i=0}^{N} (1+x)^{w_i p^i} ;$$

$$\sum_{i=0}^{w} \binom{w}{i} x^i = \prod_{i=0}^{N} (1+x^{p^i})^{w_i} = \prod_{i=0}^{N} \left(\sum_{j=0}^{w_i} \binom{w_i}{j} x^{jp^i} \right) .$$

Since this yields a formal identity, the coefficients of x^{p^i} on both sides of the equation must be equal. Since the x^{p^i} term on the right-hand side can only arise by taking

$$\binom{w_i}{1} x^{p^i}$$

and multiplying it by $(N-1)$ ones, we obtain

$$\binom{w_i}{p} = \binom{w_i}{1} \equiv w_i \pmod p ,$$

which proves the lemma.

The weight of the vector a is therefore given by

$$w(a) = \sum_{i=0}^{N} \Gamma_{2^i}(a) 2^i .$$

We shall express the $\Gamma_{2^i}(a)$ as functions of the coefficients of the Mattson–Solomon polynomials for cyclic codes (Ref. 2). This formula greatly simplifies the exponential weight formula given in Ref. 3 since, in that formula, all the Γ_i were needed.

3. Cyclic Codes and Mattson–Solomon Polynomials

Let A be an (n,k) cyclic code generated by the recursion

$$\sum_{i=0}^{k} u_i a_{n+i} = 0$$

whose associated polynomial

$$f(x) = \sum_{i=0}^{k} u_i x^i$$

divides $x^n + 1$. Let β be a fixed primitive nth root of unity. Then, with every code word a in A, there is associated a polynomial

$$g_a(x) = \sum_{i=0}^{n-1} c_i x^i$$

with the following six properties (details given in Ref. 2):

(1) If $a = (a_0, a_1, \cdots, a_{n-1})$, then $g_a(\beta^j) = a_j$.

(2) The c_j are given by the Reed formula:

$$c_j = \sum_{i=0}^{n-1} a_i \beta^{-ij}, \qquad j = 0, 1, \cdots, n-1.$$

(Note that the c_j are thus contained in the field containing the nth roots of unity, that

$$c_0 = \sum_{i=0}^{n-1} a_i \equiv w(a) \pmod 2,$$

and that $c_{2j} = c_j^2$ for all j.)

(3) The $c_j = 0$ whenever $f(\beta^j) \neq 0$; i.e., the only c_j that can enter in the expression for $g_a(x)$ are those c_j such that $f(\beta^j) = 0$.

(4) If $a = (a_0, a_1, \cdots, a_{n-1})$ is given by $(g_a(\beta^j))$, $g_a(x) = \Sigma\, c_j x^j$, then the cyclic shift of a, denoted by $Ta = (a_1, a_2, \cdots, a_{n-1}\ a_0)$, is given by $(g_{Ta}(\beta^j))$, where $g_{Ta}(x) = \Sigma\, c_j \beta^j$. (Shifting a vector is thus equivalent to multiplying each c_j by β^j.)

(5) If $f(x) = f_1(x) f_2(x) \cdots f_r(x)$, where each $f_i(x)$ is irreducible over F, then there are r independent parameters c_{i_r} in $g_a(x)$ whose subscripts i_r satisfy $f_j(\beta^{i_j}) = 0$, but $f_j(\beta^{i_k}) \neq 0$ if $j \neq k$, and $g_a(x)$ itself satisfies

$$g_a(\beta^i) = \mathrm{Tr}(c_{i_1}\beta^{ii_1}) + \mathrm{Tr}(c_{i_2}\beta^{ii_2}) + \cdots + \mathrm{Tr}(c_{i_r}\beta^{ii_r}),$$

where

$$\mathrm{Tr}(c\beta^j) = \sum_{k=1}^{m} (c\beta^j)^k$$

is the *trace* of $c\beta^j$, and m is the degree of the minimum polynomial for β^j.

We note that the c_0 term in $g_a(x)$ depends on the root $\beta^0 = 1$ of $f(x)$; hence, c_0 corresponds to the factor $(x + 1)$ in $f(x)$. Thus, $c_0 = 1$ if and only if the weight of a is odd. If $(x + 1)$ divides $f(x)$, then clearly every odd-weight vector is the sum of the all-ones vector and an even-weight vector, thus proving the result.

(6) The odd weights are computed from the even weights w_{2i} by taking $n - w_{2i}$; thus, we need a formula for the even-weight vectors only.

4. Coefficients $\Gamma_{2^i}(a)$ as Polynomials in c_j

We have shown that

$$w(a) = \sum_{i=1}^{N} \Gamma_{2^i}(a) 2^i$$

for $N = \log_2 n$, and, as above, we may assume that $w(a) \equiv 0 \pmod 2$. Now

$$\begin{aligned}\Gamma_{2^i}(a) &= \sum_{j_1 < j_2 \cdots < j_{2^i}} a_{j_1} a_{j_2} \cdots a_{j_{2^i}} \\ &= \sum g_a(\beta^{j_1})\, g_a(\beta^{j_2}) \cdots g_a(\beta^{j_{2^i}}) \\ &= P(c_{j_1}, c_{j_2} \cdots c_{j_{2^i}}),\end{aligned}$$

which is a polynomial in the c's. Since $w(Ta) = w(a)$, and since Property (4) of Section 3 above holds, we have

$$P(c_{j_1}, c_{j_2}, \cdots, c_{j_{2^i}}) = P(\beta^{j_1} c_{j_1}, \beta^{j_2} c_{j_2}, \cdots, \beta^{j_{2^i}} c_{j_{2^i}}).$$

Consequently,

$$\sum_{i=1}^{2^i} j_i \equiv 0 \pmod n.$$

This result is obtained by another method in Section 7.

5. Computation of $\Gamma_2(a)$

Using elementary methods, we now evaluate $\Gamma_2(a)$ and, with this value, obtain some new results about cyclic codes.

Theorem 1: If $w(a) \equiv 0 \pmod 2$, then

$$\Gamma_2(a) = \sum_{i=1}^{(n-1)/2} c_i c_{n-i}.$$

Proof: Again we recall that, according to Property (6) in Section 3, we may assume $w(a) \equiv 0 \pmod 2$, i.e., that $c_0 = 0$. Substituting $x_i = \beta^i$ and $y_j = \beta^j$, we have

$$\begin{aligned}\Gamma_2(a) &= \sum_{i<j} a_i a_j = \sum_{i<j} g_a(\beta^i)\, g_a(\beta^j) \\ &= \sum_{i<j} \left(\sum_{r=1}^{n-1} c_r x_i^r\right)\left(\sum_{s=1}^{n-1} c_s y_j^s\right) \\ &= \sum_{i<j} \sum_{r,s} c_r c_s \left(x_i^r y_j^s + x_i^s y_j^r\right).\end{aligned}$$

The second sum in the last equation is taken over all possible unordered terms $c_r c_s$. As a result, $\Gamma_2(a)$ takes the form

$$\Gamma_2(a) = \sum_{r,s} A_{r,s}\, c_r c_s,$$

where the $A_{r,s}$ are given by

$$\begin{aligned}A_{r,s} &= \sum_{i<j} \left(x_i^r y_j^s + x_i^s y_j^r\right) \\ &= \sum_{i \neq j} x_i^r y_j^s . \qquad (1)\end{aligned}$$

But, since for any x_i, y_j it is true that

$$\left(\sum_{i=0}^{n-1} x_i^r\right)\left(\sum_{i=0}^{n-1} y_j^s\right) = \sum_{i \neq j} x_i^r y_j^s + \sum_{i=0}^{n-1} x_i^{r+s}; \qquad (2)$$

since, if $x_0, \cdots, x_{n-1}$ are the nth roots of unity over any field,

$$\sum_{i=0}^{n-1} x_i^r = \begin{cases} 0 \text{ if } r \not\equiv 0 \pmod n, \\ n \text{ if } r \equiv 0 \pmod n; \end{cases}$$

and since we may assume $r \not\equiv 0$, $s \not\equiv 0 \pmod n$, it follows that

$$0 = A_{r,s} + \sum_{i=0}^{n-1} x_i^{r+s}.$$

Thus,

$$A_{r,s} = \begin{cases} 0 \text{ unless } r + s \equiv 0 \pmod n, \\ -n \equiv 1 \pmod 2 \text{ if } r + s \not\equiv 0 \pmod n. \end{cases}$$

So, from Eq. (1) there now remains

$$\Gamma_2(a) = \sum_{i=1}^{(n-1)/2} c_i c_{n-i},$$

the assertion of the theorem.

6. Corollaries of the Γ_2 Formula

Corollary 1: Let A be a cyclic code generated by $f(x)$. If, for every γ such that $f(\gamma) = 0$, we have $f(\gamma^{-1}) \not\equiv 0$, then Γ_2 for the code A is 0.

Proof: Note that c_j corresponds to $f(\beta^j) = 0$. Thus, c_i and c_{-i} are never simultaneously non-zero by hypothesis.

Corollary 2: All cyclic codes A with $\Gamma_2(a) = 0$ for all $a \in A$ have, for their even weights w, $w \equiv 0 \pmod 4$.

Proof: If $w(a) \equiv 2 \pmod 4$, then $\Gamma_2(a) = 1$, since

$$\begin{aligned}\Gamma_2 &\equiv \binom{w}{2} \equiv \frac{(4m+2)(4m+1)}{1 \cdot 2} \\ &\equiv (2m+1)(4m+1) = 1 \pmod 2.\end{aligned}$$

Corollary 3: All cyclic codes A of length $n \equiv 3 \pmod 4$ with $\Gamma_2(a) = 0$ for all $a \in A$ have only weights $w \equiv 0$ or $3 \pmod 4$ occurring. Similarly, if $n \equiv 1 \pmod 4$ and $\Gamma_2 = 0$ for all $a \in A$, only weights $w \equiv 0$ or $1 \pmod 4$ can occur.

Proof: This follows readily from Corollary 2 and Property (6) of Section 3.

Theorem 2: Every vector in the class of $[p, (p+1)/2]$ cyclic codes, with $p \equiv -1 \pmod 8$ and

$$x^p + 1 = (x+1) f_0(x) f_1(x),$$

where $f_0(x)$ and $f_1(x)$ are irreducible of degree $(p-1)/2$, has weight $w \equiv 0$ and $3 \pmod 4$.

Proof: It can be shown that this class of codes satisfies Corollary 1. Hence, $\Gamma_2 = 0$ for even-weight vectors. (The details are omitted.)

Theorem 2, when applied to the famous Golay (23, 12) code (Ref. 1, p. 70), proves that the code is triple-error-correcting. (See also an independent proof that $\Gamma_2 = 0$ for this code in Section 9 below.) For proof, note that the Bose–Chaudhuri bound indicates that the minimum weight of code words is at least 5. But, $5 \not\equiv 3$ (4), so 7 is the minimum possible odd weight. Further, no words of weight 6 occur, since $6 \not\equiv 0$ (4). This weight of 7 must be attained, for the sphere-packing bound limits the minimum weight to at most 7. Other proofs of this result have been given by computer search, as well as by projective geometry methods[1], but the present proof is the shortest known.

[1]Gleason, A., Harvard University, Cambridge, Massachusetts (Unpublished results).

7. Higher-Order Γ's

We now present a method for computing Γ_r for any integer r. Γ_r may be defined as follows:

$$\Gamma_r(a) = \sum_{i_1<i_2\cdots<i_r} a_{i_1} a_{i_2} \cdots a_{i_r},$$

In terms of the Mattson–Solomon polynomials for cyclic codes, this is

$$\Gamma_r(a) = \sum_{i_1<i_2\cdots<i_r} g_a(\beta^{i_1}) g_a(\beta^{i_2}) \cdots g_a(\beta^{i_r}),$$

where β is a primitive nth root of unity. For reasons which will become clear later in this discussion, we shall let $\beta^j = x_j$ in what follows, and consider $\Gamma_r(a)$ as a polynomial in the n variables $x_0, x_1, \cdots, x_{n-1}$. By definition,

$$g_a(x) = \Sigma c_j x^j;$$

thus,

$$\Gamma_r(a) = \sum_{i_1<i_2\cdots<i_r} \left[\prod_{k=1}^{r} \left(\sum_{j=0}^{n-1} c_j x_{i_k}^j \right) \right]$$

$$= \sum_{i_1<i_2\cdots<i_r} \left[\sum_{(A)} c_{j_1} c_{j_2} \cdots c_{j_r} \left(\sum_{(B)} x_{k_1}^{j_1} x_{k_2}^{j_2} \cdots x_{k_r}^{j_r} \right) \right],$$

where the summation indicated as (A) above is taken over all

$$\binom{n+r-1}{r}$$

unordered r-tuples (repetition allowed) from the set $\{0, 1, \cdots, n-1\}$, and the summation indicated as (B) is taken over all *distinct* terms obtainable by permutations of the variables $x_{i_1}, x_{i_2}, \cdots, x_{i_r}$. We now consider $\Gamma_r(a)$ as a polynomial in the c's and collect terms:

$$\Gamma_r(a) = \sum_{(A)} A_{j_1, j_2, \cdots, j_r} c_{j_1} c_{j_2} \cdots c_{j_r},$$

where (A) is as described, and

$$A_{j_1, j_2, \cdots, j_r} = \sum x_{i_1}^{j_1} x_{i_2}^{j_2} \cdots x_{i_r}^{j_r},$$

which shall be denoted as (*) in what follows. This summation is taken over all distinct terms obtainable by selections of the r distinct variables $x_{i_1}, x_{i_2}, \cdots, x_{i_r}$. We observe that $A_{j_1, j_2, \cdots, j_r}$ is a symmetric polynomial in the n variables $x_0, x_1, \cdots, x_{n-1}$. We now need some information about expressions of the type (*).

Consider polynomials in n variables $x_1, x_2, \cdots, x_n$ over a field F. Let $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_r)$ be a collection of non-negative integers, and define the symmetric polynomial

$$k(\lambda) = \sum x_{i_1}^{\lambda_1} x_{i_2}^{\lambda_2} \cdots x_{i_r}^{\lambda_r}, \qquad (1)$$

where the sum is taken over all distinct terms which can be obtained by arbitrary combinations (without repetition) of the variables. If λ has only one element, say $\lambda = (\lambda_1)$, then $k(\lambda) = \Sigma x_i^{\lambda_1}$ is an ordinary power-sum symmetric function. In what follows, if m is a non-negative integer, $k(m)$ will always mean $k(\lambda)$, where $\lambda = (m)$ is a one-element collection. We shall now develop $k(\lambda)$ as a polynomial in these power-sum symmetric functions $k(m)$.

We assume initially that the characteristic of F is zero. Let m_i be the multiplicity of the integer i in λ. Then,

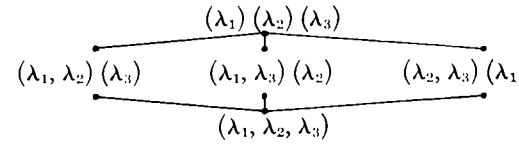
$$\pi(m_i!)\, k(\lambda) = \sum x_{i_1}^{\lambda_1} x_{i_2}^{\lambda_2} \cdots x_{i_r}^{\lambda_r} = h(\lambda), \qquad (2)$$

where now the sum is taken over *all*

$$\binom{n}{r}$$

combinations of the variables x_{i_j}. In effect, then, we can regard $h(\lambda)$ as a $k(\lambda)$ in which the λ_i are regarded as formally distinct.

Now consider the set P_λ of all partitions of the set $\lambda = \{\lambda_1, \lambda_2, \cdots, \lambda_r\}$. (It will always be clear whether we are considering λ as a collection of integers or a set with r elements, so we need not formally distinguish these logical differences.) We introduce a partial ordering $\geqslant$ on P_λ as follows: if p and q are partitions of λ, then $p \geqslant q$ if and only if a set of p is contained in some set of q (i.e., $p \geqslant q$ if p is a *refinement* of q). Thus, if $r = 3$, we could diagram P_λ as below:



It is easy to see that, under $\geqslant$, P_λ becomes a lattice with top element $t_\lambda = (\lambda_1)(\lambda_2)\cdots(\lambda_r)$.

Next, we introduce two functions on P_λ: namely, if $p \in P_\lambda$ is the partition

$$\lambda = \bigcup_{i=1}^{t} S_i,$$

and if

$$\sigma_i = \sum_{\lambda_j \in S_i} \lambda_j,$$

then we define

$$f(p) = h(\sigma_1)\, h(\sigma_2) \cdots h(\sigma_\ell)\,,$$

$$g(p) = h(\sigma), \text{ where } \sigma = (\sigma_1, \sigma_2, \cdots, \sigma_\ell)\,.$$

For example, if $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6)$ and $p = (\lambda_1, \lambda_3, \lambda_4)$ (λ_2, λ_6) (λ_5), then

$$f(p) = h(\lambda_1 + \lambda_3 + \lambda_4)\, h(\lambda_2 + \lambda_6)\, h(\lambda_5)\,,$$

$$g(p) = h(\lambda_1 + \lambda_3 + \lambda_4,\ \lambda_2 + \lambda_6,\ \lambda_5)\,.$$

The following lemma is crucial to our development.

Lemma: $f(p) = \sum_{q \le p} g(q)\,.$

Proof: With notation as introduced above,

$$f(p) = \left(\sum x_{\alpha_1}^{\sigma_1}\right)\left(\sum x_{\alpha_2}^{\sigma_2}\right) \cdots \left(\sum x_{\alpha_\ell}^{\sigma_\ell}\right)$$

$$= \sum x_{\alpha_1}^{\sigma_1} x_{\alpha_2}^{\sigma_2} \cdots x_{\alpha_\ell}^{\sigma_\ell}\,,$$

the sum being taken over all n^ℓ possible ℓ-tuples $(\alpha_1, \alpha_2, \cdots, \alpha_\ell)$. To each such ℓ-tuple, we assign a partition of the collection $(S_1, S_2, \cdots, S_\ell)$ as follows: S_i and S_j are in the same set if and only if $\alpha_i = \alpha_j$. We call two monomials from the above sum *equivalent* if their subscripts determine the same partition of $(S_1, S_2, \cdots, S_\ell)$. Thus, the terms of $f(p)$ divide themselves into equivalence classes, and, to each such equivalence class, there corresponds in a natural way a partition of λ with $q \le p$; i.e., to a partition of $(S_1, S_2, \cdots, S_\ell)$, such as $(S_{i_1}, \cdots, S_{i_r})$ $(S_{j_1}, \cdots, S_{j_s})$ $\cdots$, we associate the partition of λ: $(S_{i_1} \cup \cdots \cup S_{i_r})$, $(S_{j_1} \cup \cdots \cup S_{j_s})$, $\cdots$. Clearly, the sum of the terms in a given equivalence class of $f(p)$ is the corresponding $g(q)$. This proves the lemma.

In view of the lemma, we may use *Möbius inversion* on the lattice P_λ to determine $g(t_\lambda) = h(\lambda)$:

$$h(\lambda) = g(t_\lambda) = \sum_{p \in P_\lambda} \mu(p)\, f(p)\,, \tag{3}$$

where μ is the Möbius function associated with P_λ. As given in Ref. 4, for example, if p is given by

$$\lambda = \bigcup_{i=1}^{\ell} S_i$$

(with no $S_i = \phi$), and if S_i has s_i elements, then

$$\mu(p) = (-1)^{r-\ell} \prod_{i=1}^{\ell} (s_i - 1)!$$

Since each $f(p)$ is a product of power-sum symmetric functions, Eq. (3) is the required expansion. Finally, from Eq. (2), we see that

$$k(\lambda) = \frac{1}{\pi(m_i!)} \sum_{p \in P_\lambda} \mu(p)\, f(p)\,. \tag{4}$$

The special cases $r = 1$, $r = 2$ of Eq. (4) appear in Ref. 5. Also, we note that Eq. (4) gives almost immediately a formula for the elementary symmetric functions in terms of the power-sum symmetric functions over a field of characteristic 0. Thus, if we let

$$\lambda = \overbrace{(1, 1, \cdots, 1)}^{n}$$

in Eq. (4), it is easy to see that

$$k(\lambda) = \sigma_n = \frac{1}{n!} \sum_{\lambda} \mu(\lambda) S_\lambda\,,$$

with the above sum extended over all *number theoretic* partitions of n. Also, if λ is of the form $(1^{r_1} 2^{r_2} 3^{r_3} \cdots)$, then

$$\mu(\lambda) = \frac{(-1)^{n-\ell}\, n!}{1^{r_1} 2^{r_2} \cdots r_1! r_2!}\,,$$

where $\ell = \ell(\lambda)$ is the number of parts in the partition λ. Finally, $S_\lambda = S_{\lambda_1} S_{\lambda_2} \cdots$, where $S_r = \Sigma x_i^r$ is the power-sum symmetric function. Thus, for example,

$$\sigma_1 = S_1\,,$$

$$\sigma_2 = \frac{1}{2}\left(S_1^2 - S_2\right)\,,$$

$$\sigma_3 = \frac{1}{6}\left(S_1^3 - 3S_1 S_2 + 2S_3\right)\,,$$

$$\sigma_4 = \frac{1}{24}\left(S_1^4 - 6S_1^2 S_2 + 3S_2^2 + 8S_1 S_3 - 6S_4\right)\,.$$

We are now able to compute the functions $k(\lambda)$ in the special case where the x_i are the nth roots of unity over F (still temporarily of characteristic zero). In this case, we know that the power-sum symmetric functions (k) are all zero except when $m \equiv 0 (\text{mod } n)$, in which case $k(m) = n$. Thus, if P_λ^0 represents the subset of P_λ of partition $\lambda = S_i$ with each $\sigma_i \equiv 0 (\text{mod } n)$, then

$$k(\lambda) = \frac{1}{\pi m_i!} \sum_{p \in P_\lambda^0} \mu(p)\, n^{\ell(p)}\,, \tag{5}$$

where $\ell = \ell(p)$ is the number of parts of the partition p.

All of our considerations so far have been made assuming the characteristic of F to be zero. We now discuss the possibility of relaxing this assumption.

The polynomial $k(\lambda)$, being symmetric, can be expressed as a polynomial in the elementary symmetric functions, i.e., functions $k(\nu)$ with $\nu = (1,1,\cdots,1)$ (e.g., Ref. 6). Further, since $k(\lambda)$ has *integral* coefficients, it is well-known that the representation of $k(\lambda)$ as such a polynomial will also have integral coefficients. So, suppose we have $k(\lambda) = P$, with P a polynomial in the elementary symmetric functions. Now, the relation $k(\lambda) = P$ is simply a polynomial identity and thus is valid over any field F, whether or not the characteristic is zero. (Of course, we must consider the coefficients modulo the characteristic of the field.) Also, if the x_i are the nth roots of unity over any field F, then all the elementary symmetric functions are zero except $k(\nu)$ with

$$\nu = \underbrace{(1,1,\cdots,1)}_{n},$$

in which case $k(\nu) = (-1)^n$. Hence, if we can compute $k(\lambda)$ when the x_i are the nth roots of unity over any field of characteristic 0, and if we reduce this value (mod p), the resulting value is the same as $k(\lambda)$ in any field of characteristic p. But, Eq. (5) is such an expression and therefore is valid whether or not the characteristic of F is zero.

Thus, Eq. (5) is exactly what we need to evaluate the coefficients in (*), since the x_i in (*) are nth roots of unity over $GF(2)$. If $\lambda = (j_1, j_2, \cdots, j_r)$ is a subscript as in (*), we define an *admissible partition* of λ as a partition of λ such that the sum of the elements in each of the parts $\equiv 0 \pmod{n}$. Thus, if $n = 7$ and $\lambda = (1,2,5,6)$, the admissible partitions of λ are (1,2,5,6) and (16)(25). We state these results in the next theorem.

Theorem 3:

$$\Gamma_r(a) = \sum_{(A)} A_{j_1, j_2, \cdots, j_r} c_{j_1} c_{j_2} \cdots c_{j_r},$$

where

$$A_{j_1, j_2, \cdots, j_r} = A_\lambda = k(\lambda) \pmod 2$$

with $\lambda = (j_1, j_2, \cdots, j_r)$ in Eq. (5).

Corollary: If $j_1 + j_2 + \cdots + j_r = n$, or, more generally, if the only admissible partition of λ is the one-part partition, then

$$A_\lambda = \frac{(-1)^{r-1}}{\pi m_i!}(r-1)!n \equiv \frac{(r-1)!}{\pi m_i!} \pmod 2.$$

Several remarks are in order here. It is clear that, if $j_1 + j_2 + \cdots + j_n \not\equiv 0 \pmod n$, no partition of λ is admissible, and so $A_\lambda = 0$. Also, as we noted earlier, we can assume without loss of generality that the vector has even weight, so $c_0 = 0$; hence, we need not compute A_λ if any $j_i = 0$. Finally, the theorem gives us as a trivial corollary the formula for Γ_2, obtained above, since $j_1 + j_2 \equiv 0 \pmod n$ if and only if $j_1 = n - j_2$. If n is odd, we can assume $j_1 \neq j_2$ by our above remarks. Thus, $A_{j_1, n-j_1} = -n \equiv 1 \pmod 2$.

8. Computation of Γ_4

We next use Theorem 3 to compute Γ_4. We divide the subscript vectors $\lambda = (j_1, j_2, j_3, j_4)$ into three classes:

Class I: The only admissible partition of λ is (j_1, j_2, j_3, j_4). Here

$$A_\lambda \equiv \frac{-6n}{\pi m_i!} \pmod 2.$$

Therefore, $A_\lambda \equiv 1 \pmod 2$ if and only if λ is of one of the following forms:

$$\lambda = (r, r, s, t)\ ; \quad 2r + s + t \equiv 0 \pmod n,$$

$$\lambda = (r, r, r, s)\ ; \quad 3r + s \equiv 0 \pmod n.$$

Class II: $\lambda = (i, j, n - i, n - j)$ with $i \neq j$. Then, λ has the admissible partitions:

$$\lambda = (i, j, n - i, n - j),$$

$$\lambda = (i, n - i)(j, n - j).$$

Hence, $A_\lambda = -6n + n^2 \equiv 1 \pmod 2$.

Class III: $\lambda = (i, i, n - i, n - i)$. Hence, λ has three admissible partitions of two types:

$$\lambda = (i, i, n - i, n - i),$$

$$\lambda = (i, n - i)(i, n - i).$$

Two of the admissible partitions of λ are of the latter type. Thus,

$$A_\lambda = \tfrac{1}{4}(-6n + 2n^2) = \frac{n^2 - 3n}{2} = \binom{n}{2} - n.$$

In this case, it follows that

$$A_\lambda \equiv 0 \pmod 2 \text{ if } n \equiv -1 \pmod 4$$

$$\equiv 1 \pmod 2 \text{ if } n \equiv 1 \pmod 4.$$

Summarizing the above results, for even-weight vectors we have

$$\Gamma_4 \equiv \sum_{3i+j\equiv 0 \pmod n} c_i^3 c_j + \sum_{2i+j+k\equiv 0 \pmod n} c_i^2 c_j c_k + \sum_{i<j} c_i c_{n-i} c_j c_{n-j} + \left[1 + \binom{n}{2}\right] \sum_i c_i^2 c_{n-i}^2 \pmod 2 .$$

Theorem 3 is applicable to the higher-order Γ's, but the enumeration of all possible classes of subscript vectors becomes increasingly tedious. Thus, for Γ_8, there are seven types of possible admissible partitions, corresponding to the 7 partitions of the integer 8 with all parts ≥ 2; the situation is further complicated by the fact that a given subscript vector can have many possible admissible partitions.

9. Applications of Γ_4

We shall compute Γ_4 for the Golay (23, 12) code. The recursion polynomial for this code is $f(x) = x^{11} + x^9 + x^7 + x^6 + x^5 + x + 1$; $(x^{23}+1)/(x+1)$ is the product of two irreducible polynomials. Using the properties of the Mattson–Solomon polynomials $g_a(x)$ listed in Section 3 above (details given in Ref. 2), we see that, for every vector a of even weight, we have the trace polynomial

$$g_a(x) = cx + c^2x^2 + c^{2^2}x^4 + c^{2^3}x^8 + c^{2^4}x^{16} + c^{2^5}x^9 + c^{2^6}x^{18} = c^{2^7}x^{13} + c^{2^8}x^3 + c^{2^9}x^6 + c^{2^{10}}x^{12}$$

for some c in $GF(2^{11})$; $g_a(\beta^i) = a_i$ for some specific root β of $f(x)$. Note that this immediately shows $c_i c_{n-i} = 0$ for all i, and so $\Gamma_2 = 0$. Hence, the weights of the code satisfy $w \equiv 0$ and 3(mod 4). Our previous calculation of Γ_4 now tells us that

$$\Gamma_4 = \sum_{\substack{i=1 \\ 3+l=23}}^{22} c_i^3 c_{li} + \sum_{\substack{i=1 \\ 2+j+k=23}}^{22} c_i^2 c_{ji} c_{ki}.$$

The first of these terms is zero. Hence,

$$\Gamma_4 = \mathrm{Tr}\left[c^2(c_9c_{12} + c_8c_{13} + c_3c_{18})\right] = \mathrm{Tr}\left[c^2(c^{2^5+2^{10}} + c^{2^3+2^7} + c^{2^6+2^8})\right].$$

Since c is in $GF(2^{11})$, we have $c^{2^{11}} = c^{23\cdot 89} = 1$ if $c \neq 0$. Therefore, if we choose γ as a primitive 89th root of unity, we know that setting $\gamma = c$ yields

$$\Gamma_4(c) = \Gamma_4(\gamma) = \Gamma_4(\gamma\beta^l), \quad l = 0, 1, \cdots, 22,$$

since $\gamma\beta^l$ simply represents a cyclic shift of the original vector. In order to find those vectors a with $\Gamma_4 = 1$, we need to find only those γ such that $\Gamma_4(\gamma) = 1$. Now,

$$\Gamma_4(\gamma) = \mathrm{Tr}\left[\gamma^{2+32+1024} + \gamma^{2+8+128} + \gamma^{2+64+256}\right] = \mathrm{Tr}\left[\gamma^9 + \gamma^9 + \gamma^9\right] = \mathrm{Tr}\left[\gamma^9\right].$$

(Here we have used the fact that the trace is linear and is the same for conjugate elements.)

By consulting tables of the factorization of x^n+1, we find that there are 56 values of γ such that $\mathrm{Tr}(\gamma^9) = 1$ (of course, γ^9 is also a primitive 89th root of unity) and with zero trace. Therefore, there are 56·23 = 1288 code words with $\Gamma_4 = 1$. All even-weight vectors thus satisfy $w = 4 \pmod 8$. Since the Bose–Chaudhuri bound assures us that $w \geq 5$, there are consequently 1228 code words of weight 12. These words have Mattson–Solomon polynomials $g_a(x) = \mathrm{Tr}(\gamma\beta^l x)$, $\beta^{23} = 1$, where $\gamma^{89} = 1$, $\mathrm{Tr}(\gamma^9) = 1$, and $l = 0,1,2,\cdots, 22$.

We note also that there are 56·23 code words of weight 11 given by $g_a(x) = 1 + \mathrm{Tr}(\gamma\beta^l x)$ in the (23,12) code. This gives us an *algebraic* method for generating a nonlinear (23,11) code with 2576 vectors, all of which have nearly the same number of zeros and ones, such that the distance between code words is at least 7. If we add a parity check bit, we obtain a nonlinear code of length 24, with all weights 12 and distance at least 8. Such codes are useful in command systems in which balance must be obtained between the number of zeros and ones in a code word.

10. Generalizations to GF(p), With p a Prime

Let A be an (n,k) cyclic code over $GF(p)$. The weight $w(a)$ of a code vector a is defined to be the number of non-zero components of a, just as in the binary case. Let $\bar{a}$ be the vector obtained from a by changing every non-zero component of a to one. Then, if the Mattson–Solomon polynomials associated with a and $\bar{a}$ are denoted by $g_a(x)$ and $h_a(x)$, respectively,[2] we see that $[g_a(x)]^{p-1} = h_a(x)$, since, if $g_a(\beta^j) \neq 0$, $g_a^{p-1}(\beta^j) = 1$. Using the polynomials $h_a(x)$, the analogy with the case $p = 2$ is almost exact.

If $w = \sum w_i p^i$ is the p-ary expansion of w, then

$$w_i \equiv \binom{w}{p^i} \pmod p$$

[2] Solomon, G., *Introduction to Algebraic Coding Theory* (In preparation).

from the lemma in Section 2. If we now define the Γ's by reference to the case $p = 2$, we have

$$\Gamma_1(a) = \Gamma_1(\bar{a}) = \sum \bar{a}_i \equiv \binom{w}{1} \pmod{p},$$

$$\Gamma_2(a) = \Gamma_2(\bar{a}) = \sum \bar{a}_i \bar{a}_j \equiv \binom{w}{2} \pmod{p},$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$\Gamma_r(a) = \Gamma_r(\bar{a}) = \sum_{i_1<i_2\cdots<i_r} \bar{a}_{i_1}\bar{a}_{i_2}\cdots\bar{a}_{i_r} \equiv \binom{w}{r} \pmod{p}.$$

Thus, our basic formula

$$w(a) = \sum \Gamma_{p^i}(a)p^i$$

remains true. As we observed earlier, Eq. (5) in Section 7 is true for fields of any characteristic, so the formula there can still be applied to calculate the Γ's. We must, however, apply that formula to

$$h_a(x) = \sum_{i=0}^{n-1} \bar{c}_i x^i .$$

Thus, in particular,

$$\Gamma_1(a) = n\bar{c}_0 ,$$

$$\Gamma_2(a) = (n^2-n)\bar{c}_0^2 - n\sum_2^0 \bar{c}_i\bar{c}_{n-i} .$$

References

1. Peterson, W. W., *Error-Correcting Codes*, John Wiley & Sons, Inc., New York, 1961.
2. Mattson, H. F., and Solomon, G., "A New Treatment of Bose–Chaudhuri Codes," *Journal of the Society of Industrial Applied Mathematics*, Vol. 9, pp. 654–669, 1961.
3. Solomon, G., "A Weight Formula for Group Codes," *IRE Transactions on Information Theory*, Vol. 1T-8, pp. 1–4, 1962.
4. Rota, G. C., "On the Foundations of Combinatorial Theory. I. Theory of Mobius Functions," *Zeitschrift fur Wahrscheinlichkeitstheorie*, Vol. 2, pp. 340–368, 1964.
5. MacMahon, P. A., *Combinatory Analysis*, Vol. I, University Press, Cambridge, England, 1915.
6. Jacobson, N., *Lectures in Abstract Algebra*, Vol. I, D. van Nostrand Company, Inc., Princeton, New Jersey, 1951.

XXV. Communications and Tracking

A. Maximum Likelihood Symbol Synchronization

J. J. Stiffler

1. Introduction

A symbol is defined as one of a (possibly infinite) set of waveforms that can be transmitted over a communication system. Each symbol lasts for a period of T seconds, at which time a new symbol is transmitted. The optimum receiver for such a symbol sequence depends upon the knowledge of the instants in time in which the transition from one symbol to the next takes place. This information can be obtained from the received sequence itself. This summary discusses the maximum likelihood procedure for doing this. It is assumed that a clock reference has been established at the receiver, so that the length T of a symbol interval is known, but that it is desired to determine the transition instant to within ΔT seconds. The general maximum likelihood symbol synchronization procedure is determined and applied to specific pulse communication systems: PAM, PSK, and FSK. Finally, the maximum-likelihood functions are re-examined with the goal of estimating from them the true synchronization position without the necessity of searching each search position individually. The interference is assumed to be additive white Gaussian noise.

2. The Maximum Likelihood Synchronization Procedure

Let the received signal $y(t)$ be passed through an ideal low-pass filter with bandwidth B, and sampled at the Nyquist rate, $2B$ samples/sec. The output is a sequence of samples, say

$$y_i = x^j_{i-\nu k} + n_i \qquad \begin{array}{l} lL + \nu k \leq i \leq (l+1)L + \nu k \\ j = 1, 2, \cdots, N \\ l = \cdots -1, 0, 1, \cdots \end{array} \tag{1}$$

where the terms n_i represent independent Gaussian random variables with zero mean, and variance $\sigma^2 = N_0 B$, N_0 being the single-sided noise spectral density. The superscript j indicates that the jth of the set of N possible symbols has been received, L is the number of samples per symbol, and $k/L = \Delta T/T = 1/M$ is the ratio of the desired accuracy of the synchronization estimate to the symbol period. The index ν, $\nu = 1, 2, \cdots, M$, corresponds to the M possible synchronization positions.

We begin by observing that the probability that the νth position is indeed the correct one, given the set of observables $Y = \{y_i\}$, is, by Bayes' rule,

$$P(\nu \mid Y) = \frac{p(Y \mid \nu)\, P(\nu)}{p(Y)}. \tag{2}$$

If we assume that the a-priori probability of the νth position is $1/M$ for all ν, or if we ignore $P(\nu)$ because it is unknown, the optimum value of ν which maximizes the a-posteriori probability $P(\nu \mid Y)$ is that for which the inverse probability $P(Y \mid \nu)$ is a maximum. Decisions based on this inverse probability exclusively are known as maximum likelihood decisions.

Since the samples y_i are Gaussian variables

$$\begin{aligned} P(Y \mid \nu) &= \prod_l \sum_{j=1}^{N} \frac{1}{\left[(2\pi)^{1/2}\sigma\right]^L} \\ &\quad \times \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=lL+\nu k}^{(l+1)L+\nu k-1} \left(y_i - x^j_{i-\nu k}\right)^2\right\} P(j) \\ &= K_1 \prod_l \sum_{j=1}^{N} K_2 \\ &\quad \times \exp\left\{\frac{1}{\sigma^2} \sum_{i=lL+\nu k}^{(l+1)L+\nu k-1} y_i x^j_{i-\nu k} - \frac{1}{2}\left(x^j_{i-\nu k}\right)^2\right\} P(j) \end{aligned} \tag{3}$$

where $P(j)$ is the probability of the jth symbol, and where both

$$K_1 = \exp\left\{-\frac{1}{2\sigma^2} \sum_{\text{all } i} y_i^2\right\}$$

and

$$K_2 = \frac{1}{\left[(2\pi)^{1/2}\sigma\right]^L}$$

are independent of ν and can therefore be ignored. Taking the limit as the bandwidth B becomes infinite (so that the sampling interval approaches zero), we find that the maximum likelihood decision is to select as the synchronization position that value of ν for which

$$\prod_l \left(\sum_{j=1}^{N} \exp\left\{\frac{1}{N_0}\int_{lT+\nu\Delta T}^{(l+1)T+\nu\Delta T} 2y(t)x^j(t-\nu\Delta T) - \left[x^j(t-\nu\Delta T)\right]^2 dt\right\} P(j)\right) \tag{4}$$

is a maximum.

3. Specific Examples: PAM, PSK, and FSK

a. Pulse amplitude modulation (PAM). Generally, in the case of PAM, the set of symbols is infinite, with any amplitude between the two amplitudes A and B equally likely. (Typically $A = -B$, or $A = 0$). Then the sum over j in Eq. (4) becomes an integral, and the probability $P(j)$ becomes a probability density. The maximum likelihood decision is to select that value of ν which maximizes

$$\prod_l \int_A^B \exp\left\{\frac{1}{N_0}\int_{lT+\nu\Delta T}^{(l+1)T+\nu\Delta T} \left[2a\, y(t) - a^2\right] dt\right\} \frac{1}{B-A}\, da$$

But this product equals

$$\begin{aligned} &\prod_l \frac{e^{\frac{Tb_l^2}{N_0}}}{B-A} \int_A^B e^{-\frac{T}{N_0}(a-b_l)^2}\, da \\ &\quad = \prod_l \frac{\pi N_0^{1/2}}{T} \frac{e^{\frac{Tb_l^2}{N_0}}}{2(B-A)} \left\{\operatorname{erf}\left[\left(\frac{T}{N_0}\right)^{1/2} \frac{(B-b_l)}{(2)^{1/2}}\right] \right. \\ &\qquad \left. + \operatorname{erf}\left[\left(\frac{T}{N_0}\right)^{1/2} \frac{(b_l - A)}{(2)^{1/2}}\right]\right\}, \end{aligned} \tag{5}$$

where

$$b_l \equiv \frac{1}{T}\int_{lT+\nu\Delta T}^{(l+1)T+\nu\Delta T} y(t)\, dt,$$

and where

$$\operatorname{erf}\frac{x}{2^{1/2}} = -\operatorname{erf}\left(\frac{-x}{2^{1/2}}\right) = \frac{1}{(2\pi)^{1/2}}\int_{-x}^{x} e^{-\frac{\xi^2}{2}}\, d\xi$$

Taking the logarithm of Eq. (5) and ignoring those terms in the resulting expression which are independent of ν, we find the decision should be based on maximizing the sum

$$\sum_l \frac{T}{N_0} b_l^2 + \log\left\{\operatorname{erf}\left[\left(\frac{T}{N_0}\right)^{1/2}(B - b_l)\right] + \operatorname{erf}\left[\left(\frac{T}{N_0}\right)^{1/2}(b_l - A)\right]\right\} \tag{6}$$

This expression, while a function of the easily determined variable b_l, is still rather formidable for practical application. An approximation, which is increasingly satisfactory as the signal-to-noise ratio increases, so that the quantities

$$\left(\frac{T}{N_0}\right)^{1/2}(B - b_l) \text{ and } \left(\frac{T}{N_0}\right)^{1/2}(b_l - A)$$

are almost always large, is to ignore all but the first term. In this case the decision is based exclusively on the term

$$\sum_l Tb_l^2 = \sum_l \frac{1}{T}\left[\int_{lT+\nu\Delta T}^{(l+1)T+\nu\Delta T} y(t)\, dt\right]^2 \tag{7}$$

With this simplification, we can estimate the number of symbols which must be received before a reliable

decision can be made. First we observe that the distribution of the variable

$$z_\nu = \sum_l \frac{1}{T}\left[\int_{lT+\nu\Delta T}^{(l+1)T+\nu\Delta T} y(t)\,dt\right]^2 \tag{8}$$

is approximately normal by the Central Limit Theorem when the number of terms in the summation is large. Furthermore, a measure of the length of time needed to make a decision is afforded by considering the number of symbols needed before the νth position can be distinguished with high reliability from the μth position, for arbitrary μ and ν. Since z_μ and z_ν are both approximately Gaussianly distributed, it follows that

$$Pr\{z_\nu > z_\mu\} = Pr\{y > 0\}$$

where $y = z_\nu - z_\mu$ also has a Gaussian distribution. Thus

$$Pr\{y < 0\} = \frac{1}{(2\pi)^{1/2}\sigma}\int_0^\infty e^{-\frac{(y+m)^2}{2\sigma^2}}\,dy = \frac{1}{2}\left\{1 - \operatorname{erf}\frac{m}{2^{1/2}\sigma}\right\} \tag{9}$$

where m $= E(y)$, and σ^2 is the variance of y. But

$$m = E(z_\nu - z_\mu) = \frac{1}{T}\sum_{l=1}^{S}\left\{E\left[\int_{lT+\nu\Delta T}^{(l+1)T+\nu\Delta T} y(t)\,dt\right]^2 - E\left[\int_{lT+\nu\Delta T}^{(l+1)T+\mu\Delta T} y(t)\,dt\right]^2\right\}$$

$$= 2S(\mu-\nu)\frac{\Delta T}{T}\,\frac{T-(\mu+\nu)\Delta T}{T}\,\mu_2 T \tag{10a}$$

and

$$\sigma^2 = E(z_\nu^2) - 2E(z_\nu z_\mu) + E(z_\mu^2) - m^2 \tag{10b}$$

$$= 2S(\mu-\nu)\frac{\Delta T}{T}\left\{\left(1-(\mu-\nu)\frac{\Delta T}{T}\right)N_0^2 + 2\left[1-(2\nu+\mu)\frac{\Delta T}{T}+3\mu\nu\frac{\Delta T}{T}\right]^2\mu_2 T N_0 + 2\left[1-(\mu+\nu)\frac{\Delta T}{T}\right]^2(\mu-\nu)\frac{\Delta T}{T}\,\mu_4 T^2\right\}, \quad \mu > \nu$$

where μ_2 and μ_4 are the second and fourth central moments, respectively, of the signal amplitude:

$$\mu_2 = \frac{1}{B-A}\int_A^B [a - E(a)]^2\,da = \frac{(B-A)^2}{12}$$

$$\mu_4 = \frac{1}{B-A}\int_A^B [a - E(a)]^4\,da = \frac{(B-A)^4}{60}; \tag{11}$$

The number of symbols observed, S, is assumed to be sufficiently large so that the effect of the partial symbol intervals at the beginning and end of the total observation interval can be ignored. From Eq. (9), the probability that one of the nearest ($\mu = 1$) sync positions results in an observable larger than that of the true ($\nu = 0$) position after S symbols have been observed is

$$Pr\{z_1 > z_0\} = \frac{1}{2}\left\{1 - \operatorname{erf}\frac{m}{2^{1/2}\sigma}\right\} \tag{12}$$

where m^2/σ^2, the output signal-to-noise ratio, is given by

$$\frac{m^2}{\sigma^2} = \frac{2S\Delta T\,(T-\Delta T)\mu_2^2}{(N_0^2 + 2\mu_2 N_0 T + 2\Delta T(T-\Delta T)\mu_4)} \tag{13}$$

$$= \frac{S\Delta T(T-\Delta T)}{T^2}\,\frac{R^2}{2\left(1+R+\dfrac{6\Delta T(T-\Delta T)}{5T^2}R^2\right)}$$

and where $R = 2\mu_2 T/N_0$ may be regarded as the input signal-to-noise ratio. For R large compared to one but small compared to $5T^2/6\Delta T\,(T-\Delta T)$ we have

$$\frac{m^2}{\sigma^2} \approx S\,\frac{\Delta T}{T}\left(1-\frac{\Delta T}{T}\right)\frac{R}{2} \tag{14}$$

and when R is small compared to unity

$$\frac{m^2}{\sigma^2} \approx S\,\frac{\Delta T}{T}\left(1-\frac{\Delta T}{T}\right)\frac{R^2}{2}. \tag{15}$$

It is observed that, as R approaches infinity, the output signal-to-noise ratio becomes $5S/6$, which suggests that synchronization errors are possible even in the absence of noise. This happens because the end effect becomes less and less ignorable as $R \to \infty$, and since y is not Gaussian, but only approaches the Gaussian distribution for large values of S. To verify that a correct decision can be made after only one symbol interval in the absence of noise, it is only necessary to observe that the quantity

$$\int_{\tau_1}^{\tau_2} y^2(t)\,dt - \frac{1}{\tau-\tau_1}\left[\int_{\tau_1}^{\tau} y(t)\,dt\right]^2 - \frac{1}{\tau_2-\tau}\left[\int_{\tau}^{\tau_2} y(t)\,dt\right]^2, \quad \tau_1 < \tau < \tau_2 \tag{16}$$

is positive unless the signal $y(t)$ is constant in each of the intervals of integration. Since the first of these terms is constant regardless of the value of τ, and since the remainder of the expression is the quantity used to make a maximum likelihood estimate of the sync position, it follows that the correct position will always correspond to a maximum value of the likelihood function. This maximum will be unique unless $y(t)$ is constant over the total interval $\tau_1 < t < \tau_2$, which can happen only if two

successive symbols are equal (an event of zero probability under the assumptions here) or if the region of integration is totally contained in one symbol interval. Consequently, in the noiseless case, a decision can be made with zero probability of error so long as the interval of observation contains infinitesimally more than one symbol interval.

b. Phase shift keying (PSK). The system we consider here is one in which any one of the N waveforms $x_i(t) = A \sin(\omega_0 + \theta_i)$ is equally likely to be transmitted over any given T second symbol time interval. Thus, from Eq. (4), we conclude that the most likely synchronization position corresponds to that value of ν maximizing the expression

$$\prod_l \sum_{j=1}^{N} \exp\left\{\frac{2A}{N_0}\int_{lT+\nu\Delta T}^{(l+1)T+\nu\Delta T} y(t)\sin(\omega_0 t + \theta_j)\right\} = \prod_l \sum_{j=1}^{N} \exp\left\{z_{\nu l}\cos\theta_j + w_{\nu l}\sin\theta_j\right\} \quad (17)$$

where $z_{\nu l}$ is given by

$$z_{\nu l} = \frac{2A}{N_0}\int_{lT+\nu\Delta T}^{(l+1)T+\nu\Delta T} y(t)\sin\omega_0 t\, dt$$

and $w_{\nu l}$ is its quadrature component

$$w_{\nu l} = \frac{2A}{N_0}\int_{lT+\nu\Delta T}^{(l+1)T+\nu\Delta T} y(t)\cos\omega_0 t\, dt$$

Often, in situations of this sort, one of the exponents can be expected to be significantly greater than all the rest so that the summation over j may be approximated by this dominant term. In this case, the logarithm of expression (17) becomes

$$\sum_l \max_j \left\{z_{\nu l}\cos\theta_j + w_{\nu l}\sin\theta_j\right\}, \quad (18)$$

a somewhat easier expression to evaluate. To estimate the acceptability of this approximation, suppose that the signal $y(t) = A\sin(\omega_0 t + \theta_0)$ was received. Then

$$z_{\nu l}\cos\theta_j + w_{\nu l}\sin\theta_j = \frac{A^2 T}{N_0}\cos(\theta_0 - \theta_j) \quad (19)$$

so that if the signal-to-noise ratio is large and if the possible phase angles θ_j are well separated, the approximation will be satisfactory. If either of these conditions does not hold, it may be necessary to use the original expression (17).

In the more interesting case, when the signal-to-noise ratio is small, the function (17) can be approximated by a much simpler expression by retaining only the first three terms in this power series expansion of the exponential:

$$\prod_l \sum_{j=1}^{q} \exp\left\{z_{\nu l}\cos\theta_j + w_{\nu l}\sin\theta_j\right\} = \prod_l \sum_{j=1}^{q}\left(1 + z_{\nu l}\cos\theta_j + w_{\nu l}\sin\theta_j + \frac{1}{2}z_{\nu l}^2\cos^2\theta_j + z_{\nu l}w_{\nu l}\sin\theta_j\cos\theta_j + \frac{1}{2}w_{\nu l}^2\sin^2\theta_j\right).$$

Generally, the q phases will be equally distributed in the interval 0 to 2π, i.e., typically $\theta_j = 2\pi/q\ j$, $j = 1, 2, \cdots, q$. In this case

$$\sum_{j=1}^{q}\cos\theta_j = \sum_{j=1}^{q}\sin\theta_j = 0 \quad (21)$$

so that the above expression becomes

$$\prod_l q\left(1 + \frac{1+\delta_{q2}}{4}z_{\nu l}^2 + \frac{1+\delta_{q2}}{4}w_{\nu l}^2\right) \approx K\sum_l\left(z_{\nu l}^2 + w_{\nu l}^2\right) \quad (22)$$

where $\delta_{q2} = 1$ when $q = 2$ and zero otherwise, and where K is independent of ν.

This last approximation can be analyzed to determine the probability of an incorrect decision entirely analogously to the preceeding analysis for PAM. Letting $u_\nu = \sum_l\left(z_{\nu l}^2 + w_{\nu l}^2\right)$ and $v = u_\nu - u_\mu$, we find that

$$Pr\{v < 0\} = \frac{1}{2}\left[1 - \operatorname{erf}\frac{m}{(2\sigma)^{1/2}}\right] \quad (23)$$

where

$$m = E(v) = \sum_l\left[E(z_{\nu l}^2) + E(w_{\nu l}^2) - E(z_{\mu l}^2) - E(w_{\mu l}^2)\right] = 2S\frac{A^4T^2}{N_0^2}(\mu - \nu)\frac{\Delta T}{T}\left(1 - (\mu + \nu)\frac{\Delta T}{T}\right) \quad (24a)$$

and

$$\sigma^2 = E(v^2) - E^2(v)\,. \quad (24b)$$

$$= \frac{4A^4T^2}{N_0^2}S\left\{\left(\frac{1+\delta_{q2}}{2}\right)(\mu-\nu)^2\left(\frac{\Delta T}{T}\right)^2\left(1 - (\mu+\nu)\frac{\Delta T}{T}\right)^2\frac{A^4T^2}{N_0^2} + 2(\mu-\nu)\frac{\Delta T}{T}\left[1 - (\mu+2\nu)\frac{\Delta T}{T} + 3\mu\nu\left(\frac{\Delta T}{T}\right)^2\right]\frac{A^2T}{N_0} + 2(2-\delta_{q2})(\mu-\nu)\frac{\Delta T}{T}\left(1 - (\mu-\nu)\frac{\Delta T}{T}\right)\right\}$$

The "output" signal-to-noise ratio, in this case, is

$$\frac{m^2}{\sigma^2} = \frac{SR^2(\mu-\nu)\frac{\Delta T}{T}\left(1-(\mu+\nu)\frac{\Delta T}{T}\right)^2}{\left(\frac{1+\delta_{q2}}{2}\right)R^2(\mu-\nu)\frac{\Delta T}{T}\left(1-(\mu+\nu)\frac{\Delta T}{T}\right)^2 + 2\left(1-(\mu+2\nu)\frac{\Delta T}{T}+3\mu\nu\left(\frac{\Delta T}{T}\right)^2\right)R + 2(2-\delta_{q2})\left(1-(\mu-\nu)\frac{\Delta T}{T}\right)} \tag{25}$$

where $R = A^2T/N_0$ is the input signal-to-noise ratio. (It has been assumed in this derivation that ω is some multiple of $2\pi/\Delta T$. Since phase coherence must be presumed in order that the system be operative as described, it is not necessary to resolve the synchronization position to finer than $2\pi/\omega$ seconds.)

c. Frequency shift keying (FSK). Here the set of signals which can be transmitted is comprised of

$$x_i(t) = A\sin\omega_i t \qquad i = 1, 2, \cdots, M. \tag{26}$$

The maximum likelihood synchronization estimator for FSK becomes, from Eq. (4),

$$\prod_l \sum_{j=1}^{M} \exp\left\{\frac{2R}{N_0}\int_{lT+\nu\Delta T}^{(l+1)T+\nu\Delta T} y(t)\sin\omega_j(t-\Delta\nu T)\,dt\right\} P(j) \tag{27}$$

in which the terms independent of ν have been omitted.

Without further information about the set of frequencies ω_i, there is little that can be done to simplify this expression. Consequently, we shall stipulate the rather common situation in which $\omega_i = 2\pi\, k_i/\Delta T$ for some integer k_i. Then

$$\begin{aligned}
&\int_{lT+\nu\Delta T}^{(l+1)T+\nu\Delta T} y(t)\sin\omega_j(t-\nu\Delta T)\,dt \\
&\quad = A\int_{lT+\nu\Delta T}^{(l+1)T}\sin\omega_r t\sin\omega_j(t-\nu\Delta T)\,dt \\
&\qquad + \int_{(l+1)T}^{(l+1)T+\nu\Delta T}\sin\omega_s t\sin\omega_j(t-\nu\Delta T)\,dt \\
&\qquad + \int_{lT+\nu\Delta T}^{(l+1)T+\nu\Delta T} n(t)\sin\omega_j(t-\nu\Delta T)\,dt \\
&\quad = \frac{A}{2}\cos\omega_j\nu\Delta T\left[\delta_{rj}(T-\nu\Delta T) + \delta_{sj}(\nu\Delta T)\right] \\
&\qquad + \int_{lT+\nu\Delta T}^{(l+1)T+\nu\Delta T} n(t)\sin\omega_j(t-\nu\Delta T)\,dt
\end{aligned} \tag{28}$$

where δ_{ij} is the Kronecker delta. With probability $M-1/M$, one of the M correlations will have an output with the expected value $A(T-\nu\Delta T)/2$, one an output with the expected value $A\nu\Delta T/2$, and all the rest with the expected value zero. (This, of course, is under the assumption that each one of the M signals has an equal probability of occurring over any interval, and that the signals over successive intervals are statistically independent.) The other $1/M$th of the time will result in $M-1$ correlations with expected outputs equal to zero and one with the expected output $AT/2$.

This observation suggests the method, mentioned earlier, of selecting the largest of the outputs of the correlations for each value of ν and using this as the statistic on which to base the decision. For values of ν near zero, the largest of these outputs will nearly always be that for which $j=r$ in Eq. (28). On the other hand, as ν approaches $N/2$, the signal contribution to this largest output decreases to approximately half its maximum value, and it will much more frequently happen that one of the zero mean outputs will exceed both of the outputs having a non-zero mean. Moreover, because the mean of the output corresponding to $j=s$ in Eq. (28) will be nearly as great as that for which $j=r$, the former will also frequently give rise to a larger output than the latter. The effect of this is to increase the mean of the statistic beyond the value $A(T-\nu\Delta T)/2$ for ν greater than zero. Generally, such a positive bias will not alter the fact that the positions most difficult to distinguish from the correct sync position ($\nu=0$) are those nearest to it ($\nu=1$, and $\nu=N-1$). These nearest positions will not be significantly biased by the preceding considerations, so that a measure of the effectiveness of the procedure is afforded by the number of symbol intervals which must be observed before the statistic having the mean $AT/2$ can be reliably distinguished from that having the mean $\frac{A}{T}(T-\Delta T)\left(\frac{M-1}{M}\right)+\frac{AT}{2}\frac{1}{M}$.

Let

$$z=\sum_l\left\{\max_j\int_{lT}^{(l+1)T}y(t)\sin\omega_j t\,dt - \max_j\int_{lT+\nu\Delta T}^{(l+1)T+\nu\Delta T}y(t)\sin\omega_j(t-\Delta T)\,dt.\right\} \quad (29)$$

The probability that the position $\nu=1$ results in a greater output than the $\nu=0$ position is, again,

$$Pr\left\{z<0\right\}=\frac{1}{2}\left[1-\operatorname{erf}\frac{m}{(2\sigma)^{1/2}}\right] \quad (30)$$

where

$$m=E(z)=S\frac{M-1}{M}\frac{A\Delta T}{2} \quad (31)$$

and

$$\sigma^2=E(z^2)-E^2(z)=S\left[\frac{A^2\Delta T^2}{4}\frac{M-1}{M^2}+\frac{M-1}{M}\frac{N_0\Delta T}{2}\right]$$

where S is the number of symbols observed. The effective output signal-to-noise ratio is

$$\frac{m^2}{\sigma^2}=\frac{SR}{\frac{R}{M-1}+2\left(\frac{T}{\Delta T}\right)\left[2\left(\frac{M}{M-1}\right)\right]}\approx\frac{SR}{2}\left(\frac{\Delta T}{T}\right) \quad (32)$$

where $R=A^2T/N_0$, and the approximation on the right is valid for large values of M.

4. Serial Synchronization and Estimation

We recall from the previous section that the expected value of the difference y [see Eq. (10)] between the outputs of the νth and μth PAM synchronizers is

$$m=2S(\mu-\nu)\frac{\Delta T}{T}\left(\frac{T-(\mu+\nu)\Delta T}{T}\right)\mu_2 T, \quad (33)$$

where the true sync position is represented by $\nu=0$. Suppose we let $\mu-\nu=k$ for some integer k. Then

$$m=\begin{cases}2S\frac{k}{N}\left(1-\frac{k}{N}-\frac{2\nu}{N}\right)\mu_2 T, & 0\le\nu\le N-k;\\ -2S\frac{N-k}{N}\left(1+\frac{N-k}{N}-\frac{2\nu}{N}\right)\mu_2 T, & N-k\le\nu\le N,\end{cases} \quad (34)$$

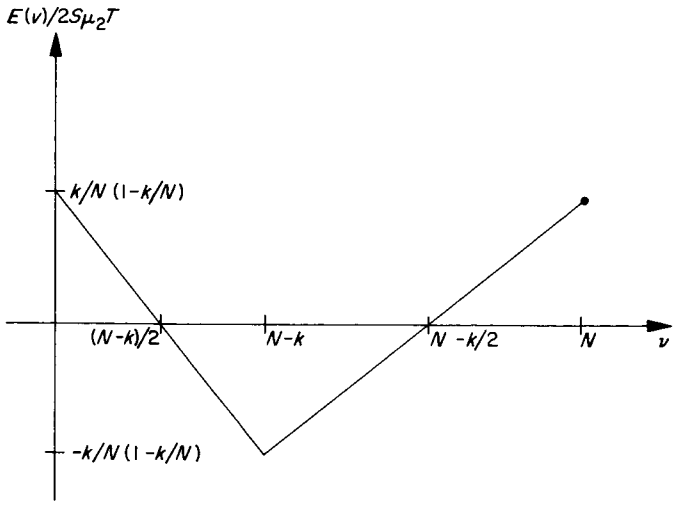


Fig. 1. The expected value of the synchronizer difference versus the sync position

where $T = N\Delta T$. This function is illustrated in Fig. 1. A specific value of the parameter $E(y)/2S\mu_2T$ corresponds to two possible values of ν, either

$$\nu = \nu_0 \text{ or } \nu = N - \frac{k}{N-k}\nu_0.$$

This suggests the use of the normalized observable $y/2S\mu_2T$ as an estimator of the value of ν. If this estimate of ν is used and the observed positions are changed accordingly, one of the correlators will be operating at the position $\nu = 0$ (if a correct estimate is made) while the other will be at $\nu = \pm k$. One of the correlators will consequently consistently have a large output whereas the output of the other will be small, and the ambiguity is resolved.

The reliability with which the correct sync position can be identified is proportional to the slope of the curve of Fig. 1. If the a-priori probability of a particular value of ν is independent of ν, then there seems little reason for the curve of Fig. 1 to be asymmetrical. In general, therefore, k will be equal to $N/2$. The estimator of ν, after S symbols have been observed, is

$$\hat{\nu} = \left(\frac{1}{2} - \frac{y}{S\mu_2T}\right)\frac{N}{2}. \tag{35}$$

We observe that[1]

$$m_\nu = E(\hat{\nu}) = \frac{N}{4} - \frac{N}{2S\mu_2T}E(y) = \nu \tag{36a}$$

since, from Eq. (10), $E(y) = S\mu_2T\left(\frac{1}{2} - \frac{2\nu}{N}\right)$. Further

$$\sigma_{\hat{\nu}}^2 = E(\hat{\nu}^2) - E^2(\hat{\nu}) = \frac{N^2}{16} - \frac{N^2}{4}\frac{E(y)}{S\mu_2T} + \frac{N^2E(y^2)}{4(S\mu_2T)^2} - \nu^2$$

$$= \frac{N^2}{S}\left[2R^2 + \left(\frac{1}{2} - \frac{3\nu}{2N} + \frac{3\nu^2}{N^2}\right)\frac{1}{R} + \frac{3}{5}\left(\frac{1}{2} - \frac{2\nu}{N}\right)^2\right] \tag{36b}$$

An estimate of the number of symbols needed can be obtained by requiring the variance of $\hat{\nu}$ to be small compared to 1 so that the probability of being in error by one or more in the estimate of ν will be sufficiently small. We then require that

$$S >> N^2\left[\frac{1}{2R^2} + \left(\frac{1}{2} - \frac{3\nu}{2N} + \frac{3\nu^2}{N^2}\right)\frac{1}{R} + \frac{3}{5}\left(\frac{1}{2} + \frac{2\nu}{N}\right)^2\right] \tag{37}$$

The comparable condition in the case of the parallel search procedure of Sect. 3 was that

$$\frac{m^2}{\sigma^2} >> 1$$

or, from Eq. (13), that

$$S >> 2N\left(\frac{N}{N-1}\right)\left(\frac{1}{R^2} + \frac{1}{R} + \frac{6}{5}\frac{N-1}{N^2}\right) \tag{38}$$

The estimation procedure takes on the order of N times as many symbols to make a decision of the same reliability, using, however, only two rather than N correlators in the process.

This estimation approach compares favorably with the alternative serial search procedure in which only one position is to be investigated at any one time. In this case, the mean of Eq. (10) remains the same, but the variance becomes

$$\sigma^2 = 25\left\{\frac{N_0^2}{2} + 2\left[1 - (\mu+\nu)\frac{\Delta T}{T} + (\mu^2+\nu^2)\left(\frac{\Delta T}{T}\right)^2\right]\mu_2N_0T + 2\left[\left(1 - \frac{\nu\Delta T}{T}\right)^2\left(\frac{\nu\Delta T}{T}\right)^2 + \left(1 - \frac{\mu\Delta T}{T}\right)^2 \times \left(\frac{\mu\Delta T}{T}\right)^2\right]\mu_4T^2\right\} \tag{39}$$

[1]We are assuming that $0 \leq \nu \leq N/2$ and that the correct one of the two estimates of ν is made. If not, the other correlator, differing by $N/2$ positions from ν, will be moved to the correct position, so the net result is the same.

yielding the condition that (when $\nu = 0$)

$$S >> \frac{\frac{1}{R^2} + \left(1 - \frac{\mu}{N} + \frac{\mu^2}{N^2}\right)\frac{1}{R} + \frac{6}{5}\left[\left(1 - \frac{\mu}{N}\right)^2\left(\frac{\mu}{N}\right)^2\right]}{\left(\frac{\mu}{N}\right)^2\left(1 - \frac{\mu}{N}\right)^2} \tag{40}$$

This is the number of symbols needed to assure that the μth position output is, with a satisfactory probability, smaller than that of the correct ($\nu = 0$) position. If each position were to be observed for a fixed length of time, S would have to be chosen so that even the most difficult to resolve positions ($\mu = 1$ and $N - 1$) could be reliably rejected. This would require that (for large N)

$$S >> N^2\left\{\frac{1}{R^2} + \frac{1}{R}\right\} \tag{41}$$

and the total number of symbols needed is proportional to N^3.

A sequential search, on the other hand, would allow each position to be rejected when the probability of its being the correct position had dropped below a certain value. In this case, the more readily distinguishable positions would be rejected more quickly than those corresponding to smaller values of μ. A measure of the total number of symbols needed in this case would, from Eq. (5), be on the order of

$$\sum_{\mu=1}^{N-1}\left[\frac{1}{R^2}\left(\frac{N^4}{\mu^2(N-\mu)^2}\right) + \frac{1}{R}\left(\frac{N^2(N^2-\mu N+\mu^2)}{\mu^2(N-\mu)^2}\right) + \frac{6}{5}\right]$$

$$\approx \frac{2N^2}{R^2}\left\{\frac{5}{3} + \frac{2\log N}{N}\right\} + \frac{2N^2}{R}\left\{\frac{5}{3} + \frac{\log N}{N}\right\} + \frac{6N}{5}. \tag{42}$$

A sequential approach can also be used with two correlators. Suppose the two correlators investigate sync positions mutually separated by $N/2$ positions. A sequential decision is made as to which of the two correlators is the larger and, when a decision is made, both correlators are shifted one unit to the right or left, depending upon the result. It is easily seen (see Fig. 1) that the shifting will continue until one of the correlators is centered at the position $\nu = N/4$ while the other is at $\mu = 3N/4$.

Letting $\mu = \nu + N/2$ in Eq. (10), we obtain

$$\frac{m^2}{\sigma^2} = \frac{S\left(\frac{1}{2} - \frac{2\nu}{N}\right)^2 R^2}{4\left\{\frac{N_0^2}{2} + \left(\frac{1}{2} - \frac{3\nu}{2N} + \frac{3\nu^2}{N^2}\right)R + \frac{3}{5}\left(\frac{1}{2} - \frac{2\nu}{N}\right)^2 R^2\right\}} \tag{43}$$

Again, a measure of the number of symbols needed for a decision is obtained by demanding that m^2/σ^2 be sufficiently large to enable a decision to be made with high reliability at each stage. (It is not apparent that a superior performance would not result when less precise decisions were made, realizing that some mistakes would be made from time-to-time. Since only unit shifts are made after each decision, those mistakes would not be very costly, and the net search time might be decreased.) The total number of symbols needed for this approach is on the order of

$$2\sum_{\nu=0}^{\frac{N}{4}-1} \frac{4\left[\frac{1}{2R^2} + \left(\frac{1}{2} - \frac{3\nu}{2N} + \frac{3\nu^2}{N^2}\right)\frac{1}{R} + \frac{3}{5}\left(\frac{1}{2} - \frac{2\nu}{N}\right)^2\right]}{\left(\frac{1}{2} - \frac{2\nu}{N}\right)^2}$$

$$\approx \frac{5}{3}\frac{N^2}{R^2} + \left(\frac{25}{24}N + \frac{3}{2}\right)\frac{N}{R} + \frac{6}{5}N \tag{44}$$

It is clear, upon comparing the discussions of PAM and PSK synchronization in the previous section, that the results just derived for PAM are directly applicable to PSK with a suitable change in the constants involved.

The FSK situation, however, is somewhat different, due to the biases introduced when the correlator with the maximum output is not that corresponding to the maximum mean (see Sect. 3). This bias, unless carefully taken into account, makes accurate synchronization estimation impossible for FSK. The method of shifting two correlators, separated by $N/2$ positions, until they both produce the same output will still be effective, however,

since the biases to the correlators will be the same when one is located at $\nu = N/4$ and the other at $\mu = 3N/4$. The asymmetry produced by the biases at other positions will nevertheless serve to increase the amount of time needed to reach the correct position.

A comparison between this method and the ordinary serial search can be made by disregarding the effect of the biases. In this case, letting

$$z = \sum_l \left[\max_j \int_{lT+\nu\Delta T}^{(l+1)T+\nu\Delta T} y(t) \sin \omega_j (t-\nu\Delta T)\, dt - \max_j \int_{lT+\mu\Delta T}^{(l+1)T+\nu\Delta T} y(t) \sin \omega_j (t-\mu\Delta T)\, dt \right], \quad (45)$$

we find that

$$m = E(z) = s\frac{M-1}{M}\frac{AT}{2}\left(1-\frac{(\mu+\nu)}{N}\right), \quad (46)$$

$$\text{and } \sigma^2 = E(z^2) - E^2(z) = S\left[\frac{A^2T^2}{4}\frac{M-1}{M^2}\left(1-\frac{(\mu+\nu)}{N}\right)^2 + \frac{N_0T}{2}\left(\frac{M-1}{M}\right)\left(1-\frac{(\mu-\nu)}{N}\right)\right].$$

Setting $\mu = \nu + N/2$ yields the condition that

$$S >> \frac{R\left(\frac{1}{2}-\frac{2\nu}{N}\right)+2M}{(M-1)R\left(\frac{1}{2}-\frac{2\nu}{N}\right)} \approx \frac{1}{R\left(\frac{1}{2}-\frac{2\nu}{N}\right)} \quad (47)$$

if a correct decision is to be made at the ν-th position. Scanning over ν as before, we find that

$$\frac{4}{R}\sum_{\nu=0}^{\frac{N}{4}-1}\frac{1}{\left(\frac{1}{2}-\frac{2\nu}{N}\right)} \approx \frac{2N}{R}\log N \quad (48)$$

provides a measure of the number of symbols needed to synchronize using this method.

If the positions are investigated serially, and z is defined as in Eq. (45), the expected value of z remains as given in Eq. (46) while its variance becomes

$$\sigma^2 = \frac{M-1}{M^2}\frac{A^2T^2}{4}\left\{1-\frac{2\mu}{N}+\frac{(\mu^2+\nu^2)}{N^2}\right\}+\frac{N_0T}{2}. \quad (49)$$

The condition that $m^2/\sigma^2 >> 1$, when $\nu = 0$, is equivalent to the requirement that

$$S >> 1 + \frac{2M^2}{M-1}\frac{1}{R\left(1-\frac{\mu}{N}\right)^2} \approx \frac{2}{R\left(1-\frac{\mu}{N}\right)^2} \quad (50)$$

The number of symbols needed to complete the search using this method is on the order of

$$\frac{2}{R}\sum_{\mu=1}^{N-1}\frac{1}{\left(1-\frac{\mu}{N}\right)^2} \approx \frac{10}{3}\frac{N^2}{R}. \quad (51)$$

B. Proper Waveforms for Subcarriers in Digital Telemetry

R. Goldstein

It is shown that the proper waveform to use for the subcarrier in phase-coherent digital telemetry in order to maximize the output power relative to the carrier power is a square wave. A gain of more than 1 db over a cosine subcarrier system is found.

1. The System

This note is an investigation of the proper waveform to use for the subcarrier in a conventional, phase-coherent digital communications link. Fig. 2 is a block diagram of the system under investigation. We consider here only the output power relative to the carrier power. Other design factors such as bandwidths and spectra are not considered.

The transmitted signal, $y(t)$, has the form:

$$y(t) = 2^{1/2}\cos\left(\omega t + x(t)\right).$$

The power P_c left in the carrier for the loop to track is

$$P_c = \left(\frac{2^{1/2}}{T}\int_0^T y(t)\cos\omega t\, dt\right)^2 + \left(\frac{2^{1/2}}{T}\int_0^T y(t)\sin\omega t\, dt\right)^2,$$

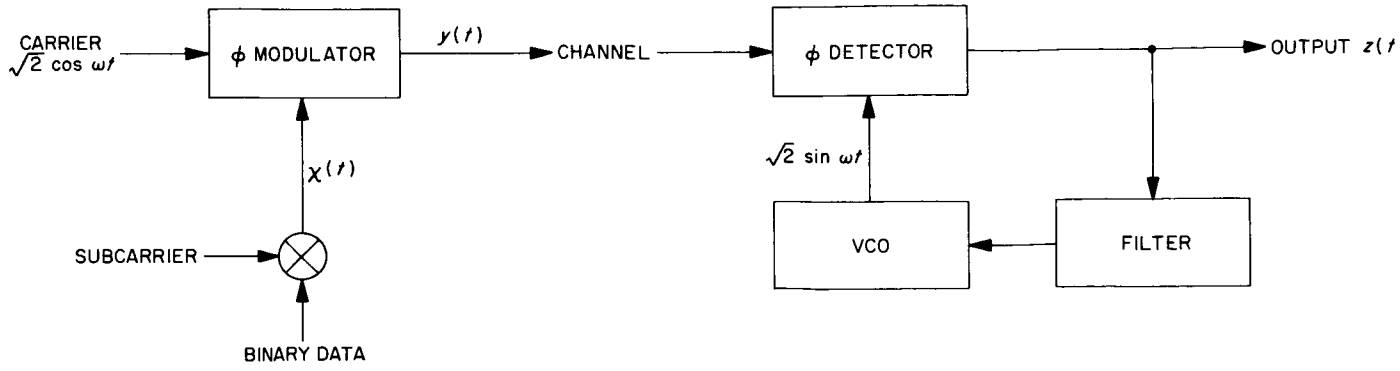


Fig. 2. Block diagram of phase-coherent system

where T is the symbol duration. Under the assumptions that $x(t)$ is independent of $\cos\omega t$ and $\sin\omega t$, and that $x(t)$ is a zero-mean waveform, this expression reduces to

$$P_c = \left(\frac{1}{T}\int_0^T \cos x(t)\,dt\right)^2 \tag{1}$$

The output $z(t)$ of the receiver phase-detector is

$$z(t) = 2\sin\omega t\cos(\omega t + x(t)) = \sin x(t),$$

if the double frequency terms are removed. The primary consideration for the subsequent detection and decoding is the power P_M of $z(t)$:

$$P_M = \frac{1}{T}\int_0^T z^2(t)\,dt = \frac{1}{T}\int_0^T \sin^2 x(t)\,dt \tag{2}$$

2. Lagrange Technique

Eqs. (1) and (2) give the carrier power and the usable modulation power, respectively. The problem can now be stated: Given a value for the carrier power P_c, what waveform $x(t)$ will maximize the modulation power, P_M? The Lagrangian multiplier technique is convenient:

$$P_M = \int \sin^2 x(t)\,dt + \lambda\left(\int \cos x(t)\,dt\right)^2$$

$$\delta(P_M) = \int 2\sin x(t)\cos x(t)\,\delta(x)\,dt$$

$$- 2\lambda\int \cos x(t)\,dt\int \sin x(t)\,\delta(x)\,dt = 0;$$

$$\int \sin x(t)\,\delta(x)\left[2\cos x(t) - 2\lambda\int \cos x(u)\,du\right]dt = 0.$$

Since the coefficient of $\delta(x)$ must be zero, we conclude that

$$\cos x(t) = \text{constant (namely, } \lambda\int \cos x(u)\,du).$$

$$\therefore x(t) = \text{square wave}. \tag{3}$$

We have reached the desired conclusion. Thus the optimum waveform is a square wave.

3. Gain in Available Power

Example (1), the cosine wave case. Suppose $x(t) = k\cos t$. Then, according to Eqs. (1) and (2)

$$P_c = \left(\frac{1}{\pi}\int_0^\pi \cos(k\cos t)\,dt\right)^2$$

$$= J_0^2(k),$$

and

$$P_M = \frac{1}{\pi}\int_0^\pi \sin^2(k\cos t)\,dt$$

$$= \frac{1}{2} - \frac{1}{2}J_0(2k).$$

Assume that a high data rate system is under consideration, and that it is desired to have 8% of the total transmitted power in the carrier. Then $J_0^2(k) = 0.08$, $k = 1.9$, and

$$P_M = 0.701.$$

Example 2, the square wave case. Suppose that $x(t) = \pm a$. Then

$$P_c = \left(\frac{1}{T}\int_0^T \cos(\pm a)\,dt\right)^2$$

$$= \cos^2 a,$$

and

$$P_M = \frac{1}{T}\int_0^T \sin^2(\pm a)\, dt$$
$$= \sin^2 a\,.$$

Notice that the total power P_T in this case is always unity,

$$P_T = P_c + P_M = \cos^2 a + \sin^2 a = 1\,.$$

As the transmitted power equals 1, there is no loss.

Assume, again, that 8% of the power is left in the carrier, then

$$P_M = 1 - 0.08$$
$$= 0.920$$

A comparison with the cosine wave case shows an improvement of

$$0.920/0.701 = 1.31 = 1.18 \text{ db}$$

XXVI. ASTROMETRICS

A. Radar Signals in Model Atmospheres

P. E. Reichley and D. O. Muhleman

1. Introduction

A problem arising from the need for high-precision angle tracking of a radar target is to find the high-precision refraction corrections to the radar signal. In order to make these corrections, an accurate model of the atmosphere is needed. This article discusses the mathematical techniques needed to construct such a model which requires the magnitude of the doppler shift of a radar signal as it passes through the atmosphere. We shall assume throughout that the atmospheric model has spherical symmetry.

2. Preliminaries

We shall construct our model using an index of refraction function. We shall not distinguish between the tropospheric index and the ionospheric index (the latter being a function of frequency) but, instead, we shall consider the model as having only one index which takes on the character of interest.

The basic equation needed in our construction is the ray equation (*SPS* 37-29, Vol. IV, p. 231)

$$\frac{d\theta}{dr} = \frac{b}{r\left[r^2n^2(r) - b^2\right]^{1/2}}, \tag{1}$$

where $n(r)$ is the index of refraction, assumed a function of r only, b is the *impact parameter* given by Bouger's Theorem (*SPS* 37-29, Vol. IV, p. 231),

$$b = rn(r)\sin i, \tag{2}$$

and r, θ, and i are shown in Fig. 1. The angle i is defined as the angle between the ray tangent vector and the radius vector at any point on the ray path. We see that b is the distance from the asymptote of the ray to the center of the coordinate system. We have chosen the ray equation in polar coordinates since the ray path lies in a plane passing through the center of the planet (Earth) in a spherically symmetric model.

We shall also assume that we have available a high-precision ephemeris of the object which we are tracking. This ephemeris provides the distance to the object from the center of the Earth as well as the angle between this vector and the station vector. Then from Eq. (1) we have the integral equation

$$\theta(t) = \int_{r_0}^{r(t)} \frac{b(t)\,dr}{r\left[r^2n^2(r) - b^2(t)\right]^{1/2}}, \tag{3}$$

where at time t, $\theta(t)$ and $r(t)$ are given by the ephemeris, r_0 is the station radius vector, and the function $b(t)$ has to be solved for (see Fig. 1).

Using Eq. (1), one easily shows (*SPS* 37-29, Vol. IV, p. 237) that the range $\rho(t)$ to the object is given by

$$\rho(t) = \int_{r_0}^{r(t)} \frac{rn^2(r)\,dr}{\left[r^2n^2(r) - b^2(t)\right]^{1/2}}. \tag{4}$$

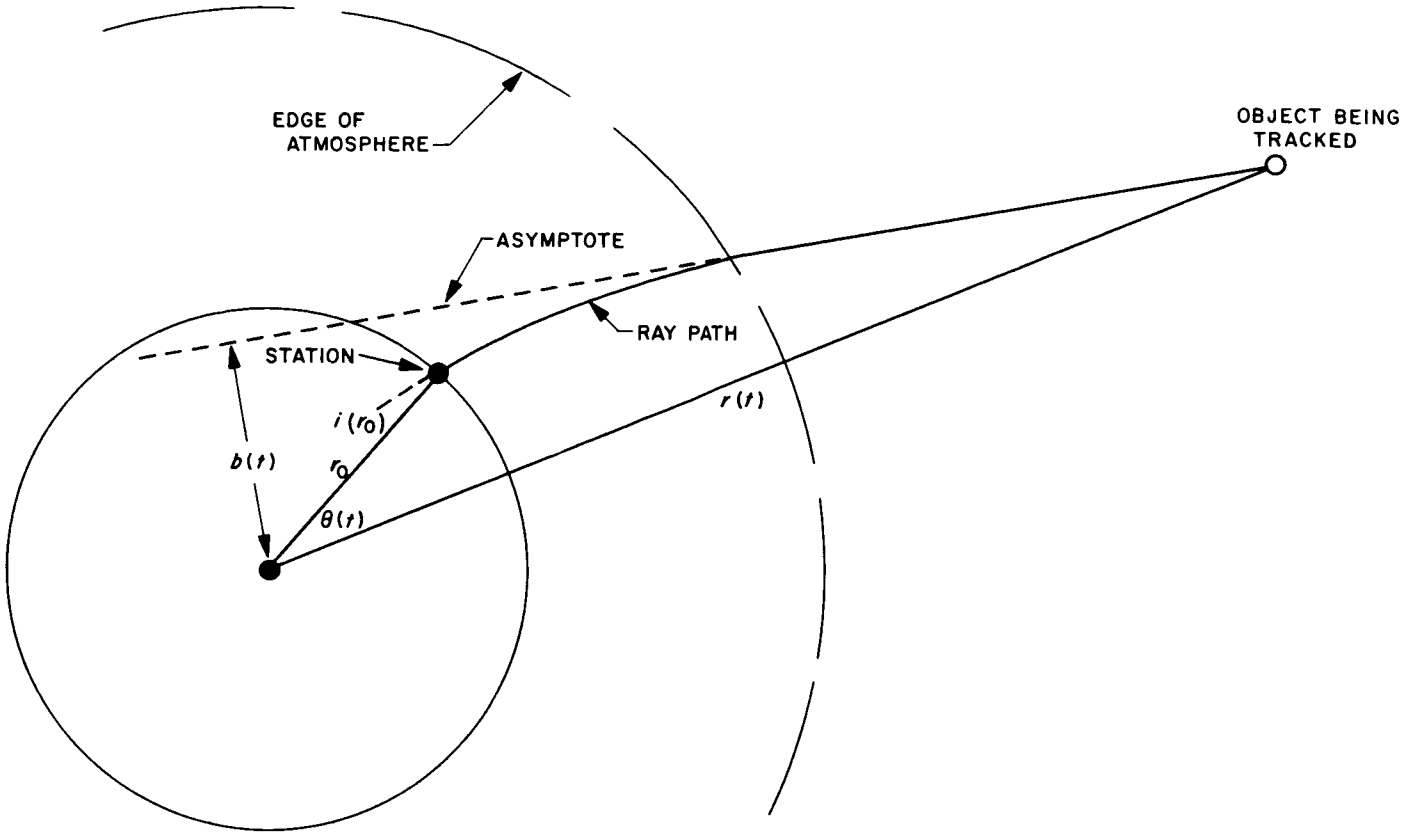


Fig. 1. Ray path geometry

To find the range rate, we differentiate Eq. (4) with respect to t and find $\dot{\rho}(t)$:

$$\dot{\rho}(t) = \int_{r_0}^{r(t)} \frac{b(t)\dot{b}(t)rn^2(r)\,dr}{[r^2n^2(r) - b^2(t)]^{3/2}} + \frac{r(t)n^2(r(t))\dot{r}(t)}{[r^2(t)n^2(r(t)) - b^2(t)]^{1/2}}.$$

But from Eq. (3) we have

$$\dot{\theta}(t) = \int_{r_0}^{r(t)} \frac{\dot{b}(t)rn^2(r)\,dr}{[r^2n^2(r) - b^2(t)]^{3/2}} + \frac{b(t)\dot{r}(t)}{r(t)[r^2(t)n^2(r(t)) - b^2(t)]^{1/2}}.$$

Hence it follows that

$$\dot{\rho}(t) = b(t)\dot{\theta}(t) + [r^2(t)n^2(r(t)) - b^2(t)]^{1/2}\left(\frac{\dot{r}(t)}{r(t)}\right). \quad (5)$$

Since, for cases of interest, the distance $r(t)$ to the object we are tracking is outside of the reaches of the atmosphere, then $n(r(t)) = 1$ and

$$\dot{\rho}(t) = b(t)\dot{\theta}(t) + [r^2(t) - b^2(t)]^{1/2}\left(\frac{\dot{r}(t)}{r(t)}\right). \quad (6)$$

Let us now define the impact parameter for the straight line path (i.e., the way the ray would travel if $n(r) = 1$) by $b_s(t)$. We can readily determine that

$$b_s(t) = \frac{r_0 r(t)\sin(\theta(t))}{[r_0^2 + r^2(t) - 2r_0r(t)\cos(\theta(t))]^{1/2}}. \quad (7)$$

For this case, the doppler residuals $f_r(t)$ are given by

$$f_r(t) = \frac{-2f_t}{c}\left[\dot{\theta}(t)(b(t) - b_s(t)) + \frac{\dot{r}(t)}{r(t)}[(r^2(t) - b^2(t))^{1/2} - (r^2(t) - b_s^2(t))^{1/2}]\right], \quad (8)$$

where f_t is the transmitter frequency and c is the speed of light.

Since $r(t)$, $\theta(t)$, $\dot{r}(t)$ and $\dot{\theta}(t)$ can all be obtained from the ephemeris, we see from Eq. (8) that the only unknown quantity needed to compute the doppler residuals is $b(t)$.

3. Solution for the Impact Parameter

If we are given $r(t)$ and $\theta(t)$ from the ephemeris, then the only unknown in Eq. (3) is $b(t)$. To solve Eq. (3) for

$b(t)$, let us form an iteratitive scheme for $\hat{i}(t)$, the angle $i(t)$ at the station. We may then find $b(t)$ from Eq. (2), i.e.,

$$b(t) = r_0 n(r_0) \sin \hat{i}(t) . \tag{9}$$

Since $i(t)$ is linearly related to θ(t) (*SPS 37-29*, Vol. IV, p. 231), we propose a linear iterative scheme:

$$i(t)_{k+1} = i(t)_k + \theta(t) - \int_{r_0}^{r(t)} \frac{b(t)_k \, dr}{r\left[r^2 n^2(r) - b^2(t)_k\right]^{1/2}} , \tag{10a}$$

where

$$b(t)_k = r_0 n(r_0) \sin i(t)_k \tag{10b}$$

As an initial guess at $i(t)$, let

$$i(t)_0 = \sin^{-1} \frac{b_s(t)}{r_0 n(r_0)} , \tag{10c}$$

where $b_s(t)$ is given by Eq. (2).

To find the error in the kth iteration of the scheme given by Eqs. (10), let us define

$$i(t)_{k+1} = f(i(t)_k) = i(t)_k + \theta(t) - \int_{r_0}^{r(t)} \frac{b(t)_k \, dr}{r\left[r^2 n^2(r) - b^2(t)_k\right]^{1/2}} \tag{11}$$

We see that $\hat{i}(t)$ satisfies Eq. (11) identically, i.e., $\hat{i}(t) = f(\hat{i}(t))$. Hence, we may write

$$i(t)_{k+1} - \hat{i}(t) = f(i(t)_k) - f(\hat{i}(t)) .$$

If we now define the error of the kth iteration as $\epsilon_k = i(t)_k - \hat{i}(t)$ and apply the mean value theorem, then

$$\epsilon_{k+1} = \epsilon_k \left| \frac{\partial f}{\partial i_k} \right| i(t)_k = \xi , \quad \xi \epsilon (\hat{i}(t), i(t)_k) . \tag{12}$$

To bound the above partial derivative of $f(i(t)_k)$, we shall need some bounds on $n(r)$. Let us choose

$$n(r_1) \leq n(r) \leq n(r_0) , \; r_1 \epsilon (r_0, r(t)) . \tag{13a}$$

We have, from Eqs. (11) and (10b),

$$\frac{\partial f}{\partial i_k} = 1 - \int_{r_0}^{r(t)} \frac{r_0 n(r_0) \cos i(t)_k \, r n^2(r) \, dr}{\left[r^2 n^2(r) - r_0^2 n^2(r_0) \sin^2 i(t)_k\right]^{3/2}} ,$$

and from Eq. (13a),

$$\frac{\partial f}{\partial i_k} \leq 1 - r_0 \frac{n^2(r_1)}{n^2(r_0)} \cos i(t)_k \int_{r_0}^{r(t)} \frac{r \, dr}{\left[r^2 - r_0^2 \sin^2 i(t)_k\right]^{3/2}} ,$$

which leads to the result

$$\frac{\partial f}{\partial i_k} \leq 1 - \frac{n^2(r_1)}{n^2(r_0)} \left(1 - \frac{r_0 \cos i(t)_k}{\left[r^2(t) - r_0^2 \sin^2 i(t)_k\right]^{1/2}} \right) .$$

Hence it follows that

$$\frac{\partial f}{\partial i_k} < 1 - \frac{n^2(r_1)}{n^2(r_0)} \left(1 - \frac{r_0}{r(t)} \right) ,$$

and from Eq. (12)

$$\epsilon_{k+1} < \epsilon_k \left(1 - \frac{n^2(r_1)}{n^2(r_0)} + \frac{n^2(r_1)}{n^2(r_0)} \frac{r_0}{r(t)} \right) . \tag{13b}$$

But for most cases of interest $n^2(r_1) = 1 - \eta_1$ and $n^2(r_0) = 1 + \eta_0$ with $\eta_1 \ll 1$ and $\eta_0 \ll 1$; hence the error of our iterative scheme depends on $r(t)$, because

$$\epsilon_{k+1} < \left(\frac{r_0}{r(t)} \right)^k \eta_0 ,$$

whenever the error ϵ_0 of our initial estimate of $i(t)$ is of the order of η_0.

4. The Differential Correction Scheme

Now that we have a method of computing doppler residuals based on a particular index of refraction, we need to utilize the observed residuals to construct the true index. We shall assume that we have available a set of observed doppler residuals and that corresponding to each residual we have $r(t)$, $\theta(t)$, $\dot{r}(t)$ and $\dot{\theta}(t)$ available. Let us further assume that the index of refraction in the model of the atmosphere involves n parameters α_i, $i = 1, \cdots, n$; that is, $n(r) = n(r;\alpha_1, \cdots, \alpha_n)$. These parameters to be estimated consist of quantities such as atmospheric scale height, refractivity at ground level, electron densities for different times of the day, temperature, and pressure, or any other quantities that would be useful. Let us denote the observed doppler residuals by $f_0(t_i)$, $i = 1, \cdots, m$, and the calculated residuals by $f_r(t_i)$, $i = 1, \cdots, m$. We then wish to expand $f_0(t_i)$ in a Taylor's expansion about our initial guess at α_i, $i = 1, \cdots, n$. Let us denote our initial guess by α_i^0, $i = 1, \cdots, n$. Retaining first-order terms in $(\alpha_i - \alpha_i^0)$ only, we obtain

$$f_0(t_i) = f_r(t_i) + \frac{\partial f_r(t_i)}{\partial \alpha_1^0} (\alpha_1 - \alpha_1^0) + \cdots + \frac{\partial f_r(t_i)}{\partial \alpha_n^0} (\alpha_n - \alpha_n^0) . \tag{14}$$

Let us denote

$$f_0(t_i) - f_r(t_i) = d(t_i) \quad i = 1, \cdots, m , \tag{15a}$$

$$(\alpha_j - \alpha_j^0) = \Delta \alpha_j^0 \quad j = 1, \cdots, n . \tag{15b}$$

Eq. (14) then becomes

$$d(t_i) = \frac{\partial f_r(t_i)}{\partial \alpha_1^0} \Delta \alpha_1^0 + \cdots + \frac{\partial f_r(t_i)}{\partial \alpha_n^0} \Delta \alpha_n^0 \quad i = 1, \cdots, m . \tag{16}$$

are able to measure a parameter involved in the process, say f. Let f be given explicitly by

$$f = f(\alpha_1, \cdots, \alpha_n) . \tag{1}$$

Suppose we make m observations of f, $(m > n)$, then let us designate the ith observation by f_i. We now have m equations in n unknowns

$$f_i = f_i(\alpha_1, \cdots, \alpha_n), i = 1, \cdots, m, (m > n) . \tag{2}$$

To form the differential correction scheme, we take the first-order terms of the Taylor's expansion of $f_i, i = 1, \cdots, m$, about our guess of $\alpha_i, i = 1, \cdots, n$. Let us call our kth estimate of α_i, α_i^k. Then

$$f_i = f_i(\alpha_1^k, \cdots, \alpha_n^k) + \frac{\partial}{\partial \alpha_1} f_i(\alpha_1^k, \cdots, \alpha_n^k)\,(\alpha - \alpha_1^k) + \cdots$$

$$+ \frac{\partial}{\partial \alpha_n} f(\alpha_1^k, \cdots, \alpha_n^k)\,(\alpha - \alpha_n^k) .$$

Let us define

$$r_i^k = f_i - f_i(\alpha_1^k, \cdots, \alpha_n^k) \tag{3a}$$

$$\Delta\alpha_i^k = \alpha - \alpha_i^k \tag{3b}$$

$$p_{ij}^k = \frac{\partial}{\partial \alpha_j} f_i(\alpha_1^k, \cdots, \alpha_n^k) . \tag{3c}$$

Then we may write our expansion as

$$r_i^k = p_{i1}^k \Delta\alpha_1^k + \cdots + p_{in}^k \Delta\alpha_n^k, i = 1, \cdots, m . \tag{4}$$

Let us now define

$$R_k^T = (r_1^k, \cdots, r_m^k)$$

$$\Delta_k^T = (\Delta\alpha_1^k, \cdots, \Delta\alpha_n^k) \tag{5a}$$

$$P_k = \begin{pmatrix} p_{11}^k & p_{12}^k & \cdots & p_{1n}^k \\ p_{21}^k & p_{22}^k & \cdots & p_{2n}^k \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ p_{m1}^k & p_{m2}^k & \cdots & p_{mn}^k \end{pmatrix} . \tag{5b}$$

Then Eq. (4) can be written

$$R_k = P_k \Delta_k .$$

Since the measurement proces of f_i, $i = 1, \cdots, m$, will undoubtedly contain errors, let us then assume that

$$E(R_k) = P_k \Delta_k , \tag{6a}$$

and

$$\mathrm{Var}(R_k) = V , \tag{6b}$$

where V is symmetric positive-definite.

To solve for Δ_k we apply the Gauss-Markoff theorem (Ref. 1) which yields

$$\Delta_k = (P_k^T V^{-1} P_k)^{-1} P_k^T V^{-1} R_k , \tag{7}$$

and is the least squares estimate of Δ_k. If M is a $p \times q$ matrix and A is a $q \times 1$ random vector then (Ref. 1)

$$\mathrm{Var}\,(MA) = M\,\mathrm{Var}(A) M^T . \tag{8}$$

Hence we have from Eq. (7)

$$\mathrm{Var}(\Delta_k) = (P_k^T V^{-1} P_k)^{-1} P_k^T V^{-1} \mathrm{Var}(R_k) V^{-1} P_k (P_k^T V^{-1} P_k)^{-1} ,$$

which yields from Eq. (6b)

$$\Sigma_k = \mathrm{Var}(\Delta_k) = (P_k^T V^{-1} P_k)^{-1} . \tag{9}$$

3. The Use of A-Priori Statistics

Let us assume that we know the variance-covariance matrix of our initial guess at α_i^0, $i = 1, \cdots, n$. Let us denote it V^*. Let us then weight Δ_0 by $V^*(V^* + \Sigma_0)^{-1}$ and denote this weighted answer by

$$\Delta_0^* = V^*(V^* + \Sigma_0)^{-1} \Delta_0 . \tag{10}$$

The term weighted by $\Sigma_0(V^* + \Sigma_0)^{-1}$ does not appear in Eq. (10) because we assume it to be zero with variance V^*. From Eqs. (7) and (9) we have

$$\Delta_0^* = V^* \left[V^* + (P_0^T V^{-1} P_0)^{-1}\right]^{-1} (P_0^T V^{-1} P_0)^{-1} P_0^T V^{-1} R_0 ,$$

which yields

$$\Delta_0^* = (P_0^T V^{-1} P_0 + V^{*-1})^{-1} P_0^T V^{-1} R_0 . \tag{11}$$

To find $V_0^* = \mathrm{Var}(\Delta_0^*)$, we use Eq. (8) which yields

$$V_0^* = (P_0^T V^{-1} P_0 + V^{*-1})^{-1} P_0^T V^{-1} \mathrm{Var}(R_0) V^{-1} P_0 (P_0^T V^{-1} P_0 + V^{*-1})^{-1} ,$$

In order to use Eq. (16) in a regression scheme, we must compute $\partial f_r(t_i)/\partial\alpha_j^0$. Since $b(t_i)$ is a function of α_j, $j = 1, \cdots, n$ by Eq. (2), then Eq. (8) leads to

$$\frac{\partial f_r(t_i)}{\partial\alpha_j^0} = -\frac{2f_1}{c}\left[\dot{\theta}(t_i) - \frac{\dot{r}(t_i)b(t_i)}{r(t_i)\left[r^2(t_i)n^2(r(t_i)) - b^2(t_i)\right]^{1/2}}\right]\frac{\partial b(t_i)}{\partial\alpha_j^0}. \quad (17)$$

Now the only unknown appearing in the right side of Eq. (17) is $\partial b(t_i)/\partial\alpha_j^0$. Since $b(t_i)$ is given implicitly by Eq. (3), let us differentiate Eq. (3) with respect to α_j^0:

$$0 = \int_{r_0}^{r(t_i)} \frac{\dfrac{\partial b(t_i)}{\partial\alpha_j^0} rn^2(r)\,dr}{\left[r^2n^2(r) - b^2(t_i)\right]^{3/2}} - \int_{r_0}^{r(t_i)} \frac{b(t_i)rn(r)\dfrac{\partial n(r)}{\partial\alpha_j^0}\,dr}{\left[r^2n^2(r) - b^2(t_i)\right]^{3/2}}.$$

Since $\partial b(t_i)/\partial\alpha_j^0$ is constant with respect to the integration process,

$$\frac{\partial b(t_i)}{\partial\alpha_j^0} = \frac{\displaystyle\int_{r_0}^{r(t_i)} \frac{b(t_i)rn(r)\dfrac{\partial n(r)}{\partial\alpha_j^0}\,dr}{\left[r^2n^2(r) - b^2(t_i)\right]^{3/2}}}{\displaystyle\int_{r_0}^{r(t_i)} \frac{rn^2(r)\,dr}{\left[r^2n^2(r) - b^2(t_i)\right]^{3/2}}}. \quad (18)$$

Using Eq. (18) we are thus able to compute $\partial f_r(t_i)/\partial\alpha_j^0$ from Eq. (17), and we now have all the terms in Eq. (16). Now to solve the system of Eqs. (16) for Δ_j^0, (i.e., for α_j since $\alpha_j^0 + \Delta_j^0 = \alpha_j$), let

$$D_0^T = \left(d(t_1), \cdots, d(t_n)\right),\ \Delta_0^T = \left(\Delta\alpha_1^0, \cdots, \Delta_n^0\right); \quad (19a)$$

$$P_0 = \begin{pmatrix} \dfrac{\partial f_r(t_1)}{\partial\alpha_1^0} & \dfrac{\partial f_r(t_1)}{\partial\alpha_2^0} & \cdots & \dfrac{\partial f_r(t_1)}{\partial\alpha_n^0} \\ \dfrac{\partial f_r(t_2)}{\partial\alpha_1^0} & \dfrac{\partial f_r(t_2)}{\partial\alpha_2^0} & \cdots & \dfrac{\partial f_r(t_2)}{\partial\alpha_n^0} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial f_r(t_m)}{\partial\alpha_1^0} & \dfrac{\partial f_r(t_m)}{\partial\alpha_2^0} & \cdots & \dfrac{\partial f_r(t_m)}{\partial\alpha_n^0} \end{pmatrix} \quad (19b)$$

Then Eq. (16) can be written

$$D_0 = P_0\Delta_0. \quad (20)$$

Assuming that we have more observations that parameters (i.e., $m > n$), we may use the Gauss-Markoff Theorem for a least-squares solution of Δ_0. If $\mathrm{Var}(D_0) = V$, where V is symmetric and positive definite, then

$$\Delta_0 = \left(P_0^TV^{-1}P_0\right)^{-1}P_0^TV^{-1}D_0. \quad (21)$$

Once we have Δ_0 from Eq. (21), we can take the new parameters $\alpha_j^1 = \alpha_j^0 + \Delta_j^0$, $j = 1, \cdots, n$, and repeat the entire process until we get sufficiently close to the true parameters α_j.

If *a-priori* information is possessed about α_j^0, $j = 1, \cdots, n$, then a conditioned set of equations may be used in lieu of Eq. (21). See "A-Priori Statistics in Differential Correction" in Sect. B.

5. Conclusion

The method discussed in this article could be used advantageously to construct a model atmosphere for the Goldstone environment. Doppler residuals from *Ranger* or *Mariner* data could be used, for example, as well as planetary tracking data. Because larger residuals occur at lower elevation angles, good data would be needed at these angles.

B. A-Priori Statistics in Differential Correction

P. E. Reichley

1. Introduction

A well-known technique in the physical sciences is the differential correction of a set of parameters related to a physical process. This differential correction is usually accomplished with respect to a set of observations of some particular observable in the physical process. When *a-priori* statistics are known about the parameters in the process, these statistics may be used to condition the normal equations of the differential correction scheme. This article shows how these statistics are employed to obtain conditioned answers and the statistics of the answers.

2. The Differential Correction Scheme

Let us consider a physical process in which we wish to estimate n parameters, α_i, $i = 1, \cdots, n$, and in which we

and

$$V_0^* = (P_0^T V^{-1} P_0 + V^{*-1})^{-1} P_0^T V^{-1} P_0 (P_0^T V^{-1} P_0 + V^{*-1})^{-1} . \quad (12)$$

We can then iterate:

$$\Delta_k^* = (P_k^T V^{-1} P_k + V_{k-1}^{*-1})^{-1} P_k^T V^{-1} R_k \quad (13a)$$

$$V_k^* = (P_k^T V^{-1} P_k + V_{k-1}^{*-1})^{-1} P_k^T V^{-1} P_k (P_k^T V^{-1} P_k + V_{k-1}^{*-1})^{-1} . \quad (13b)$$

Looking at Eq. (13), we see that the weighting by V_{k-1}^* has the effect of conditioning the normal equations such that if $P_k^T V^{-1} P_k$ is near singular then V_{k-1}^* moves it away from the singularity.

If Var $(R_k) = \sigma^2 I$, then Eq. (13) takes the form

$$\Delta_k^* = (P_k^T P_k + \sigma^2 V_{k-1}^{*-1})^{-1} P_k^T R_k$$

$$V_k^* = (P_k^T P_k + \sigma^2 V_{k-1}^{*-1})^{-1} P_k^T P_k (P_k^T P_k + \sigma^2 V_{k-1}^{*-1})^{-1} \sigma^2 .$$

4. Discussion

No attempt has been made to compare Eq. (13) with Eq. (7) and Eq. (9). This is a good problem. The weighted scheme using *a-priori* statistics and the normal scheme have both been used in a problem constructing model atmospheres, and the *a-priori* scheme yields the best answers (Sect. A in this issue).

Reference

1. Zelen, M., *The Role of Constraints In The Theory of Least Squares*, Mathematics Research Center, Technical Summary Report No. 314, University of Wisconsin, May 1962.

QUALITY ASSURANCE AND RELIABILITY OFFICE

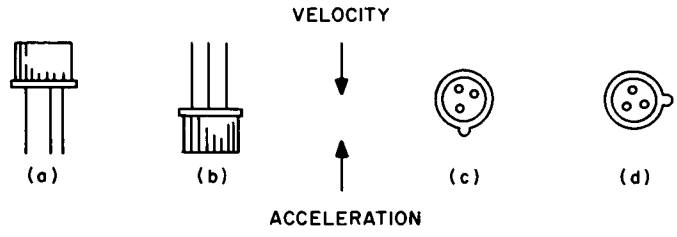
XXVII. Parts Reliability

A. Effect of High Impact Shock on Seven Types of Transistors

K. Martin

Due to the hard-landing requirements anticipated for future planetary exploration, it is highly desirable to determine the electronic parts capable of withstanding high impact shock as soon as possible. Some early tests conducted on seven types of transistors are discussed here. At first, several of the more common parameters were measured and recorded before and after the first two shocks. It became apparent that shock was not causing any discernible parameter drift. All failures were "catastrophic" due to internal lead failure or bond separation from the chip pad. Consequently, discreet parametric measurements were deleted, and a Tektronix, Inc., Type 575 curve tracer was substituted so that only gross changes in the remaining impacts would be observed.

Four of each of the following types of transistors were tested: Texas Instruments Incorporated Types 2N930, 2N2432, 2N2605, 2N3329, and 2N3350; Fairchild Controls Corp. Type 2N3117; and Solid State Products, Inc., Type 2N2843. The four parts were hard-mounted in holes pre-drilled in a single aluminum block and were shocked in the positions illustrated below:



The amplitude of each impact, the velocity, and the test results are shown in Table 1. All damage occurred in Positions (c) and (d). Apparently, lateral impact is the most severe due to shearing or, perhaps, leverage developed by the internal lead mass.

From the results of this very limited testing, it appears that planar-type transistors, using ball bonding techniques with proper bonding and no defects in the internal leads, are capable of withstanding a 10^4-g shock as applied in these tests. Observations indicate that the failures were

Table 1. High-impact-shock test results

Acceleration, g (approximate)	Velocity, ft/sec	Failed transistor(s)	Position
4.4×10^3	107	None	—
6.3×10^3	108	Texas Instruments Type 2N3329	(c)
10^4	183	Texas Instruments Types 2N930 and 2N3350	(d)
10^4	175	Fairchild Type 2N3117	(c)
10^4	173	None	—
10^4	183	None	—

caused by poor bonding and, in the case of Type 2N930, a weakened internal lead due to the bonding process. Consequently, a screening method to eliminate these weak bonds and leads must be established. Additional and more comprehensive testing is planned, including the determination of a suitable screening method and the evaluation of other bonding techniques.

B. Diode and Resistor Screening Evaluation

M. Johnston

During late-1964, two test programs were initiated to evaluate the effects of various degrees of screening upon particular types of parts. One program involved testing of two diode types, the Fairchild Semiconductor Corp. Type 1N916 and the Pacific Semiconductors, Inc., Type 1N485B, and the other program involved testing of a Texas Instruments Incorporated Type CG¼ resistor. Each program has now accumulated 5000-hr test time.

To fulfill the purpose of the test, diodes of each type were purchased to each of three screening levels, with Level 1 corresponding to no screening (i.e., a commercial grade part) and Levels 2 and 3 corresponding to increasingly stringent degrees of screening. Resistors were purchased to Levels 1 and 2 only. Two-hundred parts of each screening level were subjected to rated conditions for a 5000-hr operating life test. The differences in the reaction of the parts to the testing can validly be assumed to be the result of the treatment prior to testing (i.e., the screening process) since each part type was subjected to the same test conditions.

Measurements of the device parameters were made at specific periods during the life test to detect significant parametric changes. Statistical comparisons of the parameter measurements were then made in order to examine the parts' behavior. None of the Type 1N916 diodes under test has exhibited either "catastrophic" or parametric failure. Preliminary analysis of the parameter drift data has not revealed any differences in the behavior of the three test groups. At the present stage of analysis, the test results are inconclusive in establishing that the particular screens employed resulted in any improvement over the performance of the Level 1 devices.

Comparison of the test results of the three levels of Type 1N485B diodes shows that a decidedly better group of parts resulted from screening. Forty-six leakage failures and three mechanical failures were eliminated by the vendor in screening. Of the 200 Level 1 diodes in the test, 15 have exhibited lead failures, but surprisingly no leakage failures have occurred. No failures of any type have occurred in the groups of screened devices.

Analysis of the test results for the Texas Instruments Type CG¼ resistor shows that, after 5000-hr operation, approximately 5% of the Level 1 devices are out-of-resistance-tolerance, whereas approximately 1.5% of the screened devices are out-of-resistance-tolerance. One lead failure in the Level 2 group occurred after 4000-hr operation; however, the failure appears to be the result of mishandling. The test results indicate that a better group of parts resulted from screening.

Present plans are to continue both test programs in order to acquire useful knowledge about the behavior of the parts over longer periods of time.