NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Space Programs Summary 37-65, Vol. II

The Deep Space Network

For the Period July 1 to August 31, 1970

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

September 30, 1970

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Space Programs Summary 37-65, Vol. II

The Deep Space Network

For the Period July 1 to August 31, 1970

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

September 30, 1970

SPACE PROGRAMS SUMMARY 37-65, VOL. II

Prepared Under Contract No. NAS 7-100
National Aeronautics and Space Administration

Preface

The Space Programs Summary is a multivolume, bimonthly publication that presents a review of technical information resulting from current engineering and scientific work performed, or managed, by the Jet Propulsion Laboratory for the National Aeronautics and Space Administration. The Space Programs Summary is currently composed of four volumes:

Vol. I. *Flight Projects* (Unclassified)

Vol. II. *The Deep Space Network* (Unclassified)

Vol. III. *Supporting Research and Advanced Development* (Unclassified)

Vol. IV. *Flight Projects and Supporting Research and Advanced Development* (Confidential)

Foreword

Volume II of the Space Programs Summary reports the results of work performed by the Deep Space Network (DSN). Information is presented, as appropriate, in the following categories:

Introduction
- Description of the DSN
- Description of DSN Systems

Mission Support
- Interplanetary Flight Projects
- Planetary Flight Projects
- Manned Space Flight Project
- Advanced Flight Projects

Advanced Engineering
- Tracking and Navigational Accuracy Analysis
- Communications Systems Research
- Communications Elements Research
- Supporting Research and Technology

Development and Implementation
- Space Flight Operations Facility Development
- Ground Communications Facility Development
- Deep Space Instrumentation Facility Development
- DSN Project and System Development

Operations and Facilities
- DSN Operations
- Space Flight Operations Facility Operations
- Ground Communications Facility Operations
- Deep Space Instrumentation Facility Operations
- Facility Engineering

In each issue, the section entitled "Description of DSN Systems" reports the current configuration of one of the six DSN systems (tracking, telemetry, command, monitoring, simulation, and operations control). The fundamental research carried out in support of the DSN is reported in Vol. III.

Contents

Contents (contd)

Contents (contd)

OPERATIONS AND FACILITIES

I. Description of the DSN

INTRODUCTION

The Deep Space Network (DSN), established by the NASA Office of Tracking and Data Acquisition under the system management and technical direction of JPL, is designed for two-way communications with unmanned spacecraft traveling approximately 10,000 mi from earth to planetary distances. It supports, or has supported, the following NASA deep space exploration projects: *Ranger*, *Surveyor*, *Mariner* Venus 1962, *Mariner* Mars 1964, *Mariner* Venus 67, *Mariner* Mars 1969, *Mariner* Mars 1971 (JPL); *Lunar Orbiter* and *Viking* (Langley Research Center); *Pioneer* (Ames Research Center); *Helios* (West Germany); and *Apollo* (Manned Spacecraft Center), to supplement the Manned Space Flight Network (MSFN).

The DSN is distinct from other NASA networks such as the MSFN, which has primary responsibility for tracking the manned spacecraft of the *Apollo* Project, and the Space Tracking and Data Acquisition Network (STADAN), which tracks earth-orbiting scientific and communications satellites. With no future unmanned lunar spacecraft presently planned, the primary objective of the DSN is to continue its support of planetary and interplanetary flight projects.

To support flight projects, the DSN simultaneously performs advanced engineering on components and systems, integrates proven equipment and methods into the network,[1] and provides direct support of each project through that project's Tracking and Data System. This management element and the project's Mission Operations personnel are responsible for the design and operation of the data, software, and operations systems required for the conduct of flight operations. The organization and procedures necessary to carry out these activities are described in SPS 37-50, Vol. II, pp. 15–17.

By tracking the spacecraft, the DSN is involved in the following data types:

(1) *Metric:* generate angles, one- and two-way doppler, and range.

(2) *Telemetry:* receive, record, and retransmit engineering and scientific data.

(3) *Command:* send coded signals to the spacecraft to activate equipment to initiate spacecraft functions.

[1]When a new piece of equipment or new method has been accepted for integration into the network, it is classed as Goldstone duplicate standard (GSDS), thus standardizing the design and operation of identical items throughout the network.

The DSN operation is characterized by six DSN systems: (1) tracking, (2) telemetry, (3) command, (4) monitoring, (5) simulation, and (6) operations control.

The DSN can be characterized as being comprised of three facilities: the Deep Space Instrumentation Facility (DSIF), the Ground Communications Facility (GCF), and the Space Flight Operations Facility (SFOF).

1. Deep Space Instrumentation Facility

a. Tracking and data acquisition facilities. A worldwide set of deep space stations (DSSs) with large antennas, low-noise phase-lock receiving systems, and high-power transmitters provide radio communications with spacecraft. The DSSs and the deep space communications complexes (DSCCs) they comprise are given in Table 1.

Radio contact with a spacecraft usually begins when the spacecraft is on the launch vehicle at Cape Kennedy, and it is maintained throughout the mission. The early part of the trajectory is covered by selected network stations of the Air Force Eastern Test Range (AFETR) and the MSFN of the Goddard Space Flight Center.[2] Normally, two-way communications are established between the spacecraft and the DSN within 30 min after the spacecraft has been injected into lunar, planetary, or interplanetary flight. A compatibility test station at Cape Kennedy (discussed later) monitors the spacecraft continuously during the launch phase until it passes over the local horizon. The deep space phase begins with acquisition by either DSS 51, 41, or 42. These and the remaining DSSs given in Table 1 provide radio communications to the end of the flight.

To enable continuous radio contact with spacecraft, the DSSs are located approximately 120 deg apart in longitude; thus, a spacecraft in deep space flight is always within the field-of-view of at least one DSS, and for several hours each day may be seen by two DSSs. Furthermore, since most spacecraft on deep space missions travel within 30 deg of the equatorial plane, the DSSs are located within latitudes of 45 deg north or south of the equator. All DSSs operate at S-band frequencies: 2110–2120 MHz for earth-to-spacecraft transmission and 2290–2300 MHz for spacecraft-to-earth transmission.

To provide sufficient tracking capability to enable useful data returns from around the planets and from the edge of the solar system, a 210-ft-diam-antenna network will be required. Two additional 210-ft-diam-antenna DSSs are under construction at Madrid and Canberra, which will operate in conjunction with DSS 14 to provide this capability. These stations are scheduled to be operational by early 1973.

b. Compatibility test facilities. In 1959, a mobile L-band compatibility test station was established at Cape Kennedy to verify flight-spacecraft–DSN compatibility prior to the launch of the *Ranger* and *Mariner* Venus 1962 spacecraft. Experience revealed the need for a permanent facility at Cape Kennedy for this function. An S-band compatibility test station with a 4-ft-diam antenna became operational in 1965. In addition to supporting the preflight compatibility tests, this station monitors the spacecraft continuously during the launch phase until it passes over the local horizon.

Spacecraft telecommunications compatibility in the design and prototype development phases was formerly verified by tests at the Goldstone DSCC. To provide a more economical means for conducting such work and because of the increasing use of multiple-mission telemetry and command equipment by the DSN, a compatibility test area (CTA) was established at JPL in 1968. In all essential characteristics, the configuration of this facility is identical to that of the 85- and 210-ft-diam-antenna stations.

The JPL CTA is used during spacecraft system tests to establish the compatibility with the DSN of the proof test model and development models of spacecraft, and the Cape Kennedy compatibility test station is used for final flight spacecraft compatibility validation testing prior to launch.

2. Ground Communications Facility

The GCF, using, in part, facilities of the worldwide NASA Communications Network (NASCOM),[3] provides voice, high-speed data, and teletype communications between the SFOF and all DSSs, except those of the Goldstone DSCC. Communications between the Goldstone DSCC and the SFOF are provided by a microwave link leased from a common carrier. Early missions were

[2]The 30-ft-diam-antenna station established by the DSN on Ascension Island during 1965 to act in conjunction with the MSFN orbital support 30-ft-diam-antenna station was transferred to the MSFN in July 1968.

[3]Managed and directed by the Goddard Space Flight Center.

Table 1. Tracking and data acquisition stations of the DSN

DSCC	Location	DSS	DSS serial designation	Antenna		Year of initial operation
				Diameter, ft	Type of mounting	
Goldstone	California	Pioneer	11	85	Polar	1958
		Echo	12	85	Polar	1962
		(Venus)[a]	13	85	Az–El	1962
		Mars	14	210	Az–El	1966
—	Australia	Woomera[b]	41	85	Polar	1960
Tidbinbilla	Australia	Weemala (formerly Tidbinbilla)[b]	42	85	Polar	1965
		Ballima[b] (formerly Booroomba)	43	210	Az–El	Under construction
—	South Africa	Johannesburg[b]	51	85	Polar	1961
Madrid	Spain	Robledo[b]	61	85	Polar	1965
		Cebreros[b]	62	85	Polar	1967
		Robledo	63	210	Az–El	Under construction

[a]A research-and-development facility used to demonstrate the feasibility of new equipment and methods to be integrated into the operational network. Besides the 85-ft-diam az–el-mounted antenna, DSS 13 has a 30-ft-diam az–el-mounted antenna that is used for testing the design of new equipment and support of ground-based radio science.

[b]Normally staffed and operated by government agencies of the respective countries (except for a temporary staff of the Madrid DSCC), with some assistance of U.S. support personnel.

supported with voice and teletype circuits only, but increased data rates necessitated the use of wide-band circuits from all DSSs.

3. Space Flight Operations Facility

Network and mission control functions are performed at the SFOF at JPL. (Prior to 1964, these functions were performed in temporary facilities at JPL.) The SFOF receives data from all DSSs and processes that information required by the flight project to conduct mission operations. The following services are provided: (1) real-time processing and display of metric data; (2) real-time and non-real-time processing and display of telemetry data; (3) simulation of flight operations; (4) near-real-time evaluation of DSN performance; (5) operations control, and status and operational data display; and (6) general support such as internal communications by telephone, intercom, public address, closed-circuit TV, documentation, and reproduction of data packages. Master data records of science data received from spacecraft are generated. Technical areas are provided for flight project personnel who analyze spacecraft performance, trajectories, and generation of commands.

The SFOF is equipped to support many spacecraft in flight and those under test in preparation for flight. Over a 24-h period in 1967, as many as eight in-flight spacecraft or operational-readiness tests for flight were supported by the SFOF.

II. Description of DSN Systems

INTRODUCTION

A. System Simulation Models, *D. C. Card*

1. Introduction

The functional design and validation of design concepts for complex systems of the magnitude of the Deep Space Network (DSN) require the aid of powerful system modeling techniques, as well as practical methods for employing models as engineering tools. Monte Carlo simulation of systems operations has provided an effective means of analyzing the influence of system structure upon the dynamic flow of traffic through the systems in many types of application. Monte Carlo simulation methods provide one effective aid to system analysis in the DSN for two basic reasons:

(1) The DSN is a complex of identifiable elements through which traffic elements of several distinct types flow.

(2) The General Purpose Simulation System (GPSS) program is readily available and provides a fairly complete and adaptable language for the construction and operation of simulation models on high-speed digital computers.

2. The GPSS Program

A useful model for system analysis must have a structure that is representative of the structural elements of the system being studied; therefore, before a simulation model is constructed, a functional flow diagram and logical flow and control diagram for the system must be available. Once these requirements are met, the interpretation of system diagrams into GPSS language blocks is reasonably direct. Of primary importance to the codes is complete understanding of the logical operation of GPSS block types and overall simulation control.

The IBM GPSS program is currently available in two versions on the JPL IBM 7094 and 360/75 systems; GPSS III is located on the 7094 system and GPSS/360 is on the 360/75 system. The GPSS/360 version is somewhat more versatile than GPSS III; greater capacity has been provided for simulating systems of greater complexity in more detail.

The GPSS is a Monte Carlo simulation system that provides a high-level structural language for interpreting

functional and logical diagram blocks into simulation model blocks. The program is most directly applicable to the simulation of systems in which the study of flow characteristics of discrete traffic units through the system is of interest. Simulation system time is discrete, and the unit of time is completely arbitrary.

The GPSS program deck provides the random number generators and the overall flow control logic for the execution of simulation models, as well as computational and output format logic for simulation model statistical output. The input deck contains the model structure and the characteristic parameters to represent the flow through the system; it also contains cards to control the accumulation and tabulation of statistics.

After the model structure has been fully determined and compiled, the program begins simulation "runs" by generating traffic units (called *transactions)* according to rules specified by the programmer. These transactions are the dynamic input to the system model. The number of transactions, first generation time, arrival rate distribution, and priority levels are controlled by the programmer's input specification.

The flow of transactions through a model is controlled by system blocks. The path of transactions through the model follows the block sequence, and may be controlled logical transfers, duplication of blocks, and termination blocks, as well.

Transactions may carry with them several distinct parametric values, which may be introduced or altered at any stage of the system. These parameters provide a powerful, direct method for identifying and controlling distinct traffic elements depending upon individual characteristics. Parameter values may be used to indicate traffic type for flow logic control, characteristic service time for one or more service elements in the system, etc.

The output from GPSS simulation runs is orientated toward the study of traffic flow characteristics. The GPSS language provides blocks which "monitor" flow of traffic at any desired point in the system model. Flow statistics of several types, such as overall transit time, transit time between specified points, arrival rate, etc., may be accumulated and tabulated for convenient output. Further, statistics are accumulated for queue contents, waiting times, and for facility utilization factors.

The following IBM publications should be referred to for detailed description and operating instructions for the GPSS:

(1) GH20-0304-4, General Purpose Simulation System/360: Introductory User's Manual.

(2) H20-0326-2, General Purpose Simulation System/360: User's Manual.

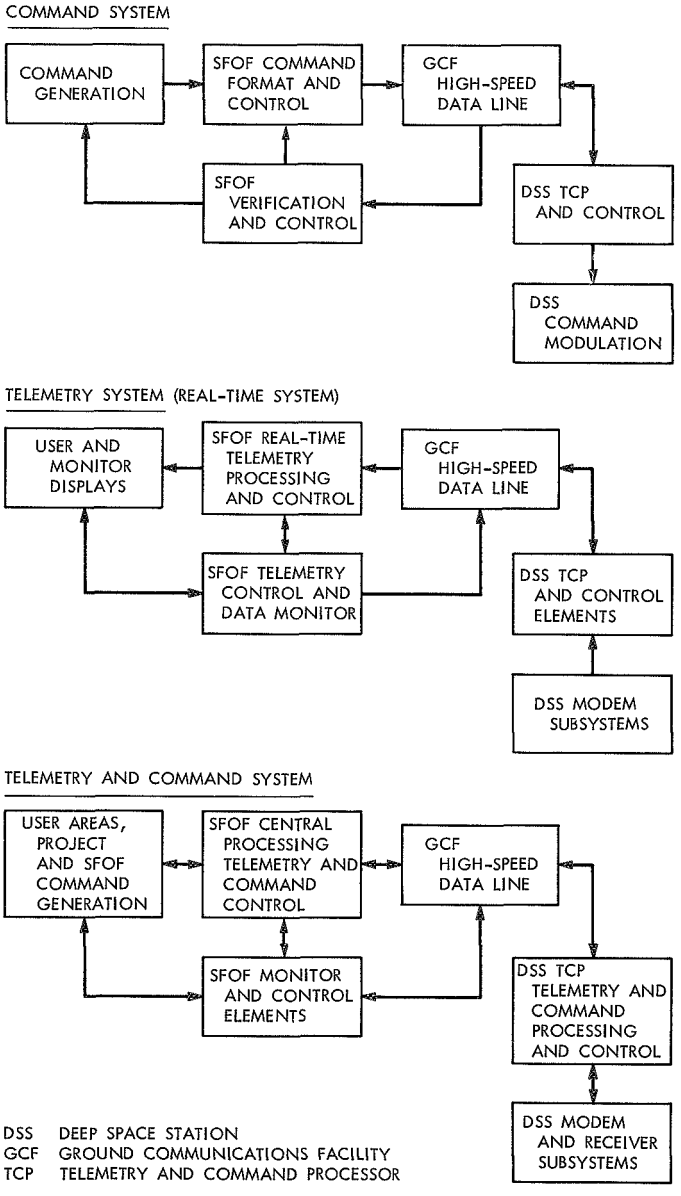


Fig. 1. Functional flow diagrams for separate and combined telemetry and command systems

3. Simulation Models for the DSN

Since the DSN is a large complex network system, it was decided that the most feasible approach to the construction of system models would be to provide separate models for each of the DSN systems, then combine the models into an overall DSN model. This method allows the functional aspects of the model system to be tested and validated in detail. The capability of GPSS models for modularity permits the separate system models to be "patched together" at the interfaces in correspondence with DSN design, and with relatively little alteration of system blocks. At this time, representative models for the Telemetry and Command Systems (Multiple-Mission Command and Multiple-Mission Telemetry Systems) are being constructed and exercised. The results should provide a good test of the modular aspects of the GPSS.

As a simple example of separate system model construction and the ultimate combination of models, consider a single branch of the Command System and the Telemetry System, as shown at a gross detail level in Fig. 1. The system elements and flow logic are assembled separately for each system, and the simulation model is constructed and executed in order to "debug" and validate the model logic. Then, the models may be combined as represented by the bottom diagram, with corresponding logical control for the various traffic elements and data handling operations in the system.

4. DSN Multiple-Mission Command System

At this writing, a fairly complete simulation model of the DSN Command System has been constructed and executed, using GPSS/360. The simulation model was constructed to be representative of the functional elements diagram and logical flow and control diagram.[1] The characteristic values for parameter, as well as some slight modification and elucidation of the original flow elements, were derived from design review meetings and direct communication with the system engineer. The results of the simulation runs provided no essential surprises, since the command system by itself puts no appreciable burden on the data processing and data handling elements of the DSN. However, it is expected that the results will be more fruitful when various systems designs are combined into unified simulation models. Work is currently proceeding on the Multiple-Mission Telemetry System model effort. The two system models will be expanded and combined in order to evaluate the design concepts and possible variation of these concepts.

[1]Contained in *Deep Space Network/System Requirements: DSN Command System* (Rev. A), December 1969 (JPL internal document).

B. Monitor System, *J. E. Allen*

1. Introduction

Historically, the Monitor System has been characterized as a data-gathering system and as such has been relegated the role of DSN system performance monitoring. The Monitor System provides the capability to determine failures to the facility subsystem level and distribute data to the DSN Tracking, Telemetry, Command, and Operations Control Systems.

The DSN Monitor System provides the capability for sensing certain characteristics of the various elements of the DSN, for processing and displaying these data for the use by DSN operations personnel, and for storing these data for later analysis. Monitor data are used for determining DSN status and configurations, for guidance in directing DSN operations, for providing alarms of nonstandard conditions, and for analysis of the quality and quantity of data provided to the projects.

DSN monitor data are defined as a selected subset of the machine-accessible facility monitor data (DSIF, GCF, and SFOF) that monitor the state, performance level, and configuration for any operationally active element of the DSN.

The scope of the Monitor System is defined as the specification, acquisition, conversion, transmission, selection, reduction, error detection, distribution, storage, and display of DSN facility and DSN system monitor data.

This article will be confined to the DSN Monitor System as the SFOF through 1971 and, in a brief sense, defines the facility interface with the Monitor System.

2. Application

The Monitor operation at the SFOF is primarily an automatic operation with personnel manning the input/output device and analysts evaluating the data. The software for the Monitor resides in the IBM 360/75 computer and is linked through the CDC 3100 computer, which in turn acts as the buffer for digital TV displays.

Raw data from the DSN sites via high-speed data lines (HSDLs) are input to the 360/75 computer and are collected in the DSN data accumulation bank. They are processed, evaluated, formatted, filed, displayed, and distributed in real time to DSN monitor data users. The stored data are validated in terms of quality and quantity within 12 h after each pass. At that time, mission data and questionable data are recalled from the individual sites and merged in the stored data file in a time-ordered fashion. Following the data recall and data merge, a Monitor System Data Record (SDR) is written on magnetic tape for storage.

3. Operational Description

The operation and control of the DSN Monitor System may be characterized as a number of open loops between the SFOF Central Processing System (CPS) and the DSIF Monitor System at the deep space station (DSS). Many of the monitor functions now performed manually or semi-automatically at the SFOF will be done automatically in the Mark III A era.

4. Basic Characteristics

The basic characteristics of the DSN Monitor System are as follows:

(1) Uniform instrumentation and reporting on status of elements of the DSN systems.

(2) Automatic and semi-automatic detection and reporting of all specified characteristics required of the DSN to operations and systems control points.

(3) Unified control and distribution of facility and DSN system standards and limits.

(4) Correlation of information to be provided by the Monitor System for facility and DSN system reports, both real time and non-real time.

(5) Adaptability to a wide range of DSN configurations, mission types, and mission loading.

5. Monitor Constraints

The following constraints will be observed by the Monitor System:

(1) The Monitor System will not preclude the data flow of other DSN systems, nor is it mandatory that the Monitor System be operational in order to deliver real-time data to Telemetry, Tracking, or Command System user.

(2) Each facility monitor system will be capable of operating independently from the SFOF Monitor.

(3) Definition of DSN configuration and capability requirements or changes shall be received by the Monitor System as changes to the monitor criteria data (MCD) sets, not as changes to the Monitor System.

6. System Operation

a. System description. The DSN Monitor System functionally resides in four processors, one in each DSN facility (DSIF, GCF, SFOF) and one in the DSN. Figure 1 is a functional data flow chart of the DSN Monitor System.

Each facility monitor processor will accept reports of configuration, load, performance, and status from the facility instrumentation and will accept reports on data accountability and quality from the DSN Tracking, Telemetry, and Command System processors. In addition, each facility processor will accept all standards and limits from the DSN monitor processor and distribute them to the appropriate error detectors. A facility analysis program will compare actual against standard configuration, load, status, accountability, and quality and will alarm nonstandard performance to the facility control and to the DSN Monitor System. Ground Communications Facility performance parameters measured at a DSS will be reported to the DSIF monitor or to the GCF monitor or both. The DSN monitor processor will provide applicable monitor data to the DSN facility chiefs and DSN system operations group. Nonstandard performance will be automatically alarmed by the DSN monitor processor. Alarms will be routed

to the system and facility involved and to the DSN Operations Control System. Reports on performance data will be made to facilities and systems as required.

Raw standards and limits will be gathered from each facility and from each DSN systems operations group and processed against a single instrumentation catalog for distribution to facilities for their system error detectors as facility-specific data sets.

A DSN monitor SDR (also called a DSN status SDR) will be made by the Monitor System from recorded DSN monitor data, which will be subsequently processed into a DSN extract data record in accordance with the processing criteria of the DSN Monitor Operations Control System.

b. Instrumentation and reports catalog. The instrumentation and reports catalog will consist of a file of DSN Operations Control and Monitor instrumentation in its current configuration. As a minimum, it will include details of standards and limits, data display formats, report requirements, and associated design data as required.

c. DSN Monitor System standards and limits. The DSN Monitor Operations Group will gather the required inputs, validate for accuracy, completeness, and compatibility between facility and DSN system subsets, and maintain the DSN standards and limits file. The types of inputs used shall be:

(1) DSN configurations from the scheduling activity.

(2) Detectable control actions expected (sequence of events) from Operations Control.

(3) DSN status display content requirements.

(4) The DSN instrumentation catalog (a description of each instrumented DSN element).

(5) Data content requirements for each summary report.

d. Initialization. The DSN Monitor Operations Group will coordinate, verify, check for consistency, and control the distribution of the standards and limits to each facility and each DSN system.

e. System performance monitoring. Each facility monitor system and the DSN Monitor System will automatically output specific alarm messages to the DSN Operations Control System based on nonconformance with standards or required performance.

f. Data quality and accountability reports. Data quality and accountability reports to other DSN systems are based on discrepancy reports from GCF monitoring and the DSN Telemetry, Tracking, and Command System status report. The reports shall identify (if possible) the records that were missed or that contain bad data. Information on known performance anomalies shall be correlated with the status report. This report allows the DSN system operations groups to make an immediate post-pass recall to fill in missing data in the master data record (MDR).

g. DSN Monitor System operational testing. The DSN Monitor System, under control of the DSN Monitor Operations Group in the SFOF, will be able to complete a Monitor System readiness and verification test in less than 15 min after startup (exclusive of communications setup and on-site calibration time). This test will be based on a standard data set for each facility monitor and on standard known output generated from test formats initialized in the DSN Monitor System.

h. Analysis. All DSN monitor data required for post-pass analysis of DSN performance will be received by the Monitor System in real time. Shipment of individual facility monitor system tapes will not be required to make any DSN monitor report. The DSN Monitor Operations Group in the SFOF shall provide access for each of the DSN systems operations groups to all monitor data concerning the performance of that group's system or any of the quality and accountability data that have been reported to the Monitor System by the data system error detectors.

i. Monitor System records. Monitor System records are shown schematically in Fig. 2. The original data records (ODRs) recorded at each DSS are compatible with both DSS and SFOF computers. They contain time-tagged data on station configuration, instrumentation performance, and spacecraft-dependent data such as ground receiver automatic gain control (AGC), static phase error (SPE), and signal-to-noise ratio (SNR). The records contain information transmitted to the SFOF and facility monitor data of local interest only. These records are available by mail or by replaying from the station.

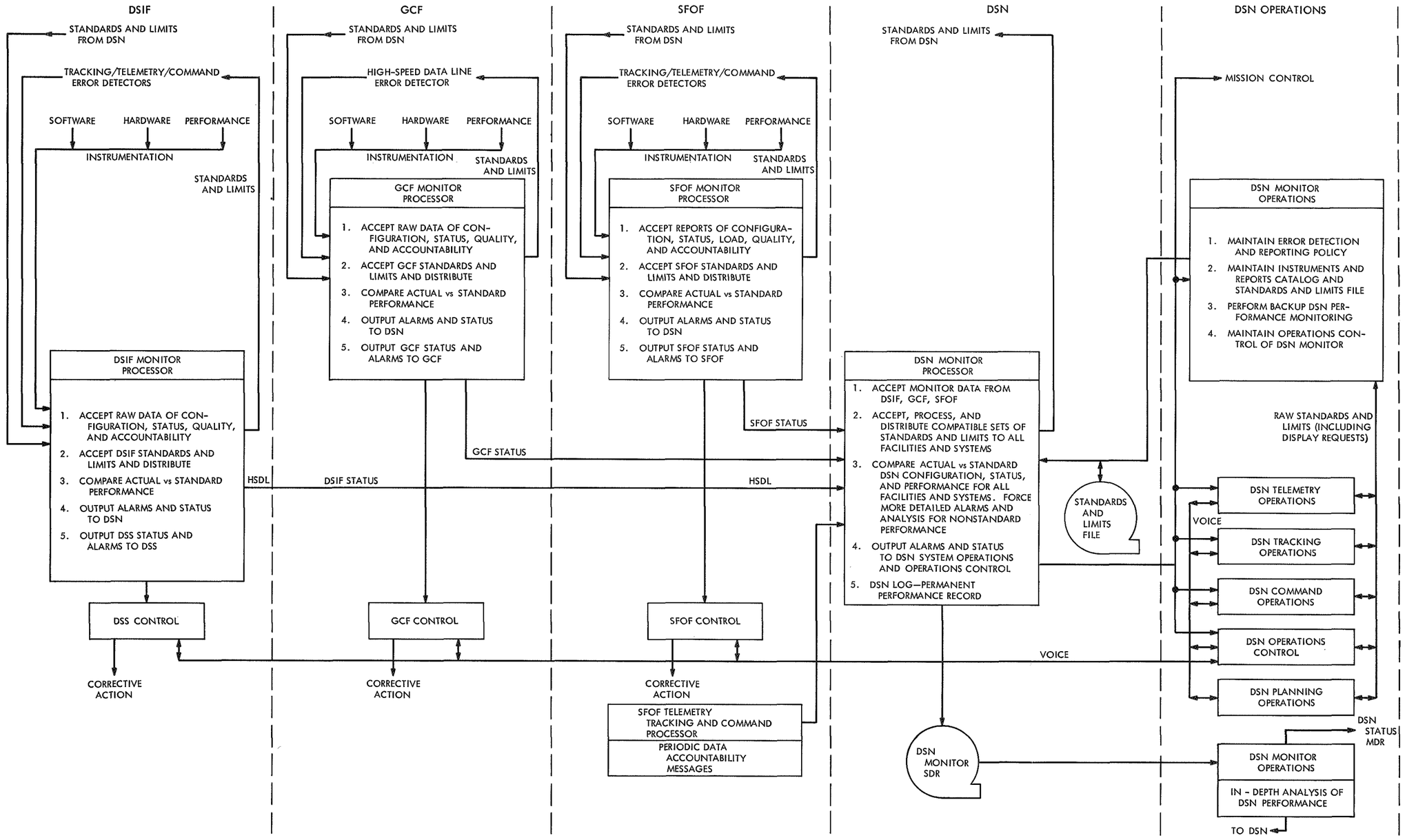


Fig. 1. DSN Monitor System functional flow diagram

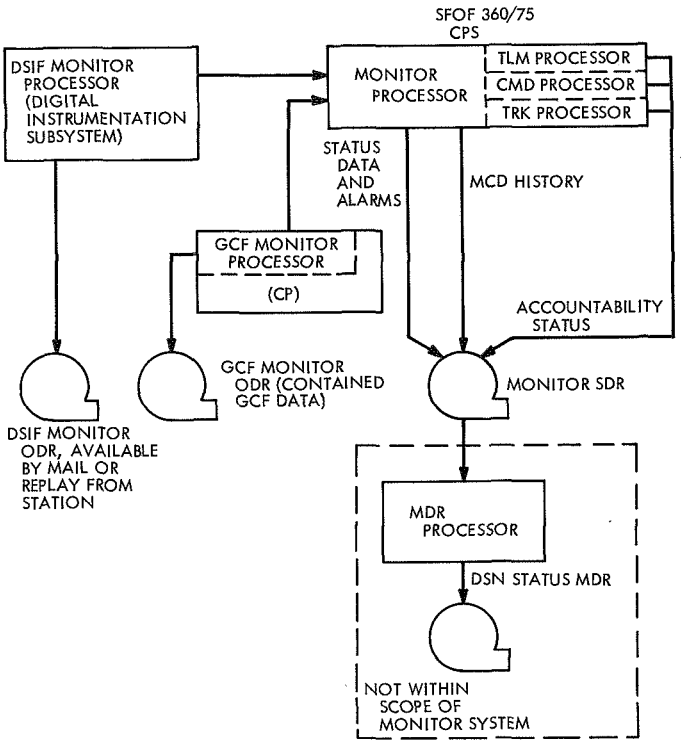


Fig. 2. Monitor data flow

The GCF monitor ODR is processed in the communications processor (CP) in the SFOF. The GCF status data are combined with data from other DSN systems. They contain data on HSDL performance and CP performance and are available by CP recall.

The SFOF monitor ODR is contained on the DSN monitor SDR. The DSN monitor SDR contains the DSIF monitor data transmitted from each DSS to the Central Processing System (CPS) in the SFOF, the GCF monitor data transmitted from the CP to the CPS, and the SFOF monitor data that originate in the SFOF and are processed by the CPS. It also contains a history of monitor criteria data (MCD) set usage and a history of all DSN monitor alarms. The DSN status MDR is obtained by processing of the DSN monitor SDR.

7. Performance Parameters

a. Monitor system support capabilities. The DSN Monitor System will provide capability for processing and displaying up to five DSIF monitor data streams and GCF and SFOF monitor data. Each facility will provide monitor processing for facility monitor and control for Tracking, Telemetry, and Command System status reports. All such monitor processing will be provided simultaneously with the tracking, telemetry, and command data handling.

The DSN Monitor System capability will be sufficient to accept, sort, store, process, and display all monitor data output from the DSS, GCF, and SFOF for all missions supported.

The DSN Monitor System is configured to:

(1) Be initialized in less than 15 min, excluding station calibration and time to activate communication circuits.

(2) Produce a validated System Data Record (SDR) within 12 h after each pass.

(3) Have a response time of less than two days to handle new requirements imposed by new operational instrumentation or display requests.

(4) Have a reliability factor based on a failure recovery time of 10 min for restart.

b. Monitor criteria data. The MCD set is a list of parameters by which the stations' configurations can be determined. Each station is configured to handle each mission in a special way, and each station has a different complement of equipment. To insure that each station is configured in the right way as specified in the sequence of events, each station's parameters are masked against a known set of parameters.

In the event that the station's configuration is not as it should be, an alarm is output to the user and he then corrects the configuration.

The MCD input processor, located in the IBM 360/75 computer at JPL SFOF, will perform the essential task of validating MCD file information, provide for correction and update to the existing MCD file, and structure the MCD master file, used by the MCD activation and assembly program, so that the required data for MCD set construction may be accessed in an efficient manner. The input processor will operate in the batch mode under JPL operating system.

The MCD activation program is the basic interface between the MCD processors and the external world, with the prime responsibility of message processing between the monitor request processor and the MCD

assembly program. The activation program validates input messages and processes operator requests according to their type and origin (i.e., card, 2260, etc.). It provides a series of information displays on MCD subsystem status for operations control. For valid requests, the activation program will request the MCD assembly program to construct a set of MCD data. Depending on the dialogue via the monitor request program, the activation program may cause MCD set creation, modification, deletion, transmission, or simply display status information.

Additionally, inputs will be received by the activation program via the tracking and telemetry processor. These data will consist of tracking standards and limits modification. The activation program will present this information to the MCD processor program for use in processing the DSN raw data file.

The purpose of the MCD sets is to provide facility configuration display by which subsystem analysis and performance can be determined.

In order to obtain a high degree of flexibility, the MCD program is contained in three main files (see Figure 3). These are:

File 1 *Subsystem configuration (hardware)*: A description of the various equipment at the DSIF and their associated switch positions.

File 2 *Configuration mode*: A description of the modes of operations of the various hardware and associated switch positions.

File 3 *Standard MCD set*: A series of numerical matrix listings comprised of hardware configurations listed in file 1 and configuration modes listed in file 2.

The MCD assembly program is responsible for receiving requests for standard MCD set activation, modification of a specified parameter within a standard set, requesting generation of a new MCD set and indexing file 3 with this request, receiving completed request from file 2, and transmission of completed sets to the DSIF for implementation.

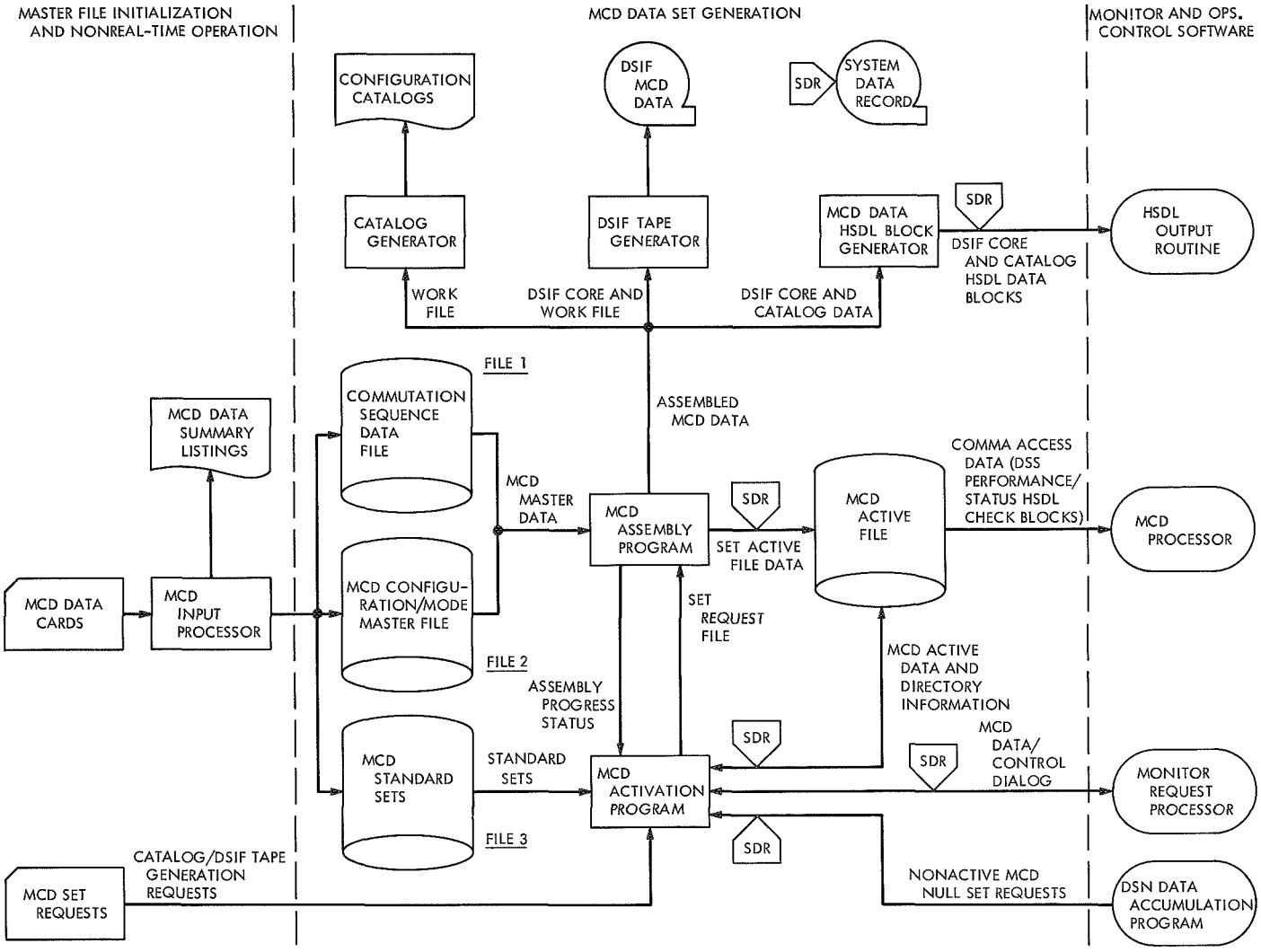


Fig. 3. Monitor and operations control MCD set generator

III. Planetary Flight Projects

MISSION SUPPORT

A. *Mariner* Mars 1969 Extended Operations Mission Support, *K. W. Linnes*

1. Introduction

The *Mariner* Mars 1969 Extended Operations Mission was formally established as a project in January 1970 as a follow-on to the main *Mariner* Mars 1969 Project. The objectives of the mission are: to test relativistic gravitational theories, to measure solar coronal and interplanetary electron density profiles, to improve certain astronomical constants and ephemerides, to demonstrate the high degree of accuracy of the ranging system at 2.6 AU, and to determine the utility and accuracy of the differenced range versus integrated doppler (DRVID) method of charged-particle calibration of metric radio tracking data.

2. Operations

DSN support of the mission is by essentially the same DSN tracking system configuration as used for the main mission. However, to this tracking system configuration an improved experimental ranging system (SPS 37-62, Vol. II, p. 34) and a new experimental 400-kW transmitter have been added, both installed at the 210-ft-diam antenna station (DSS 14). Success of the experiment is critically dependent upon accurately establishing the actual spacecraft orbit in order to distinguish from the apparent relativistic perturbations to the orbit. When the spacecraft are near superior conjunction, the radio signals passing near the sun appear to be delayed due to the relativity effect. An essential correction must be made that tends to mask the effect on the signal by charged particles, both in the earth's atmosphere and interplanetary space. Calibration of this effect is achieved by the DRVID method (SPS 37-62, Vol. II, p. 34). In order to establish the orbit accurately, the project requirements for metric data are:

Date	*Requirement*
11/1/69 to 12/20/69	1 pass per week
12/21/69 to 12/10/70	2 passes per week
12/11/70 to 4/15/70	1 pass every other day
4/16/70 to 5/16/70	1 pass per day (a shared pass with the *Pioneer* Project every other day)
5/17/70 to 7/1/70	1 pass every other day
7/2/70 to 1/1/71	2 passes per week

This requirement for metric data is for use of the 210-ft antenna, the planetary ranging system, and the high-power transmitter.

The 210-ft antenna was returned to service early in March 1970 and tracking was resumed, using the 20-kW transmitter. The high-power transmitter became operable late in March, with power levels limited to 200 kW. The transmitter was pressed into service at the earliest possible date, and higher levels will not be feasible until relaxation of the tracking schedule makes sufficient time available to properly debug the transmitter power supply and to make other adjustments.

3. Flight Support

The DSN tracking system provided doppler data in precisely the same manner as in the main mission. Figure 1 shows the tracking configuration at the DSIF; Fig. 2 shows the configuration for transmitting the data through the GCF and processing it in the SFOF. Figure 3 shows the functions performed by the tracking system in the SFOF. The DSN provided the project with a tape master data record that was the output of the tracking data processor run on the IBM 7094 computer. R&D ranging data was handled separately due to the manual operations necessary. As currently implemented, the data can only be outputted on a page printer. After being examined by the ranging experimenter for blunder points, the data is reformatted to be machine-processible and is added to the master data file (Fig. 4).

Tracking coverage was provided generally in accordance with the requirements. Table 1 summarizes coverage provided for the last several months. DSS 62 at Robledo, Spain, was used to transmit commands to the spacecraft necessary to update the central computer and sequencer (CC&S) and prepare the spacecraft for the tracking pass by the 210-ft station. In March, the command link threshold was reached for *Mariner VI*; it was

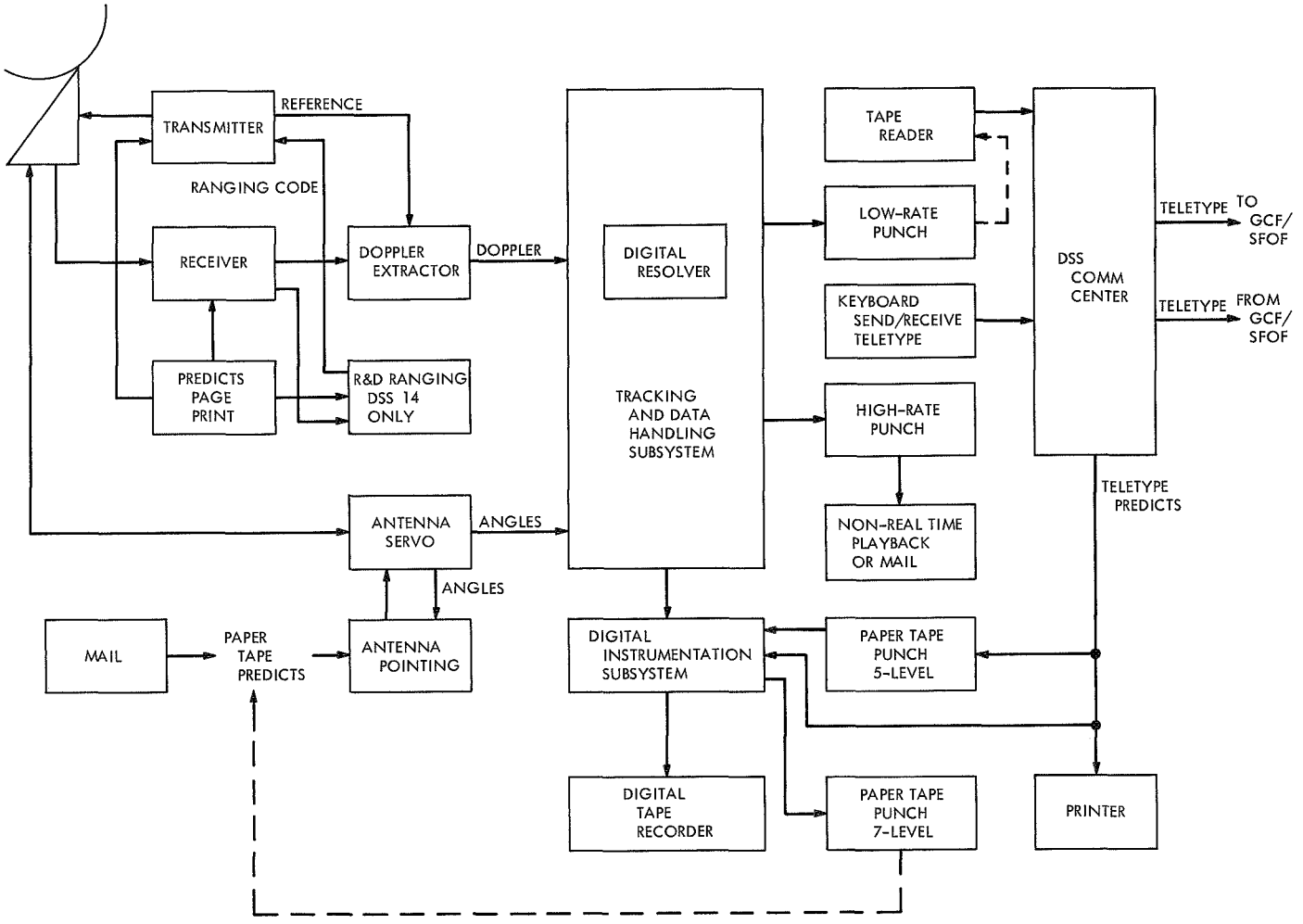


Fig. 1. DSIF tracking configuration

Table 1. DSN coverage of *Mariners VI* and *VII*

Month (1970)	Tracking coverage			Deep space station	Transmitter power, kW	Remarks
	Spacecraft	Number of passes	Number of commands sent			
February	*Mariner VI* *Mariner VII*	9 11	0 0	12	10	DSS 14 out of service for tri-cone and high-power transmitter installation. Good two-way data obtained in the listen-only mode (transmitter off) for one round-trip light time (about 43 min), followed by about 50 min of one-way data received while transmitting to the spacecraft in the diplex mode.
March	*Mariner VI* *Mariner VII*	15 15	28 20	62 and 14	20	Tri-cone installation at DSS 14 completed. High-power transmitter not ready. Command by DSS 62 to conserve coverage time from DSS 14.
April	*Mariner VI* *Mariner VII*	22 19	174 0	62 and 14	200	High-power transmitter output limited to 200 kW. DSS 62 support discontinued March 18 for *Mariner VII* and April 21 for *Mariner VI*; command link threshold reached. *Mariner VI* superior conjunction on April 29.
May	*Mariner VI* *Mariner VII*	27 28	80 0	14	200	20-kW transmitter used on three days when high-power transmitter was down. Superior conjunction for *Mariner VII* on May 10.
June	*Mariner VI* *Mariner VII*	15 16	3 49	14	200	20-kW transmitter used on three days.

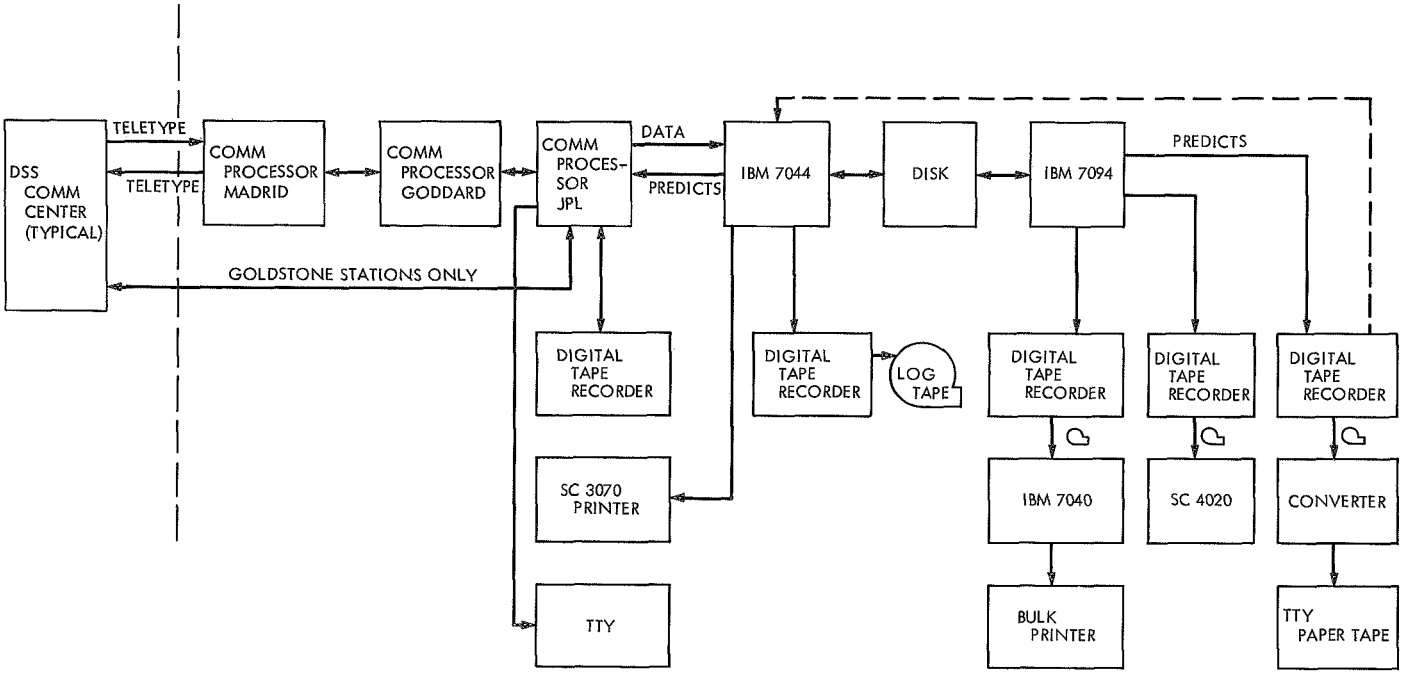


Fig. 2. SFOF/GCF tracking configuration

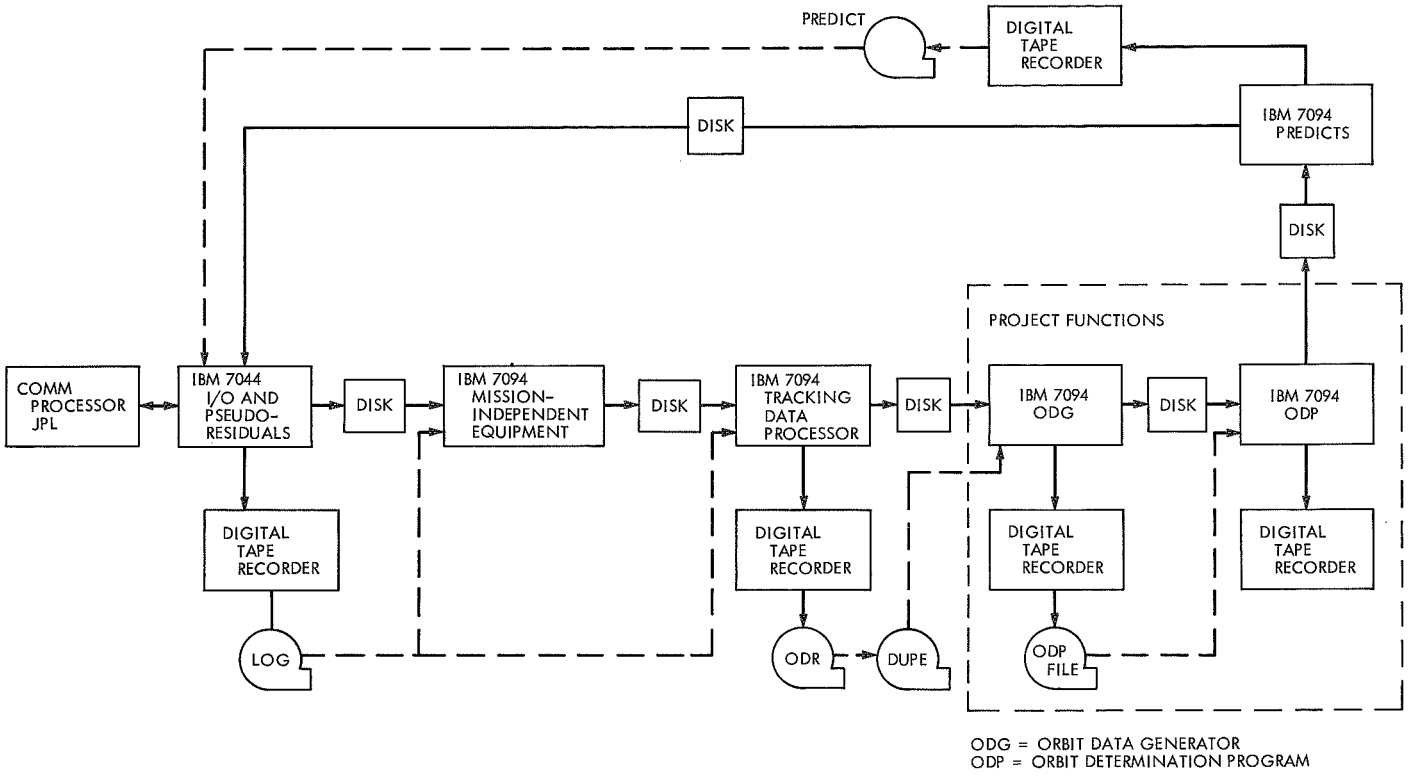


Fig. 3. SFOF tracking data functions

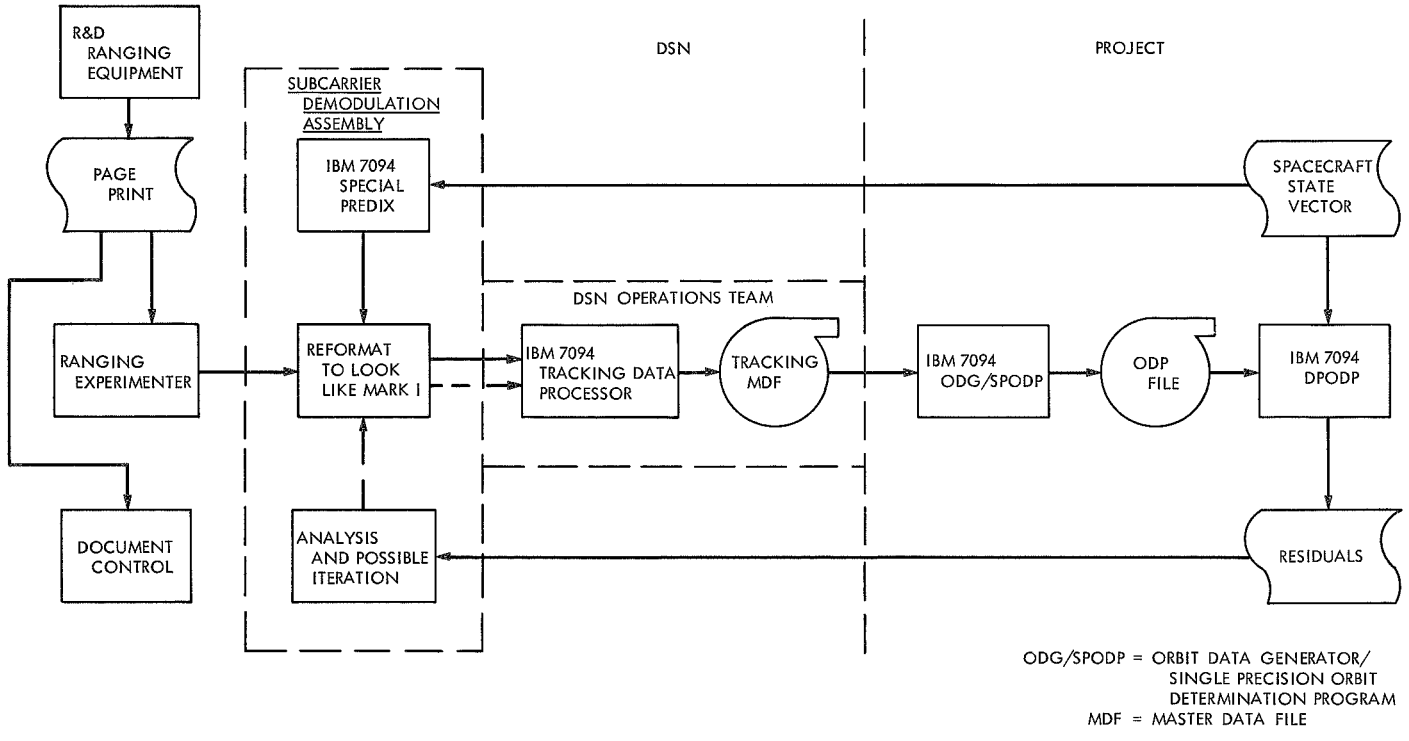


Fig. 4. Mu R&D ranging data handling

reached for *Mariner VII* in April. Since the spacecraft could no longer be commanded from an 85-ft antenna station, DSS 62 was no longer used.

Prior to the availability of the 210-ft antenna, tracking was accomplished at DSS 12. The 85-ft antenna was equipped with an ultra low noise cone which could be switched into either a listen-only or diplex mode. After two-way lock had been established, the transmitter was shut off, and the station continued to receive good two-way doppler data for one roundtrip light time (about 43 min). The transmitter was then turned on for about 50 min; one-way doppler data was received under noisier conditions for another roundtrip light time; two-way lock was established, and the procedure was repeated. This "checkerboard" pattern of two-way tracking increased the number of good data points and reduced noise in the data.

4. Tracking at Superior Conjunction

The crucial part of the experiment and mission occurred in the vicinity of superior conjunction in the latter part of April and first part of May. As the radio signal, traveling about 2.5 AU from spacecraft to earth, passed closer and closer to the sun, the ground system noise temperature increased as successively higher and higher antenna side lobes scanned across the sun. Figures 5 and 6 show the predicted and measured system noise temperatures for *Mariners VI* and *VII*. The predicted values were estimated from average antenna patterns. The measured values show not only the increase with nearness to the sun but also a variation in the values taken on one day. The diurnal variation is due to the fact that the 210-ft antenna employs an az-el mount. As the antenna tracks the spacecraft through its pass, its

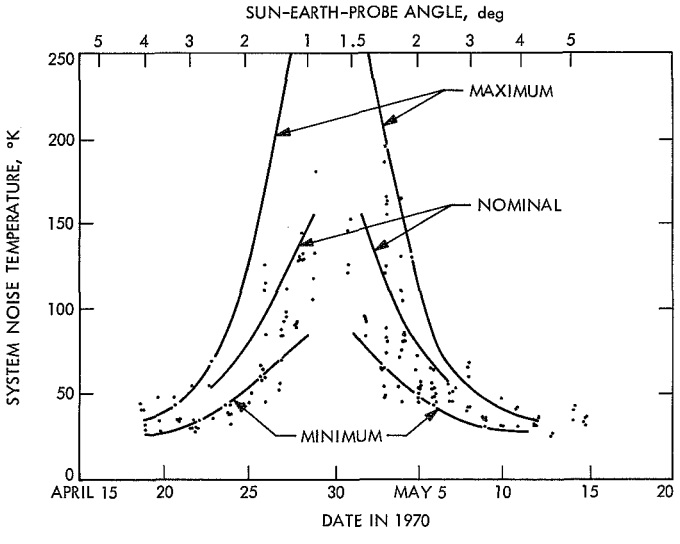


Fig. 5. *Mariner VI*, measured system noise temperature about superior conjunction at DSS 14

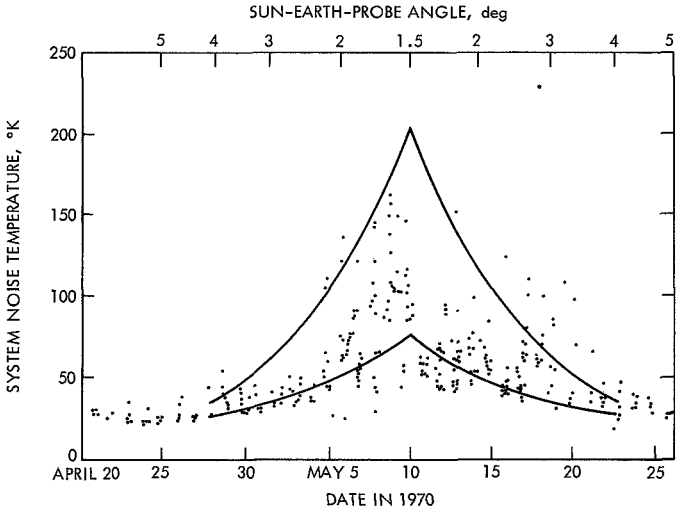


Fig. 6. *Mariner VII*, measured system noise temperature about superior conjunction at DSS 14

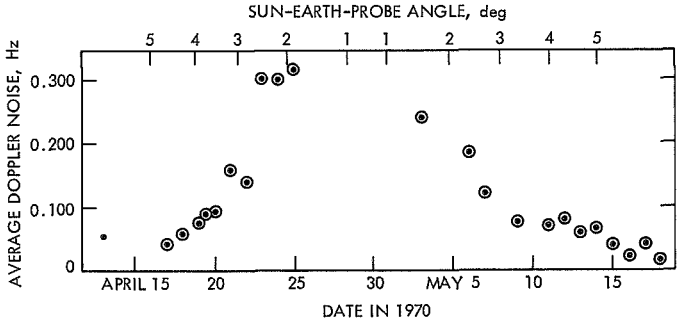


Fig. 7. *Mariner VI* average doppler noise at superior conjunction

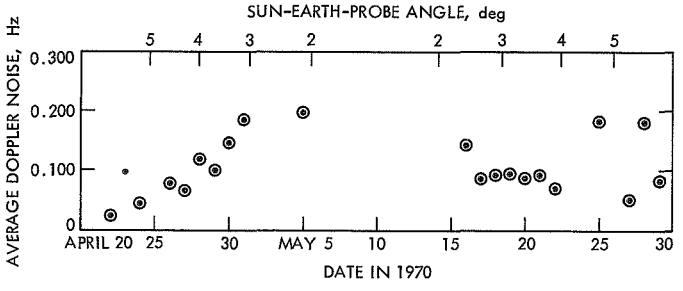


Fig. 8. *Mariner VII* average doppler noise at superior conjunction

principal E–H planes rotate slightly with respect to celestial coordinates. Since the side lobes are not cylindrically symmetrical about the antenna axis, the sun moves through the antenna pattern slightly, and its contribution to the system noise temperature varies according to the side lobe structure.

As the system noise temperature rose, so did the average doppler noise as the receivers maintained lock with increasing difficulty. Figures 7 and 8 show how the average doppler noise increases as the sun–earth–spacecraft angle decreases until the receivers can no longer maintain lock a significant percentage of the time.

Throughout the tracking coverage, the experimental ranging system was employed. Not all attempts at locking of the ranging system were successful, however. Figures 9 and 10 show the degree of success in the vicinity of superior conjunction. Successfully acquired ranging points thereupon supplied data to the project according to the flow shown in Fig. 4.

Since superior conjunction, the amount of tracking coverage has been reduced in accordance with the requirements. The data has become less noisy, and the acquisition of ranging data and good two-way doppler data has resumed a routine nature.

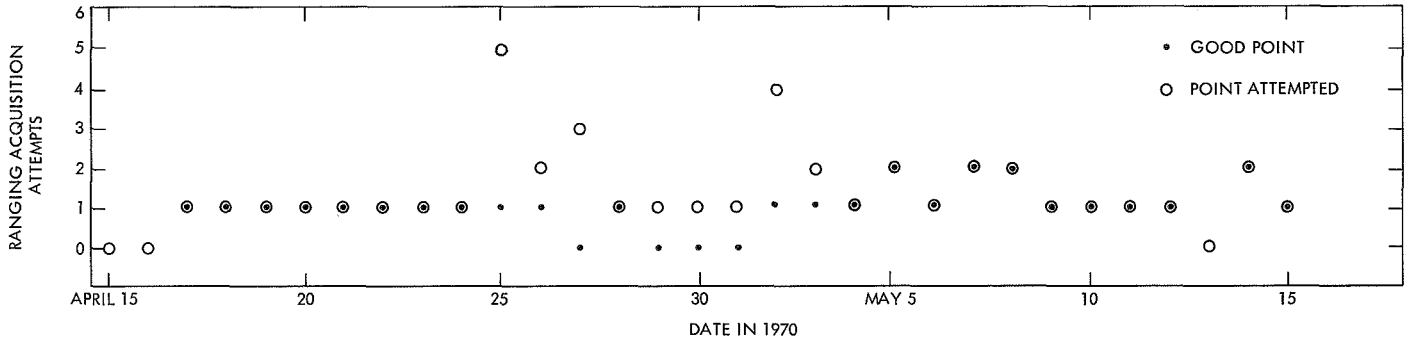


Fig. 9. *Mariner VI* ranging attempts at superior conjunction

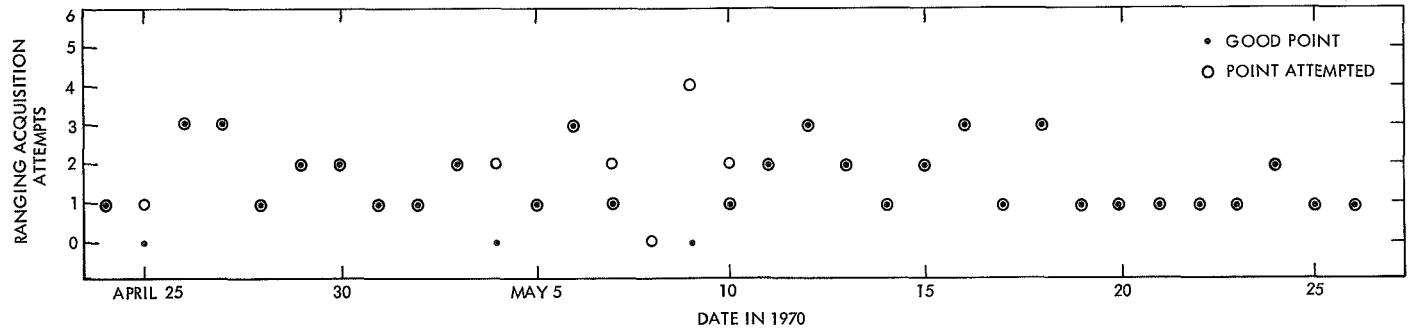


Fig. 10. *Mariner VII* ranging attempts at superior conjunction

B. *Mariner* Mars 1971 Mission Support,

R. P. Laeser

The calibration of radio metric data (generated by the DSN while tracking a spacecraft) to compensate for the effects of propagation phenomena and the variations of a DSS location in space is a DSN responsibility in some cases; in other cases, the DSN provides analytic calibration assistance to each supported project. This effort, called tracking system analytic calibration (TSAC) has been formalized for support starting with the *Mariner* Mars 1971 mission and is described in this article.

1. The Problem

The calibration of radio metric data to account for certain observer-related phenomena has become a necessity if more stringent navigation requirements are to be met. In order to extract an optimum amount of information from this data, it must be processed to remove, as much as possible, any bias or random noise effects which may have corrupted it. Currently, several error sources have been identified as contributing significantly to the composition of the noise signature. These observer-related errors can be classified into two main categories:

(1) Errors in locating the observer (tracking station) with respect to the inertial space: DSS locations with respect to the earth's crust, polar motion, and variation in earth rotation rate (A.1 − UT1).

(2) Errors caused by the transmission media that corrupt or distort the actual information content of the observation: charged particle effects (ionosphere and space plasma) and neutral particle effects (troposphere refraction).

The refractive effect of the troposphere causes the retardation and bending of an electromagnetic beam. This makes the observation appear as if the beam has traveled through a longer distance. This difference is such that radio measurements of range and doppler are significantly affected. The model currently used to measure the effects of refractivity assumes a fixed distribution of refractivity in the atmosphere over each site. Actually, the refractivity distribution goes through seasonal and diurnal changes which, if not properly detected, can cause errors in estimating the location of a spacecraft. To eliminate this error, two tropospheric models will be employed which account for the different seasonal variations of the wet (includes water vapor) and the dry components of the troposphere.

The radio signals traveling between a tracking station and a spacecraft pass through the charged-particle media of the interplanetary space plasma, the ionosphere of the earth, and possibly the ionosphere of other planets. The interaction between the radio signal and the charged particles in the medium causes an increase in the phase velocity and a decrease in the group velocity.

The earth's ionosphere is caused by ultraviolet light from the sun ionizing the upper atmosphere. Consequently, the ionosphere above a fixed location on earth increases and decreases with, roughly, a diurnal period. If the ionospheric effect cannot be measured or modeled, it cannot be distinguished from errors in tracking station location and may result in significant errors for in-flight orbit determination.

The space plasma effect is due to streamers of electrically charged particles emanating from the sun that intersect the radio tracking beam to the spacecraft. The effect on the radio beam is similar to that described for the ionosphere, except that the electron activity in interplanetary space is more of a steady-state nature with fluctuations occurring mainly as the radio beam enters and leaves a particular streamer or as a result of sudden outbursts of plasma from solar flares.

The *total* charged-particle effect, that is, the effect of both the ionosphere and space plasma, can be measured by techniques such as dual-frequency transmission and the comparison of the doppler and ranging signals. This latter method will be used for the *Mariner* Mars 1971 mission. It involves no extra hardware for the spacecraft.

This method makes use of the differential between phase and group effects in a charged-particle medium. Since doppler depends on phase propagation and ranging on group propagation, a comparison of differenced range versus integrated doppler yields the time rate of change of the error due to variations in the total columnar charged-particle content. This quantity can be used directly to correct doppler for charged-particle effects.

2. Requirements

The *Mariner* Mars 1971 Project has placed the following accuracy requirements on the DSN:

DSS location[a]	$\sigma_{r_s} = 0.5$ m and $\sigma_\lambda = 1.0$ m
Polar motion[a]	$\sigma_x = 0.7$ m and $\sigma_y = 0.7$ m
Variation in earth rotation rate[a]	2.5 ms
Neutral particle effects	0.5 m
Charged particle effects	1.0 m

[a]The DSS provides analytic calibration assistance to the project in order to achieve these accuracy requirements.

3. Planned TSAC Capabilities

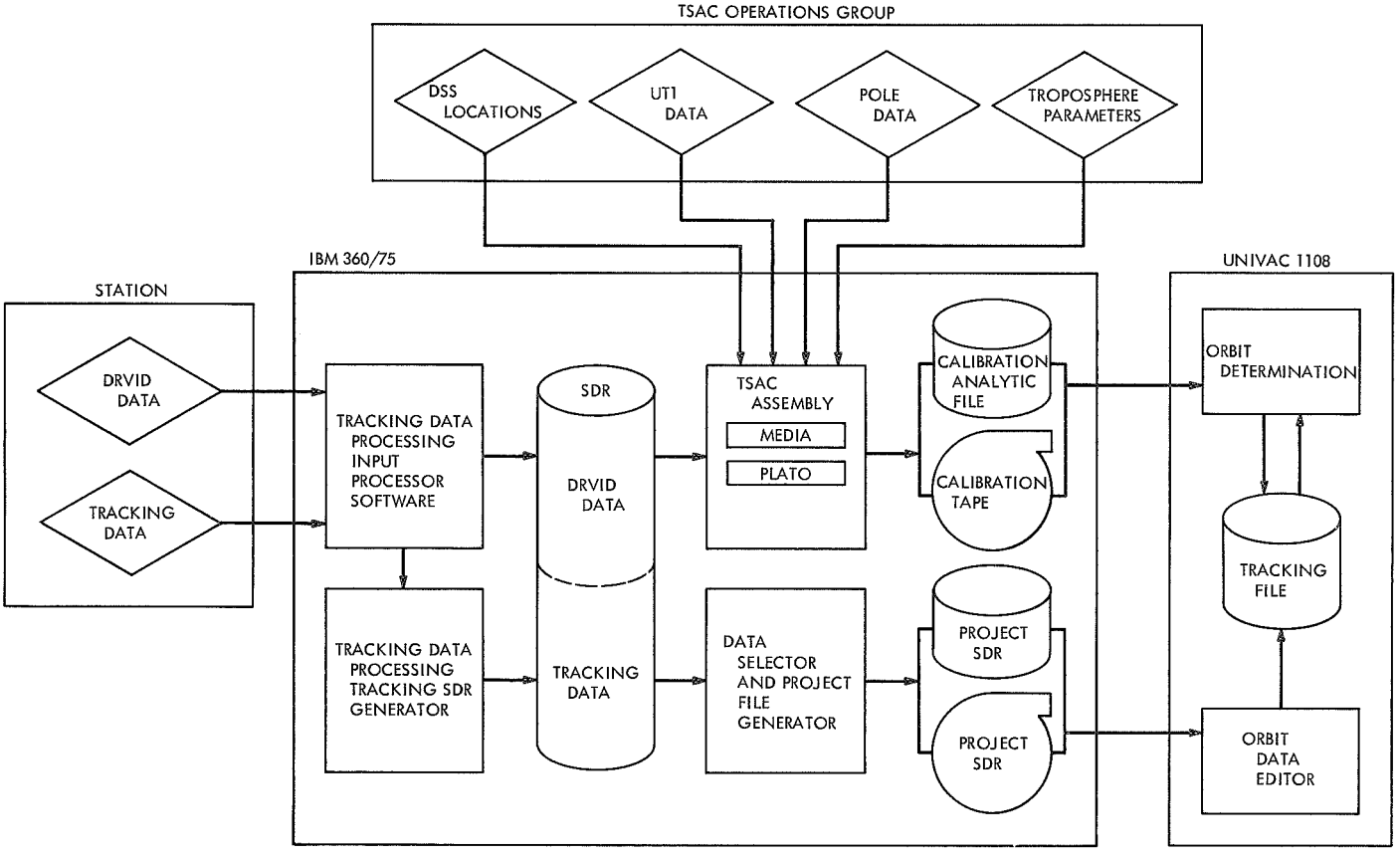
The calibration software will be termed the tracking system analytic calibration software assembly. This assembly will consist of two subassemblies residing in the IBM 360/75:

(1) The platform observables subassembly (PLATO)

(2) The transmission MEDIA subassembly (MEDIA)

The first will provide calibrations for DSS location, polar motion, and the variation in the earth's rotation rate (A.1 − UT1). The second will provide calibrations for the troposphere and a combined calibration for the ionosphere and space plasma using the differenced range versus integrated doppler (DRVID) technique. Figure 1 depicts the interfaces that the TSAC assembly will have to satisfy to meet its objectives. As shown in the figure, the data input will be from two sources; the DSSs will provide the DRVID observable, and the tracking operations group of the DSN operations organization will input the required troposphere parameters, time and polar motion data, and DSS locations which have been previously prepared.

The DRVID observable will be obtained via teletype or high-speed data from planetary ranging systems at DSS 14 and one 85-ft-diam antenna DSS (probably DSS 41). It will be placed on the system data record (SDR) by the tracking data processor. The MEDIA subassembly will access the SDR to obtain the unprocessed DRVID data, compute the doppler corrections, fit them with a polynomial, and place the polynomial coefficients on a calibration file accessible to the double precision orbit determination program in the Univac 1108 computer. The DRVID computation compares the differenced range to integrated doppler and provides a measure of exactly that quantity required to calibrate doppler for



SDR = SYSTEM DATA RECORD
DRVID = DIFFERENCED RANGE VERSUS INTEGRATED DOPPLER
TSAC = TRACKING SYSTEM ANALYTIC CALIBRATION

Fig. 1. TSAC tracking orbit software subsystems flow diagram

charged-particle effects. This quantity is usually expressed in terms of the range change error $\Delta\rho_e$:

$$\Delta\rho_e(t) = \frac{C}{2}\left\{R(t) - R(t_0) - \frac{K}{f_q} \times [D(t) - D(t_0) - f_b(t - t_0)]\right\}$$

where

$R(t_0)$ = spacecraft range at initial time t_0, sec

$R(t)$ = spacecraft range at arbitrary time t, sec

$D(t_0)$ = cumulative doppler count at t_0, cycles of S-band

$D(t)$ = cumulative doppler count at t, cycles of S-band

$K = 96\left(\frac{240}{221}\right)^{-1}$

f_q = transmitter reference frequency, Hz (nominally 22 MHz)

f_b = bias frequency of doppler extractor 1.0×10^6 Hz

c = speed of electromagnetic propagation in vacuum 2.997926×10^8 m/s

The MEDIA subassembly will also compute a troposphere polynomial for every pass of a spacecraft over a DSS. This polynomial will define the zenith range correction to the radio beam as a function of time. The double precision orbit determination program will evaluate the zenith range correction polynomial at specific times and map the correction to the spacecraft's line of sight. The approximation to the zenith range error which will be used is:

$$\Delta\rho_z = \Delta P_0\left[\frac{R}{g}\right] + \left\{\frac{C_1 C_2 (RH)_s}{\gamma}\right\}\left\{\frac{1 - \frac{C^2}{T_0}}{B - AC}\right\} \times \exp\left\{\frac{AT_0 - B}{T_0 - C}\right\}$$

where

P_0 = surface pressure, mbar

γ = temperature lapse rate, °K/10^3 ft

T_0 = extrapolated surface temperature, °K

$(RH)_s$ = surface relative humidity, % of 1.0

$A = 77.6$

$B = 2034.28 \ln 10$

$C = 38.45$

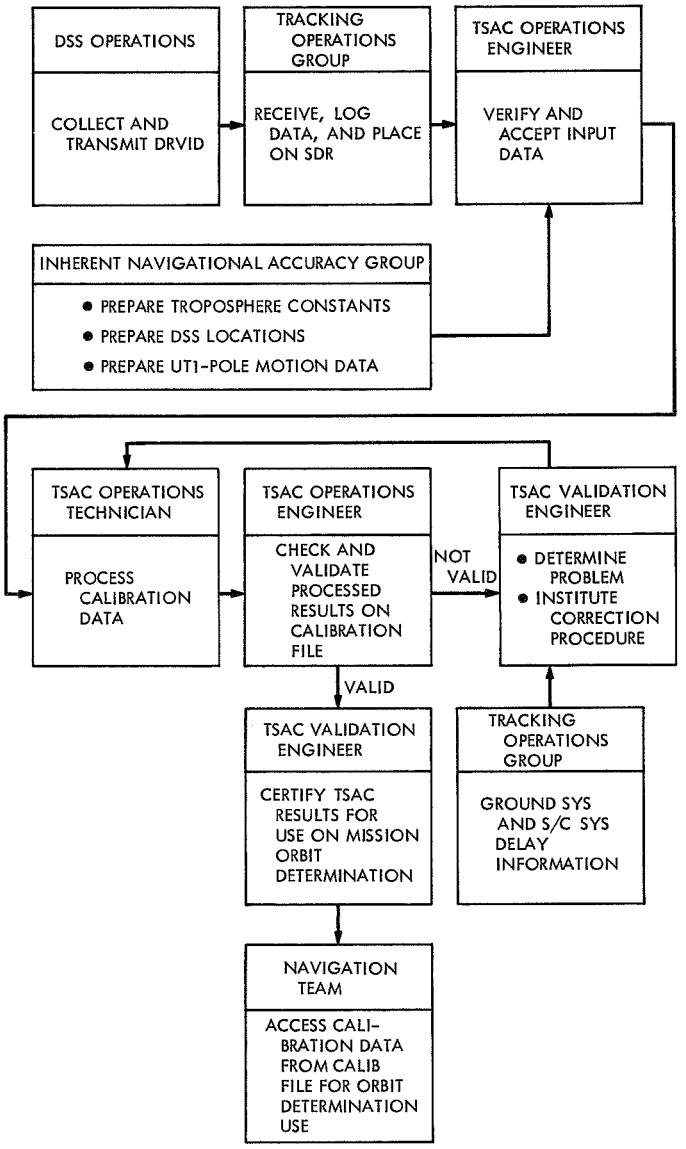
$g/R = 34.1$°C/km

g = gravitational force

R = gas constant

$C_1 = 77.6$

$C_2 = 29341.0$



TSAC = TRACKING SYSTEM ANALYTIC CALIBRATION
SDR = SYSTEM DATA RECORD

Fig. 2. Planned TSAC operations

The first part of the expression being due to the dry component and the second part to the wet component. A set of constants P_0, γ, T_0, $(RH)_s$ will be provided at distinct times, and $\Delta\rho_z(t)$ will be computed for each of those times.

It is expected that the TSAC effort will aid the DSN tracking system in providing the following accuracies:

DSS locations	$\sigma_{r_s} = 0.6$ m and $\sigma_\lambda = 2.0$ m at Goldstone DSCC $\sigma_{r_s} = 1.1$ m and $\sigma_\lambda = 3.0$ m at other DSSs
Polar motion	$\sigma_x = 0.7$ m and $\sigma_\gamma = 0.7$ m
Variation in earth rotation rate	5 ms
Charged particle calibration	0.5 to 0.75 m
Neutral particle calibration	0.5 m

4. Phased TSAC Implementation

The PLATO subassembly will probably operate as a non-real-time IBM 360/75 batch program. If manpower constraints permit, PLATO will be modified and integrated into the IBM 360/75 real-time system.

MEDIA is being programmed to operate under the IBM 360/75 real-time system. It will be completed for model 2 (launch/cruise support) except that a separate (not SDR) output file will be generated. The full capability will be available in model 3.

5. Planned TSAC Operations

The tentatively planned TSAC operations concept is depicted in Fig. 2. The TSAC operations engineer and TSAC operations technician are part of the DSN tracking system operations group. The TSAC validation engineer is a development engineer who works with the tracking system operations group.

IV. Tracking and Navigational Accuracy Analysis

ADVANCED ENGINEERING

A. Introduction, *T. W. Hamilton and B. D. Mulhall*

The DSN Inherent Accuracy Project was formally established by the DSN Executive Committee in July 1965. The objectives of the project are:

(1) Determination (and verification) of the inherent accuracy of the DSN as a radio navigation instrument for lunar and planetary missions.

(2) Formulation of designs and plans for refining this accuracy to its practical limits.

Achievement of these goals is the joint responsibility of the Telecommunications and Mission Analysis Divisions of JPL. To this end, regular monthly meetings are held to coordinate and initiate relevant activities. The project leader and his assistant (from the Mission Analysis and Telecommunications Divisions, respectively) report to the DSN Executive Committee, and are authorized to task project members to (1) conduct analyses of proposed experiments, (2) prepare reports on current work, and (3) write descriptions of proposed experiments. The project is further authorized to deal directly with those flight projects using the DSN regarding data-gathering procedures that bear on inherent accuracy.

The various data types and tracking modes provided by the DSIF in support of lunar and planetary missions are discussed in SPS 37-39, Vol. III, pp. 6–8. Technical work directly related to the Inherent Accuracy Project is presented in SPS 37-38, Vol. III, and in subsequent *Deep Space Network* SPS volumes, and is continued in the following sections of this volume.

For most upcoming planetary missions, such as *Mariner* Mars 1971, the tightest bounds on the allowable errors for a number of parameters arise from the navigational accuracy requirements during encounter support. In particular, encounter navigational accuracy is most sensitive to error sources that cause a diurnal signature on the radio tracking data (SPS 37-39, Vol. III, pp. 18–24). These sources of error are of two classes: (1) those parameters that define the locations of the DSS in inertial space, and (2) those phenomena that directly affect the DSS tracking data. The first category includes the locations of the DSS with respect to earth's crust; Universal Time (UT1); polar motion (the motion of the earth's crust with respect to the spin axis); precession and nutation (orientation of the earth's spin axis with respect to inertial space); and the ephemerides of the earth, moon, and target body. Of these, uncertainties in the first three are currently the

major limitations to the encounter support of navigation accuracy.

The dominant sources of error in the second category are those affecting the tracking data directly. These include frequency system instability, electrical phase path variations (through both the spacecraft and the DSS), and the transmission media (the troposphere and the charged particles in the ionosphere and space plasma).

The two articles in this chapter are concerned with the effects of the transmission media on DSN navigational accuracy. The first article, *Section B,* concentrates on tropospheric effects on navigation errors. This article discusses the temporal variation of tropospheric errors as measured by radiosonde balloon data. A cursory examination of surface weather measurements as a calibration technique and their correlations with the zenith range effect is also discussed.

The second article, *Section C,* deals with the effects of the earth's ionosphere on tracking data. The sensitivity of navigation errors to the ionospheres for differences in seasons, solar cycle, and mission geometries (inner vs outer planet missions) is described.

In the investigation of transmission media effects on navigation, the need for data not normally collected by the DSN has arisen. Consequently, the Jet Propulsion Laboratory has worked with outside agencies in obtaining and exchanging atmospheric data. Examples of this are the studies discussed in *Sections B* and *C*. The radiosonde data used in tropospheric studies were obtained from the National Bureau of Standards. The ionospheric data used in the study discussed in *Section C* were obtained from the Center for Radar Astronomy, Stanford Electronics Laboratories, and from devices installed at the Goldstone DSCC. Through Stanford's cooperation with JPL, a Faraday rotation polarimeter designed and built by Stanford is currently on loan to JPL and is being used to monitor the ATS-1 satellite to obtain ionospheric calibration data for the Goldstone complex.

B. Variations in the Zenith Tropospheric Range Effect Computed From Radiosonde Balloon Data, *V. J. Ondrasik and K. L. Thuleen*

1. Introduction

For many years it has been recognized that the passage of a tracking signal through the troposphere will significantly corrupt the data used in determining the orbit of a distant spacecraft. At the present time (July 1970), a tropospheric refractivity model, which is independent of time, is being used in attempts to reduce the tropospheric errors. Some idea of the temporal behavior of errors in radio tracking data due to the troposphere is obtained by calculating the tropospheric zenith range effect from measured refractivity profiles collected during 1967. In addition, a cursory examination of the surface weather measurements is undertaken to see if it may be worthwhile to try to predict the zenith range effect from such measurements.

2. Zenith Range Effect

It is convenient to think of the range measurement ρ, divided by c, as the time it takes a radio signal to travel from the spacecraft to the tracking station. This concept is expressed by

$$\rho = c \int_{\text{path}} \frac{ds}{v(s)} \tag{1}$$

where

ds = an element of the ray path over which the signal travels

$v(s)$ = the velocity of the radio signal on the path

c = the velocity of light in a vacuum

In general the presence of the troposphere will cause the path to curve, as well as retarding the velocity of the radio

signal. However, when the spacecraft is at the zenith, there will be no bending and Eq. (1) reduces to

$$\rho_z = \rho\,(\text{zenith}) = \int_0^\infty \frac{c}{v(h)}\,dh \tag{2}$$

where h is the height above the station. By definition

$$n = \frac{c}{v} = 1 + N \times 10^{-6} \tag{3}$$

where

n = index of refraction

N = refractivity

Clearly, then, the zenith range effect $\Delta\rho_z$, which is the difference between the measured and geometric ranges to an object at the zenith, is given by

$$\Delta\rho_z = \rho_z - \rho_z\,(\text{geo.}, v = c) = 10^{-6}\int_0^\infty N(h)\,dh \tag{4}$$

Thus, once the refractivity profile $N(h)$ is known, $\Delta\rho_z$ may be calculated from Eq. (4).

3. Tropospheric Refractivity Profile

The refractivity of the troposphere is commonly determined from the following equation (Ref. 1):

$$N = \frac{77.6}{T'}\left(P + \frac{4810e}{T'}\right) \tag{5}$$

where

P = pressure, mb

T' = temperature, °K

e = vapor pressure, mb

and

$$e = \left[6.1 \times \exp_{10}\left(\frac{7.4475T}{234.7 + T}\right)\right] \times \frac{R}{100}$$

where

R = relative humidity, %

T = temperature, °C

$\exp_{10} y = 10^y$

The first and second terms of N will be referred to as the dry and wet refractivities, respectively. An approximation to the errors in N resulting from the use of Eq. (5), due both to the errors in the equation itself and to errors in the meterological measurement(s), is given by Bean and Dutton (Ref. 1), and is summarized in Table 1.

Table 1. Errors in the equation for refractivity

Error source	Ratio of rms error in refractivity to refractivity[a]
Equation (5)	0.5%
Surface weather observations (±1 mb, ±0.1 °C, ±1% R)	0.2%
Radiosone observations (±2 mb, ±1 °C, ±5% R)	1.3%

[a] p = 1013 mb, T = 15 °C, R = 60%.

Presently the best source of obtaining the necessary weather measurements to determine the refractivity profile (i.e., N versus h) from Eq. (5) is by using radiosonde balloon observations. The data which are generally retained from these observations are the pressure, temperature, and humidity at both the standard pressure levels and at any pressure level where the observer notices any significant changes in the data from a nominal. An example of this type of data is shown in Table 2. Generally the overseas data are not as complete and terminate at approximately 350 mb.

A determination of the height above sea level corresponding to the standard and significant pressure levels may be found by using the hydrostatic equation and a particular gas law. The computation of the heights in this way and an interpolation between the resulting points to obtain the pressure, temperature, vapor pressure, and relative humidity at the surface and every 1000 ft above mean sea level is done in the program CNVTWR. This program is a slight modification of the program AFFTC which is used at Edwards AFB and was supplied by Capt. G. Fredricks, U.S. Air Force, 6th Weather Wing. An abbreviated example of the output of CNVTWR is given in columns 2 through 6 of Table 3.

These sets of the pressure temperature and vapor pressure for every 1000 ft of altitude may be substituted into Eq. (5) to give the corresponding total, dry, and wet refractivities. The calculations for determining the refractivity profiles in this manner were performed and an

abbreviated example is given in the seventh, eighth, and ninth columns of Table 3.

4. Calculation of the Zenith Range Effect

According to Eq. (4) the zenith range effect may be written as

$$\Delta\rho_z \doteq 10^{-6}\int_0^{\infty} N(h)\,dh \tag{6}$$

To accommodate the results of the previous section, the above integral is replaced by the following summation:

$$\Delta\rho_z = \sum_{i=1} \Delta\rho_i \tag{7}$$

where

$$\Delta\rho_i = 10^{-6}\int_{h_{i-1}}^{h_i} N(h)\,dh \tag{8}$$

and i is the index of the refractivity profile table. Before the integral in Eq. (8) can be integrated, it is necessary to specify $N(h)$. This study calculates $\Delta\rho_i$ by assuming that $N(h)$ assumes the tabular values N_{i-1} and N_i at h_{i-1} and h_i, respectively, and has an exponential behavior with inverse scale height, B_i, between these two heights. For this type of atmosphere, Eq. (8) may be written as

$$\Delta\rho_i = 10^{-6}\frac{1}{B_i}\left(N_{i-1} - N_i\right) \tag{9}$$

where

$$B_i = \frac{\ln\left[N(h_{i-1})\right] - \ln\left[N(h_i)\right]}{h_i - h_{i-1}} \tag{10}$$

An example of the zenith range effect and inverse scale heights computed according to Eqs. (7–10) is given in the last nine columns of Table 3.

5. Data

To determine what type of values the zenith range effect may assume, weather data profiles for 1967 were obtained from the radiosonde balloon data site nearest to each of the tracking stations. The Edwards AFB data was supplied by Capt. G. Fredricks and the overseas data was obtained from the National Weather Records Center in Asheville, N. C., in cooperation with Mr. Wayne Hensley and Mr. Richard Davis of that agency.

Table 2. Standard and significant pressure level radiosonde balloon data observed at Edwards AFB, August 2, 1968

Standard pressure level data			Significant pressure level data		
Pressure, mb	Temperature, °C	Relative humidity, %	Pressure, mb	Temperature, °C	Relative humidity, %
931.2[a]	19.0	49.7	931.2[a]	19.0	49.7
900.0	23.5	46.3	920.0	23.2	48.5
850.0	22.2	42.9	880.0	23.9	44.1
800.0	19.2	41.3	802.0	19.4	40.9
750.0	13.9	51.7	709.0	9.4	60.7
700.0	8.9	58.0	625.0	5.3	33.4
650.0	6.5	41.9	454.0	−10.3	35.2
600.0	3.3	33.7	400.0	−19.0	36.0
550.0	−1.0	34.2	289.0	−39.5	36.5
500.0	−5.6	34.7	233.0	−51.7	9.0[b]
450.0	−10.9	35.3	163.0	−63.0	9.0[b]
400.0	−19.0	35.0	112.0	−68.6	9.0[b]
350.0	−27.4	36.2	100.0	−66.5	9.0[b]
300.0	−37.2	36.4	86.0	−68.6	9.0[b]
250.0	−47.7	18.0	70.0	−60.4	9.0[b]
200.0	−56.5	9.0[b]	23.0	−49.7	9.0[b]
175.0	−60.7	9.0[b]	10.0	−37.9	9.0[b]
150.0	−64.2	9.0[b]	8.0	−40.0	9.0[b]
125.0	−66.9	9.0[b]	4.0	−26.1	9.0[b]
80.0	−65.7	9.0[b]			
70.0	−60.4	9.0[b]			
60.0	−59.0	9.0[b]			
50.0	−57.2	9.0[b]			
40.0	−55.1	9.0[b]			
30.0	−52.3	9.0[b]			
25.0	−50.5	9.0[b]			
20.0	−47.8	9.0[b]			
15.0	−43.7	9.0[b]			
10.0	−37.9	9.0[b]			
7.0	−37.3	9.0[b]			
5.0	−30.6	9.0[b]			
4.0	−26.1	9.0[b]			

[a]Surface.

[b]Calculated values.

Table 4 lists the location of the radiosonde balloon sites, the distance from the nearest tracking station, and the elevation of the tracking station.

Generally the overseas data went up to only 30,000 ft while the Edwards AFB data usually went above 80,000 ft.

Table 3. Weather measurements, refractivities, zenith range effect, and inverse scale heights as a function of altitude for radiosonde balloon observations at Edwards AFB on May 24, 1967

Index	Altitude, ft	Pressure, mb	Temperature, °C	Vapor pressure, mb	Relative humidity, %	Refractivity			Zenith range effect from surface, m			Inverse exponential scale height, km^{-1}		
						Total	Dry	Wet	Total	Dry	Wet	Total	Dry	Wet
0	2375	927.2	24.0	7.5	25	273.8	242.1	31.7	—	—	—	—	—	—
1	3000	907.2	24.2	7.6	25	268.8	236.8	32.1	0.052	0.046	0.006	0.097	0.118	−0.062
2	4000	876.1	23.5	7.4	25	260.5	229.2	31.4	0.132	0.117	0.016	0.103	0.107	0.072
3	5000	845.9	21.6	6.8	26	251.9	222.7	29.2	0.210	0.185	0.025	0.111	0.094	0.235
4	6000	816.6	19.6	6.2	27	243.4	216.5	27.0	0.286	0.252	0.034	0.112	0.093	0.258
5	7000	788.1	17.5	5.7	28	235.6	210.4	25.2	0.359	0.317	0.041	0.108	0.093	0.229
6	8000	760.3	14.9	5.8	34	230.9	204.8	26.1	0.430	0.381	0.049	0.066	0.088	−0.116
7	9000	733.3	12.2	5.7	40	225.5	199.4	26.1	0.500	0.442	0.057	0.077	0.088	−0.005
8	10000	707.0	9.6	5.4	45	219.2	194.0	25.2	0.567	0.502	0.065	0.093	0.090	0.117
10	12000	656.7	5.2	3.6	41	200.4	183.1	17.3	0.695	0.617	0.078	0.135	0.095	0.529
12	14000	609.0	−0.7	3.2	56	189.5	173.5	16.1	0.814	0.726	0.088	0.093	0.089	0.129
14	16000	564.0	−5.1	2.9	71	178.3	163.3	15.1	0.926	0.828	0.097	0.107	0.101	0.168
16	18000	521.6	−9.7	2.5	85	167.1	153.6	13.4	1.031	0.925	0.106	0.108	0.100	0.195
18	20000	481.8	−14.4	1.7	85	154.0	144.5	9.5	1.129	1.016	0.113	0.148	0.102	0.788
23	25000	392.9	−23.4	0.7	73	126.3	122.1	4.2	1.340	1.218	0.121	0.123	0.114	0.394
28	30000	317.4	−36.7	0.2	61	105.5	104.2	1.3	1.516	1.390	0.125	0.103	0.105	−0.077
33	35000	253.2	−49.3	0	43	87.8	87.8	0	1.663	1.537	0.126	0.133	0.133	—
38	40000	199.6	−60.6	0	12	72.9	72.9	0	1.785	1.659	0.126	0.124	0.124	—
43	45000	156.1	−61.4	0	9	57.2	57.2	0	1.884	1.758	0.126	0.179	0.179	—
48	50000	122.4	−61.8	0	9	44.9	44.9	0	1.962	1.836	0.126	0.165	0.165	—
58	60000	74.7	−62.4	0	9	27.5	27.5	0	2.070	1.944	0.126	0.166	0.166	—
68	70000	45.9	−57.3	0	9	16.5	16.5	0	2.136	2.010	0.126	0.171	0.171	—
78	80000	28.6	−52.2	0	9	10.0	10.0	0	2.175	2.050	0.126	0.158	0.158	—
88	90000	18.0	−49.2	0	9	6.2	6.2	0	2.200	2.074	0.126	0.156	0.156	—
98	100000	11.4	−44.3	0	9	3.9	3.9	0	2.215	2.089	0.126	0.144	0.144	—
108	110000	7.3	−39.0	0	9	2.4	2.4	0	2.224	2.089	0.126	0.136	0.136	—
118	120000	4.7	−35.9	0	9	1.5	1.5	0	2.230	2.104	0.126	0.141	0.141	—

Table 4. Radiosonde balloon site parameters

Radiosonde station	Latitude	Longitude	Elevation, m	Nearest DSS	DSS elevation, m	Distance from DSS, km
Edwards AFB	34°55′ N	117°54′ W	724	Goldstone[a]	1032	100
Madrid	40°28′ N	3°34′ W	606	Madrid[b]	789	70
Wagga	35°10′ S	147°28′ E	214	Canberra	656	140
Woomera	31°9′ S	136°48′ E	165	Woomera	151	
Pretoria	25°44′ S	28°11′ E	1330	Johannesburg	1398	50

[a]DSS 14.
[b]DSS 61.

6. Computed Zenith Range Effect

After each month's weather data has been used to calculate the total, dry, and wet zenith range effects from the surface to various altitudes, $\Delta\rho_z$ is plotted as a function of time of the month and the monthly averages and standard deviations are calculated. For example, Fig. 1 shows the daily variations of the zenith range effect produced by the portion of the troposphere between the surface and 80,000 ft over Edwards AFB during March 1967, with the average, the average plus the standard deviation, and the average minus the standard deviation also included. The corresponding values of the total and dry zenith range effects for the total troposphere may be obtained by adding 6 cm to the surface to 80,000-ft level results.

To better illustrate the daily variations in the zenith range effect the difference of the zenith range effects from their monthly averages are also plotted. An example of these deviations is given in Fig. 2, where the average, and the average plus and minus the standard deviation are plotted again.

Similar plots for the overseas stations which would give representative values of the total and dry zenith range effects may not be obtained from the radiosonde balloon data alone because the data usually stops at about 30,000 ft. However, representative values of the complete dry zenith range effect may be found by using the high correlation between this quantity and the surface pressure. This correlation will be illustrated in the next section.

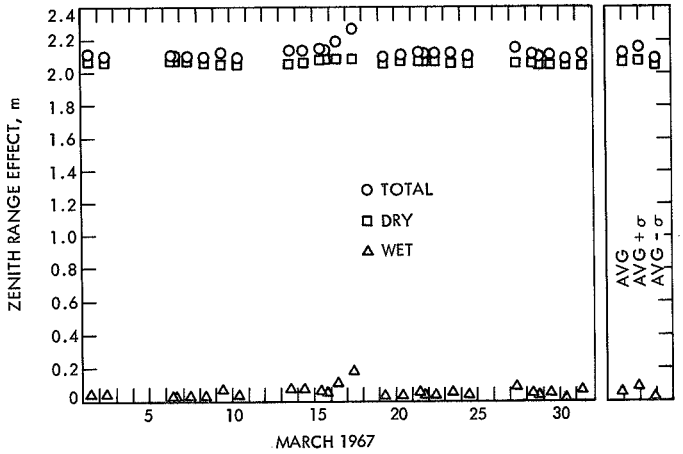


Fig. 1. Total, dry, and wet zenith range effects for portion of troposphere between surface and 80,000 ft over Edwards AFB during March 1967

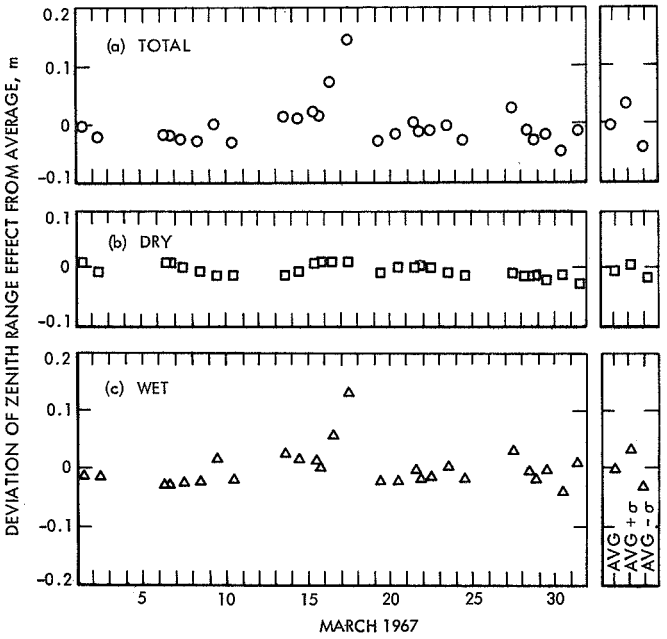


Fig. 2. Deviations from monthly averages of total, dry, and wet zenith range effects over Edwards AFB during March 1967

Figures 3 through 7 summarize the preceding monthly results by plotting for each month the monthly average, the average plus and minus the standard deviation, and the maximum and minimum values of the zenith range effects. These graphs illustrate the important feature that the total zenith range effect seldom varies by more than $\pm 5\%$ from the yearly average.

Figures 3 through 7 also include the zenith range effect obtained by using the profile which currently is used to make tropospheric corrections in the orbit determination program. This profile is exponential in character with a sea level value of 340 and scale height of 7.042 km. The total zenith range effect computed by the following equation agrees to within approximately $\pm 4\%$ of the yearly average computed from the radiosonde balloon data:

$$\Delta\rho_z\,(\text{const exp}) = 340 \times 10^{-6} \int_{x_s}^{\infty} \exp\left[-\frac{x}{7.042}\right] dx$$

where

x_s = height of station above sea level, km

x = height above sea level, km

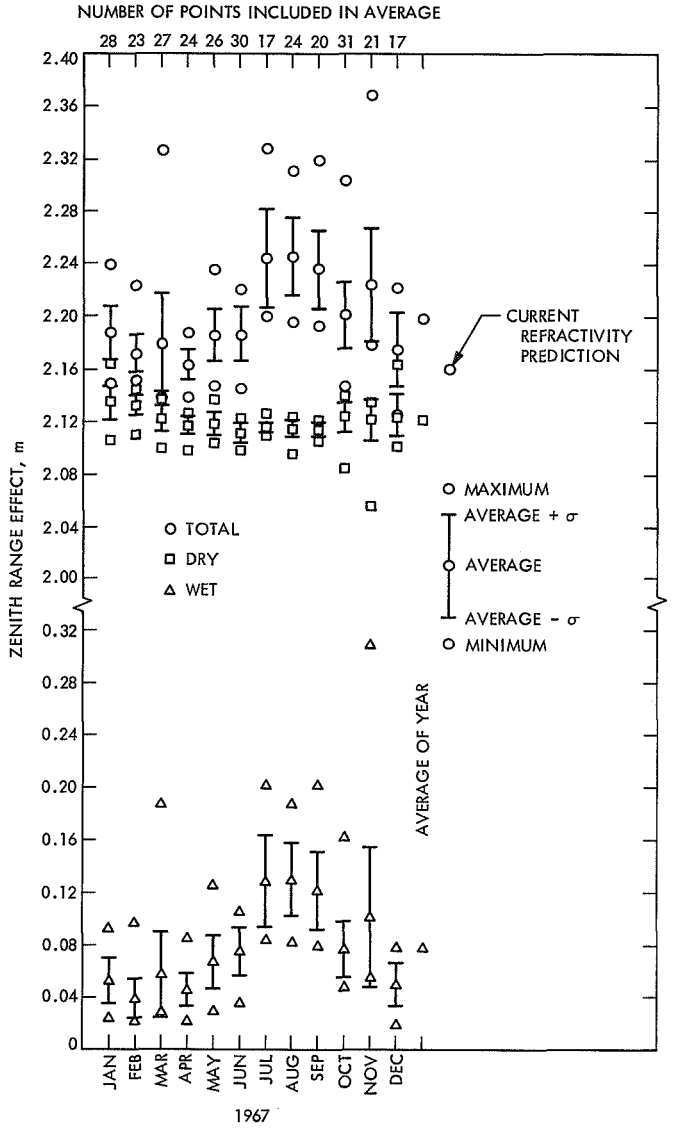


Fig. 3. Monthly averages and standard deviations of total, dry, and wet zenith range effects over Edwards AFB during 1967

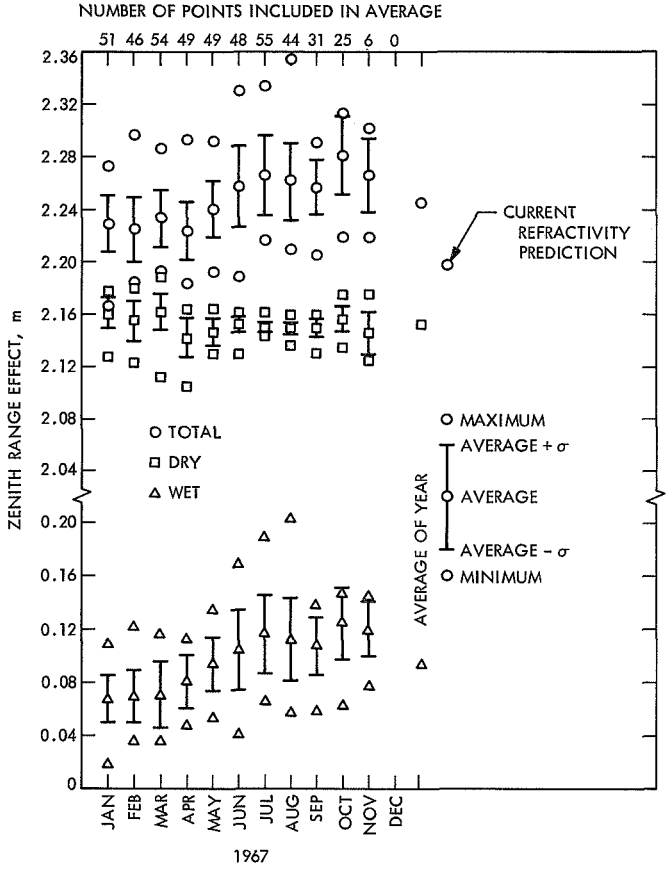


Fig. 4. Monthly averages and standard deviations of total, dry, and wet zenith range effects over Madrid during 1967

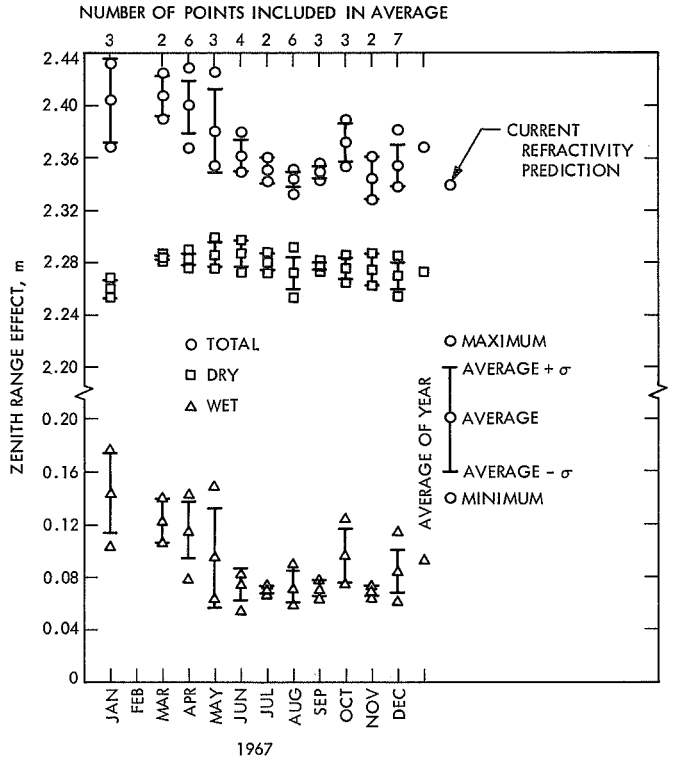


Fig. 5. Monthly averages and standard deviations of total, dry, and wet zenith range effects over Woomera during 1967

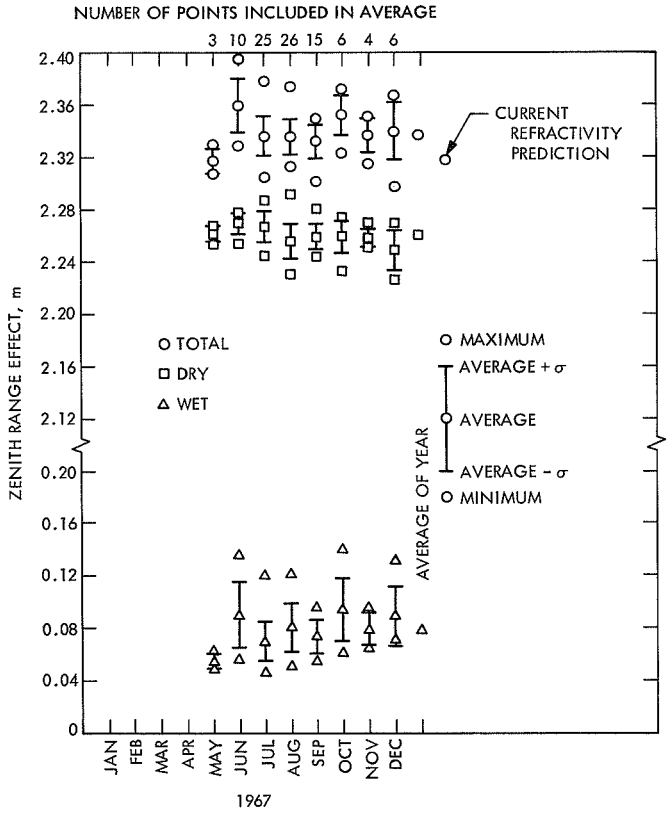


Fig. 6. Monthly averages and standard deviations of total, dry, and wet zenith range effects over Wagga during 1967

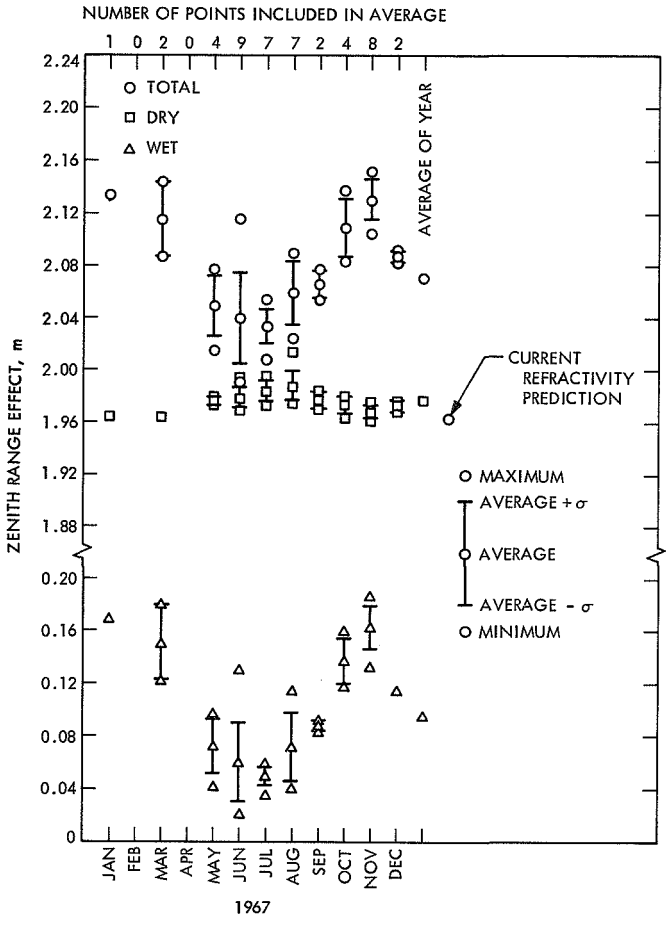


Fig. 7. Monthly averages and standard deviations of total, dry, and wet zenith range effects over Pretoria during 1967

7. Prediction of Dry Zenith Range Effect From the Surface Pressure

From Eqs. (4) and (5), the dry zenith range effect may be written as

$$\Delta\rho_z\,(\text{dry}) = 77.6\int_{\text{surface}}^{\infty}\frac{P(h)}{T(h)}\,dh \tag{11}$$

Berman (see *Chapter XIII-D*, this volume) has shown that by using the hydrostatic equation

$$dP = -\bar{\rho}gdh \tag{12}$$

where $\bar{\rho}$ is the density and g is the gravitational acceleration (assumed constant with altitude), and the perfect gas law

$$P = \bar{\rho}\bar{R}T \tag{13}$$

where $\bar{R}$ is the gas constant, Eq. (11) reduces to

$$\Delta\rho_z^*(\text{dry}) = 77.6\frac{\bar{R}}{g}P_s \tag{14}$$

where P_s is the surface pressure and the * indicates that $\Delta\rho_z$ (dry) has been calculated under certain conditions. Equation (14) predicts a linear correlation between the surface pressure and $\Delta\rho_z$ (dry). To see if this predicted correlation is present, scatter diagrams of $\Delta\rho_z$ versus P_s were plotted and the correlation between them calculated. The results of this process for the Edwards AFB, as illustrated in Fig. 8, always show a high (i.e., > 0.97) correlation between the dry zenith range effect and surface pressure.

The overseas data cannot be used to verify the correlation according to Eq. (14) since the data generally stops below 30,000 ft. However, this correlation may be shown to be present if the upper limit of Eq. (11) is replaced by $h_{\max}$ so that

$$\Delta\rho_z^*\ (\text{dry}, h_{\max}) = 77.6\frac{\bar{R}}{g}[P_s - P(h_{\max})]$$

where $\Delta\rho_z$ (dry, $h_{\max}$) is the zenith range effect produced by the troposphere between the surface and $h_{\max}$.

It is not surprising that correlation between the dry zenith range effect and surface pressure exists because the hydrostatic equation and gas law used to obtain Eq. (14) are very close to the equations used to determine the height of a particular set of weather measurements.

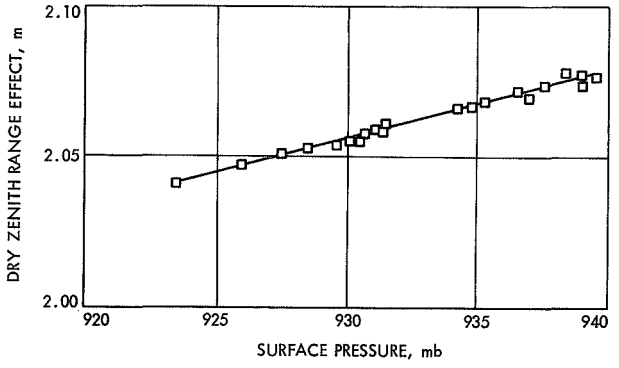


Fig. 8. Scatter diagram showing correlation between dry zenith range effect and surface pressure for data taken over Edwards AFB during March 1967

8. Cursory Examination of Methods for Predicting the Wet Zenith Range Effect From Surface Measurements

From Eqs. (5) and (6) the wet zenith range effect is calculated by evaluating the following integral:

$$\Delta\rho_z\,(\text{wet}) = 2280\int_0^\infty \frac{1}{T^2}\exp_{10}\left[\frac{7.4475(T-273.2)}{T-38.5}\right]R\,dh \tag{15}$$

Obviously, to perform this integral based upon surface measurement, it is necessary to specify some models for $T(h)$ and $R(h)$ based upon these quantities. The difficulty in specifying models that are representative of what the physical situation may actually be lies not with $T(h)$, but almost entirely with $R(h)$. For example, the difference between the representative wet refractivity profiles shown in Fig. 9 is due for the most part to differences in the relative humidity profile. Note that although the March 9 wet surface refractivity is only 70% of the March 10 value, the March 9 $\Delta\rho_z$ (wet) is twice as large as the March 10 value.

Berman (*Chapter XIII-D*, this volume) has made an attempt at determining an equation for $\Delta\rho_z$ (wet) in terms of surface measurements by assuming that the temperature falls off according to a constant lapse rate and the relative humidity is constant at its surface values. Berman's equation for the wet zenith range effect, $\Delta\rho_z(B)$, which results from his assumptions, is given as

$$\Delta\rho_z(B) = \left[\frac{C_1C_2R_s}{\gamma(T_s)}\right]\left[\frac{(1-T_sC)^2}{B-AC}\right]\exp\left[\frac{AT_s-B}{T_s-C}\right] \tag{16}$$

where

$C_1 = 77.6$

$C_2 = 29341.0$

T_s = surface temperature

R_s = surface relative humidity

$A = 7.4475\ \ln(10)$

$B = 2034.28\ \ln(10)$

$C = 38.45$

$$\gamma(T_s) = \frac{T_s - 216.65}{11 - h_0}$$

h_0 = station elevation, km

To see how well Berman's wet model predicts the wet zenith range effects, monthly plots were made with $\Delta\rho_z(B)$ computed from Eq. (16), $\Delta\rho_z$ (wet) computed from the radiosonde balloon measurements, and the difference between them. An example of these plots is shown in Fig. 10.

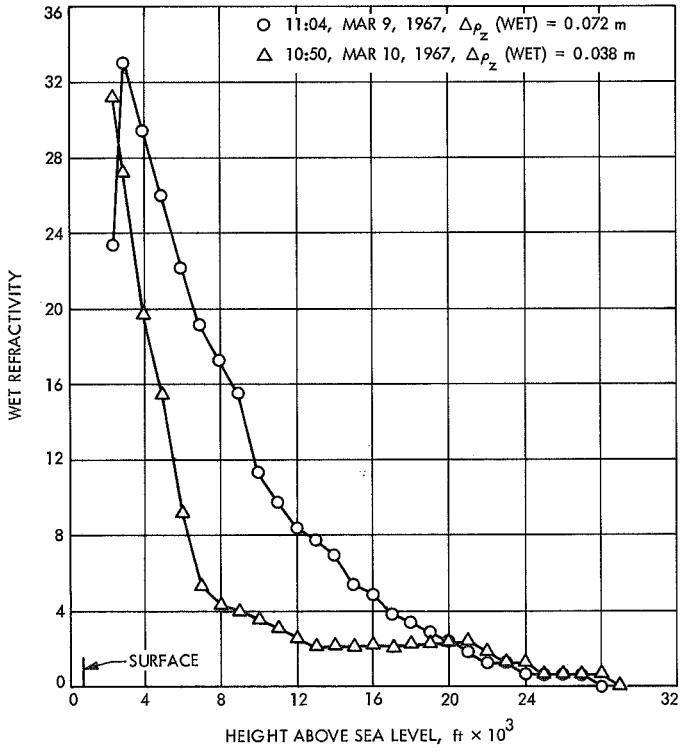


Fig. 9. Wet refractivity profiles over Edwards AFB

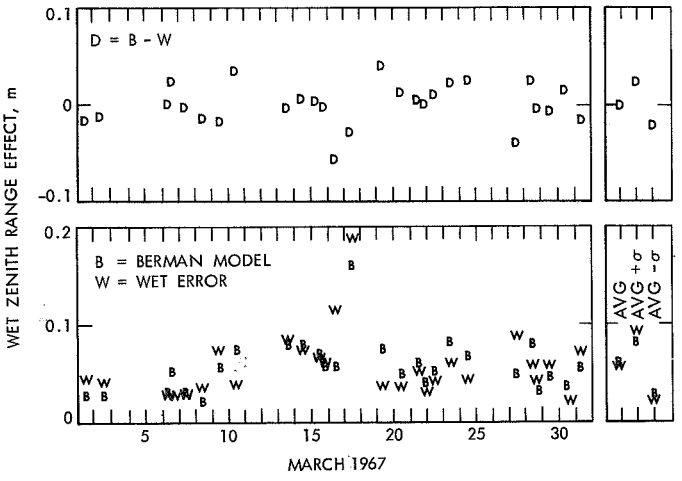


Fig. 10. Actual, predicted, and differenced wet zenith range effect for Edwards AFB during March 1967

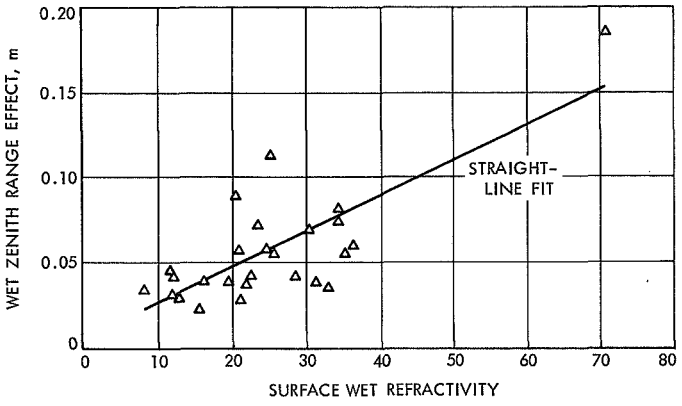


Fig. 11. Surface wet refraction vs wet zenith range effect with correlation 0.76 for Edwards AFB during March 1967

To investigate the correlation between the wet zenith range effect and the wet surface refractivity, the monthly $\Delta\rho_z$ (wet) versus N_w (surface) scatter diagram was plotted, the correlation calculated, and a least-squares straight-line fit made. An example of this output is shown in Fig. 11. Unfortunately, the slope and intercept of this fitted straight line vary widely throughout the year.

A summary of the results of trying to predict $\Delta\rho_z$ (wet) from the Edwards AFB data is given in Fig. 12. This figure is representative of the overseas data and indicates that nothing substantial is gained by using surface measurements and Berman's wet model in its present form. However, if the monthly variations of the average wet zenith range effect are repetitive from year to year, it may be possible to reduce the standard deviations from the yearly average of approximately 0.06 m by a factor of two.

9. Apparent Changes in Station Locations Produced by the Troposphere

For the elevation angles larger than approximately 5 deg, a good (10%) approximation to the range effect at a particular elevation angle, γ, produced by a spherically symmetric troposphere is given by

$$\Delta\rho(\gamma) = \frac{\Delta\rho_z}{\sin\gamma} \tag{17}$$

If the tracking station is viewing a zero-declination spacecraft, the time behavior of the tropospheric range rate correction associated with the above equation will be

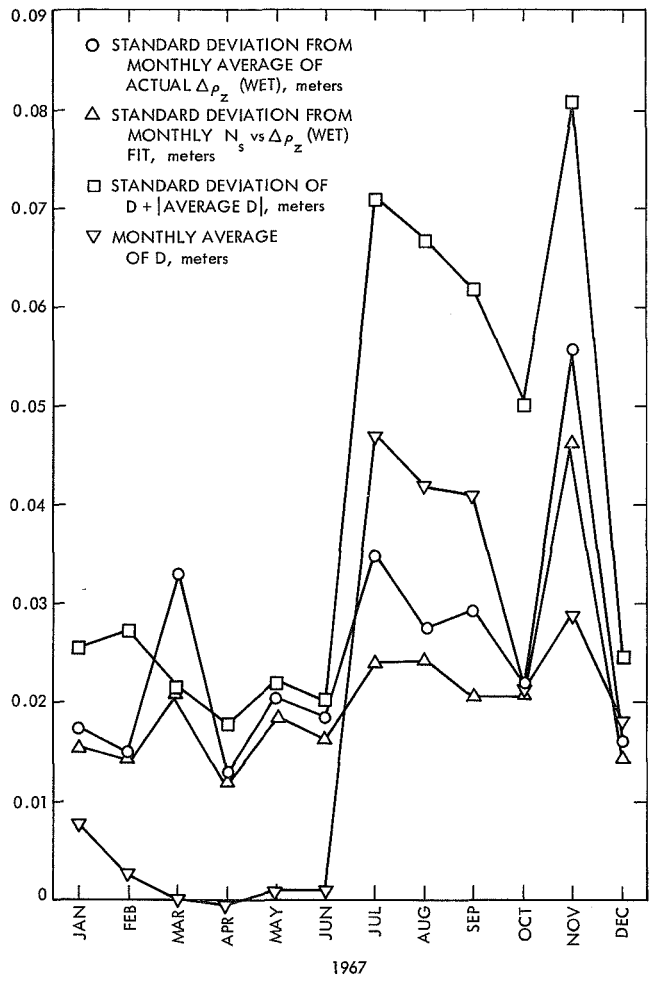


Fig. 12. Monthly summary of wet zenith range effect predicted from surface weather measurements taken at Edwards AFB during 1967

given by

$$\Delta\rho(t) = \frac{\omega\,\Delta\rho_z}{\cos\Phi}\,\frac{\sin\omega t}{\cos^2\omega t}$$

where

t = time past meridian crossing

ω = angular rotation rate of the earth

Φ = latitude of tracking station

In the previous sections it has been shown that the standard deviation of $\Delta\rho_z$ from a constant model is approximately 0.07 m and if surface pressure measurements are used in conjunction with the dry Berman model this standard is reduced to approximately 0.06 m. An examination of the plots showing $\Delta\rho_z$ as a function of time has

given an estimate that the average of the absolute value of the rate of change of $\Delta\rho_z$ is approximately 0.02 m/day.

A useful artifice for investigating navigational errors, such as the ones produced by the troposphere, is to describe them in terms of equivalent errors in tracking station locations. As described in SPS 37-57, Vol. II, pp. 24–29, an effect which corrupts tracking data can be decomposed into parameters, one of which is the apparent change in the station distance off the spin axis, Δr_s, and another is the apparent change in the station's longitude, $\Delta\lambda$. The results of performing such an analysis for the various unmodeled tropospheric effects discussed above are given in Table 5. The numbers in the table were computed under the assumption that a tracking station situated at a latitude of 0° was viewing a zero-declination spacecraft and taking range rate data throughout a pass which was symmetric about the meridian crossing and had a minimum elevation angle, γ_{min}.

Table 5. Estimated standard deviations of apparent daily change in station locations[a]

Minimum elevation angle, deg	Δr_s, m	$\Delta\lambda$, m
15	0.26	0.12
10	0.42	0.19
5	0.85	0.42

[a]For zero-degree latitude station viewing a zero-declination spacecraft under assumption that $\Delta\rho_z(t) = 0.06$ m + (0.02 m/day)t days.

10. Summary and Conclusion

An examination of radiosonde balloon data taken during 1967 has shown that the tropospheric zenith range effect is quite a stable quantity and will almost always stay within ±5% of its yearly average. The variation produced by the dry constituents can be eliminated almost entirely if use is made of surface pressure measurements. The remaining 80 or 90% of the variations may possibly be reduced up to a factor of two if the seasonal variations of the wet zenith range effect is repeatable over the years. The effects of the zenith range effect upon radio tracking data may be grossly represented by errors in tracking station locations of approximately $\Delta r_s = 0.3$ m and $\Delta\lambda = 0.1$ m if a minimum elevation of 15 deg is used.

These estimates of the navigational errors produced by the variation of the zenith range effect should not be taken to represent the total effect of the troposphere. Unfortunately, there are other possible tropospheric error sources which may contribute significantly to the errors produced by the troposphere. Two such error sources arise from mapping zenith values to arbitrary elevation angles, and from effects produced by the nonhomogeneous structure of the troposphere. These other tropospheric error sources are currently under investigation.

Reference

1. Bean, B. R., and Dutton, E. J., *Radio Meteorology*, National Bureau of Standards, Monograph 92, p. 11, 1966.

C. The Effect of the Diurnal Variation of the Earth's Ionosphere on Interplanetary Navigation, *B. D. Mulhall and K. L. Thuleen*[1]

1. Introduction

The need to calibrate the effect of charged particles in the earth's ionosphere on radio tracking data has been recognized and was undertaken on a research basis for the *Mariner* Mars 1969 encounter.[2] To better understand the effect of the ionosphere on navigation and to predict what the effect might be for future missions, a parametric study was performed. The diurnal variation of the ionosphere was modeled. The seasonal and solar cycle changes in the model parameters were determined by fitting the model to actual ionospheric measurements sampled from different seasons

[1]The authors wish to acknowledge the assistance of S. Finley in developing programs used in this study.

[2]Mulhall, B. D., et al., *Tracking System Analytic Calibration Activities for the Mariner Mars 1969 Mission*, Technical Report 32-1499. Jet Propulsion Laboratory, Pasadena, Calif. (in press).

and years. To represent different mission geometries, the ionospheric effect was computed for various sun–earth–probe angles.

2. Ionospheric Model

The diurnal variation of the charged particle content in the earth's ionosphere can be represented by

$$E_c = \begin{cases} D_c + A_c \cos \omega (t - \Phi) \\ \text{whichever is greater} \\ D_c \end{cases}$$

where

E_c = modeled columnar electron content

ω = rotational rate of the earth

t = meridian time

and D_c, A_c, and Φ are the model parameters.

The model parameters represent the nighttime average electron content, D_c, the peak concentration during the day, $D_c + A_c$, and the time of the peak concentration, Φ. This model has been used for some time in the reduction of Faraday rotation measurements made at Goldstone, and was intended as a scheme for predicting the diurnal variation. However, its prediction capability is limited (Mulhall, SPS 37-64, Vol. II, pp. 25–27).

The model was fitted to measurements of the ionosphere made by Stanford Radio Science Laboratory[3] and by JPL at the Venus site (DSS 13). Figure 1 shows the seasonal variations in the values of D_c and A_c. Table 1 lists the sample mean and standard deviation of each parameter for spring, summer, fall, and winter. From the figure and the table, it can be seen that the nighttime concentrations of charged particles are generally higher in summer than in winter, probably

[3]The authors wish to thank M. Davis and T. Howard of the Center for Radar Astronomy, Stanford Electronics Laboratories, for providing the ionospheric data.

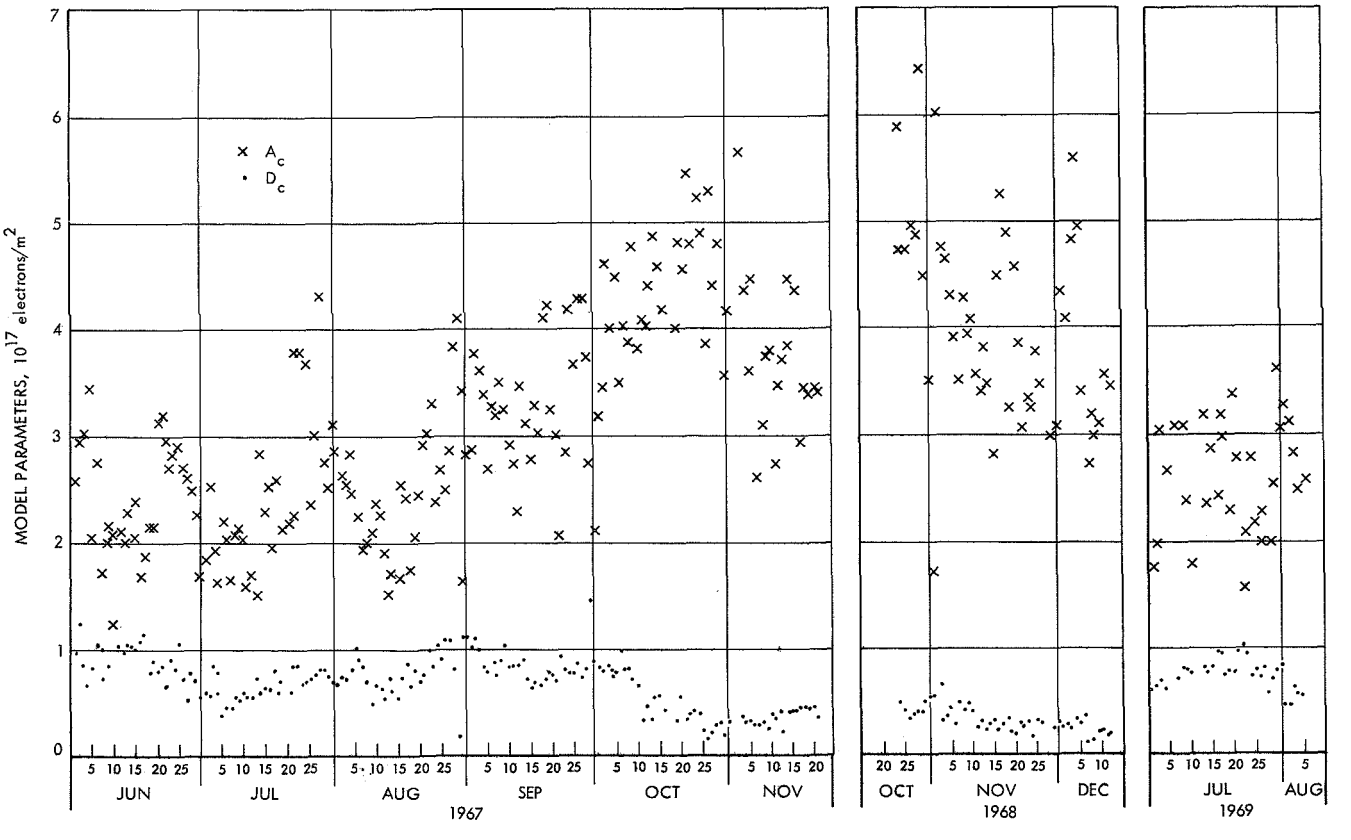


Fig. 1. Ionospheric conditions

Table 1. Seasonal variation of ionospheric model parameters (northern hemisphere)

Parameter	Summer			Spring/Fall		Winter
	1965	1967	1969	1967	1968	1965
A_c						
Mean	1.32	2.44	2.65	3.77	4.06	1.49
Standard deviation	0.09	0.62	0.57	0.79	0.95	0.15
D_c						
Mean	0.39	0.79	0.78	0.60	0.34	0.29
Standard deviation	0.14	0.23	0.15	0.27	0.11	0.15
Φ, hours						
Mean	13.17	13.90	13.03	13.79	12.61	12.93
Standard deviation	2.38	1.82	2.18	1.52	1.48	1.51
Number of days in sample	27	90	32	78	48	32

due to the warmer nighttime temperatures. However, daytime peak concentrations are higher in fall, winter, and spring than in summer. There are several theories proposed to explain this seemingly contradictory situation; however, none has been completely satisfactory. The uncertainty in Φ (see Table 1) increases in summer and, as discussed in SPS 37-64, Vol. II, the summer ionosphere stays at a maximum concentration for a fairly long period of time.

The 11-yr solar cycle reached a maximum in 1968. By comparing the average values of nighttime and daytime peak concentrations (D_c and A_c) listed in Table 1 and Fig. 1 for 1965 versus the values for 1967, 1968, and 1969, it can be seen that electron concentrations tend to increase with the solar cycle.

3. Navigational Effect

To determine the ionospheric effect on navigation for various missions, the sun–earth–probe angle (χ) was varied as the parameters of the model were varied. By using the three-parameter navigation filter developed by Hamilton and Melbourne (reference footnote 2; SPS 37-39, Vol. III, pp. 18–23; and SPS 37-57, Vol. II, pp. 24–29), the ionospheric effect was resolved into apparent changes in the tracking station longitude, the station distance from the earth's spin axis, and changes in apparent spacecraft range rate:

$$\dot{\rho} = a + b \sin \omega t + c \cos \omega t$$

where

$$a = \dot{r}$$

$$b = r_s \omega \cos \delta$$

$$c = r_s \omega \Delta\theta \cos \delta$$

and

$\dot{\rho}$ = spacecraft station range rate

$\dot{r}$ = spacecraft geocentric range rate

r_s = tracking station spin radius

δ = spacecraft declination

and $\Delta\theta$ can be considered the error in tracking station longitude.

Using the program ION (reference footnote 2), doppler effects of the ionosphere were computed for various sun–earth–probe angles. The doppler effects were fitted to the three-parameter navigation filter by least-squares estimation. The empirical sensitivities of a, b, and c with respect to A_c, D_c, and Φ were computed. Figures 2 and 3 show the partial derivatives of a, b, and c with respect to A_c and D_c. Figures 4, 5, and 6 show the partial derivatives of a, b, and c with respect to Φ for various values of Φ since these partials are nonlinear.

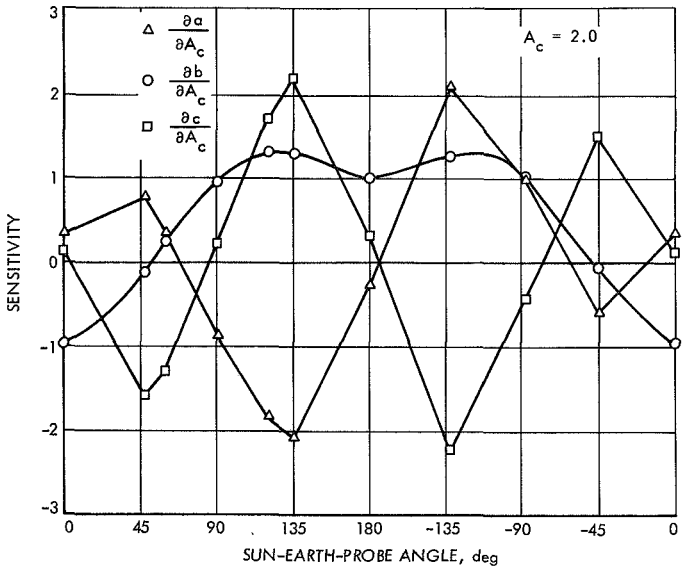


Fig. 2. Sensitivity of a, b, c with respect to A_c versus sun–earth–probe angle

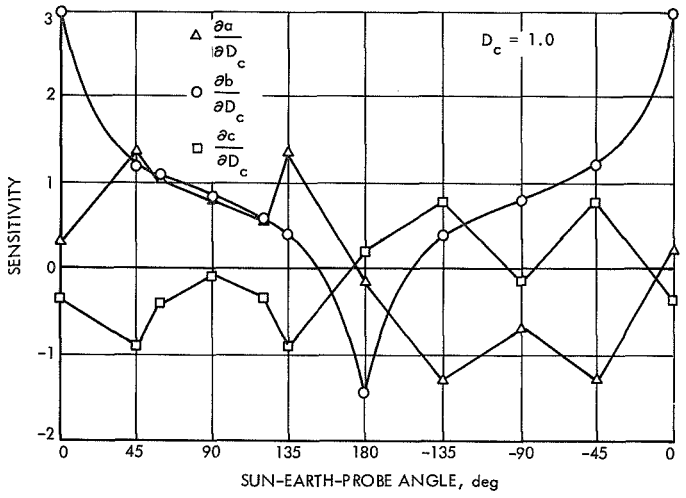


Fig. 3. Sensitivity of *a*, *b*, *c* with respect to D_c versus sun–earth–probe angle

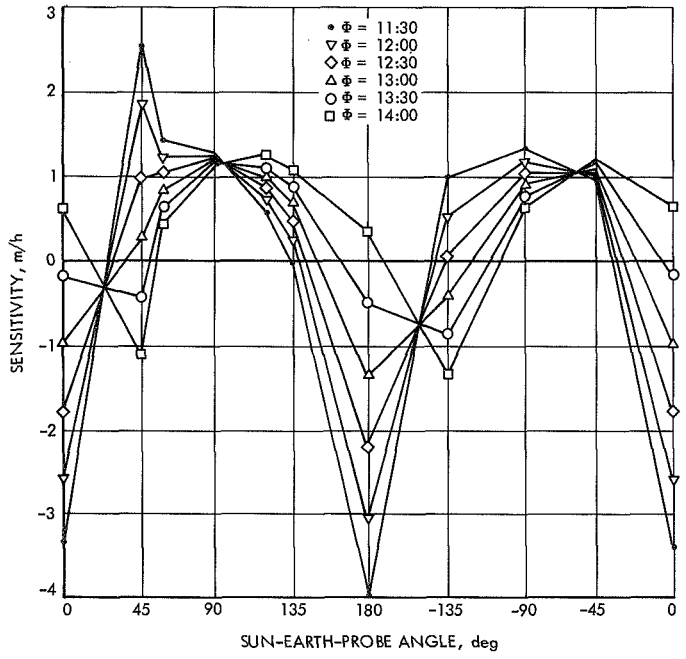


Fig. 4. Sensitivity of *a* with respect to Φ versus sun–earth–probe angle

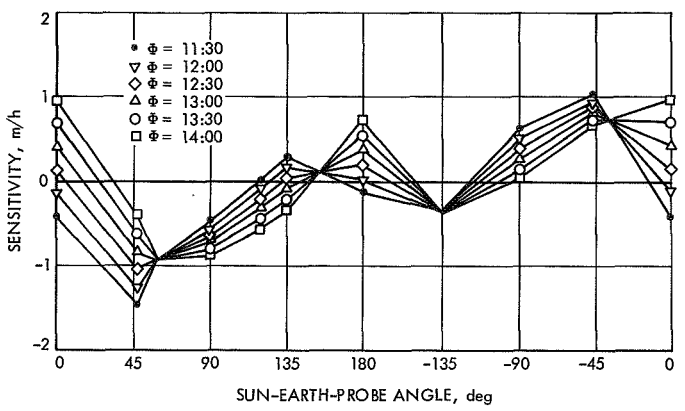


Fig. 5. Sensitivity of *b* with respect to Φ versus sun–earth–probe angle

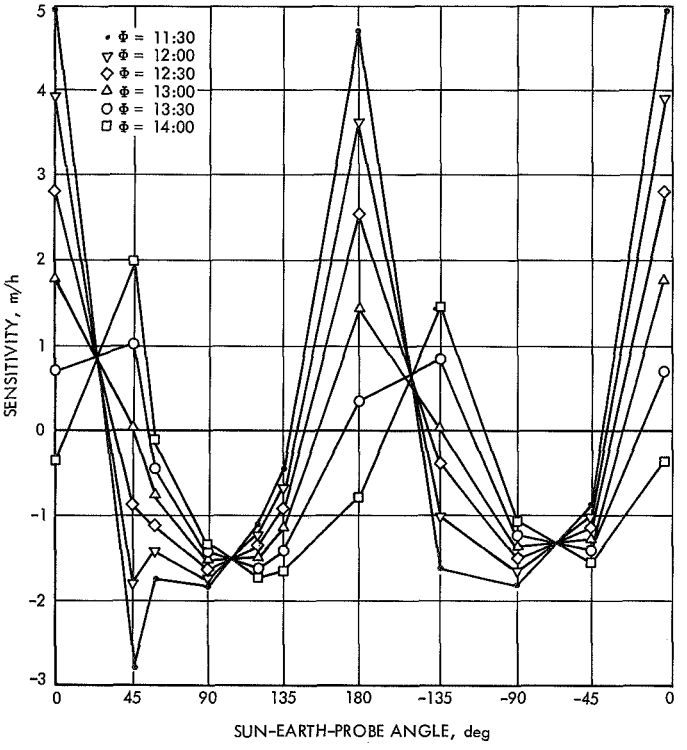


Fig. 6. Sensitivity of *c* with respect to Φ versus sun–earth–probe angle

The high negative correlation between *a* and *c* is apparent in these figures. A radical change in *a* is accompanied by an opposite and nearly equal change in *c*. An interesting comparison, which will be performed later, will be to compute the partials for the eigenvalues of *a* and *c*. Presumably, these will resemble the more smoothly varying partials of *b*, since *b* is independent of *a* and *c*.

4. Estimation of Ionospheric Effect on Navigation

The intended use of the partial derivative is to estimate the sensitivity of orbit determination to ionospheric effects. The sun–earth–probe angle (χ) at critical mission phases characterized the ionospheric effect since the relationship of the spacecraft rise and set times to the diurnal variation of the ionosphere is typified by χ. For

example, the effect of the ionosphere on b (and, consequently, on r_s or spacecraft declination) with respect to the amplitude of the peak electron concentration (A_c) is at a maximum at $\chi = 120$ deg. *Mariner VI* encounter occurred in late July 1969 at about a sun–earth–probe angle of 120 deg. In July 1967, χ for *Mariner V* was approximately 15 deg. The average ionospheric effect on r_s for the month of July for *Mariner VI* was 3.3 m while for *Mariner V* the effect was 1.1 m. Even though the ionospheric concentrations of electrons were similar during July 1967 and July 1969, the increased sensitivity of b due to χ caused an increased navigational error.

V. Communications Systems Research

ADVANCED ENGINEERING

A. Digital Acquisition and Detection: Digital Frequency Doubler, *G. Morris*

1. Introduction

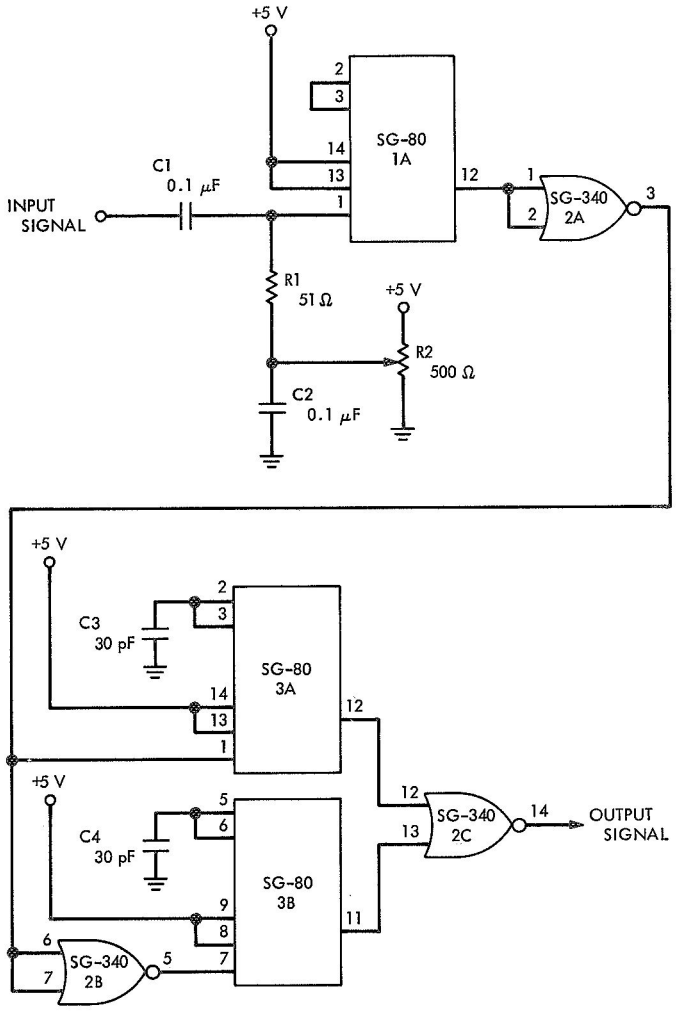
In the design of digital systems, it is often necessary to have several different clock frequencies. It is quite easy to design digital frequency dividers with flip-flops. Here is given a presentation of a digital frequency doubler using only three integrated circuits and six discrete components (Fig. 1). This circuit has been evaluated in some experimental timing equipment in the DSIF frequency and timing subsystem at DSS 11. In this application, the doubler multiplied the available 5 MHz to 10 MHz to drive a counter. Previously, this doubling would have been done in an RF doubler which requires tuned circuits and would only operate at one frequency. The advantage of the digital doubler is that it can handle any frequency up to the maximum of the integrated circuits used, in this case, 20 MHz.

2. Theory of Operation

The active circuits are all Sylvania SUHL line integrated circuits. The SG-80 is a dual pulse shaper/delay *and* gate and the SG-340 is a quad two-input *nor* gate. The input signal is assumed to be a sine wave from a source which can deliver approximately 1 V rms into a 50-Ω load. This source is the DSIF standard. The SG-80 contains a regenerative Schmitt trigger which causes the output to snap to a logic "1" when the input signal rises to approximately 1.4 V. The output will snap to logic "0" when the input falls to approximately 0.8 V. This hysteresis gives a positive action which prevents multiple pulses when the input signal contains noise. The input coupling circuit of C1, C2, and R1 provides a 50-Ω load for the signal source. The potentiometer R2 is adjusted to obtain a square wave output from pin 12 of 1A.

The *nor* gate 2A is not absolutely necessary but is desirable so that 3A and 3B are both driven by identical circuits to obtain better symmetry and, hence, stability. The output from the Schmitt trigger is inverted by *nor* gate 2A for input to 3A. The output of 2A is inverted by 2B for input to 3A. Therefore, complimentary square wave signals are input to delay gates 3A and 3B. The delay circuits in 3A and 3B are controlled by the external capacitors C3 and C4. The delay is approximately 1.5 ns/pF. The 30-pF capacitors shown result in a delay of approximately 50 ns.

The frequency doubling is obtained by the delays in 3A and 3B. When 3A turns off, there is a delay determined by C4 before 3B turns on. During this interval, the output of *nor* gate 2C is high. When 3B turns off, there is a delay determined by C3 before 3A turns on. During this interval, the output of *nor* gate 2C is also high. Therefore, an output pulse is generated at both zero crossings of the input signal. The pulse widths of the output signal are



SG-80 DUAL PULSE SHAPER/DELAY AND GATE

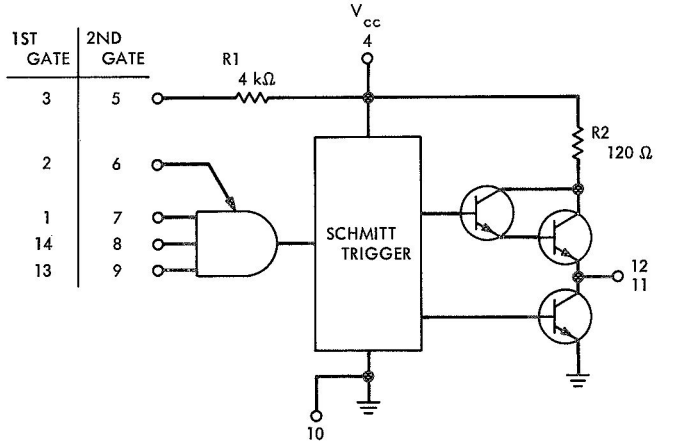


Fig. 1. Digital doubler schematic

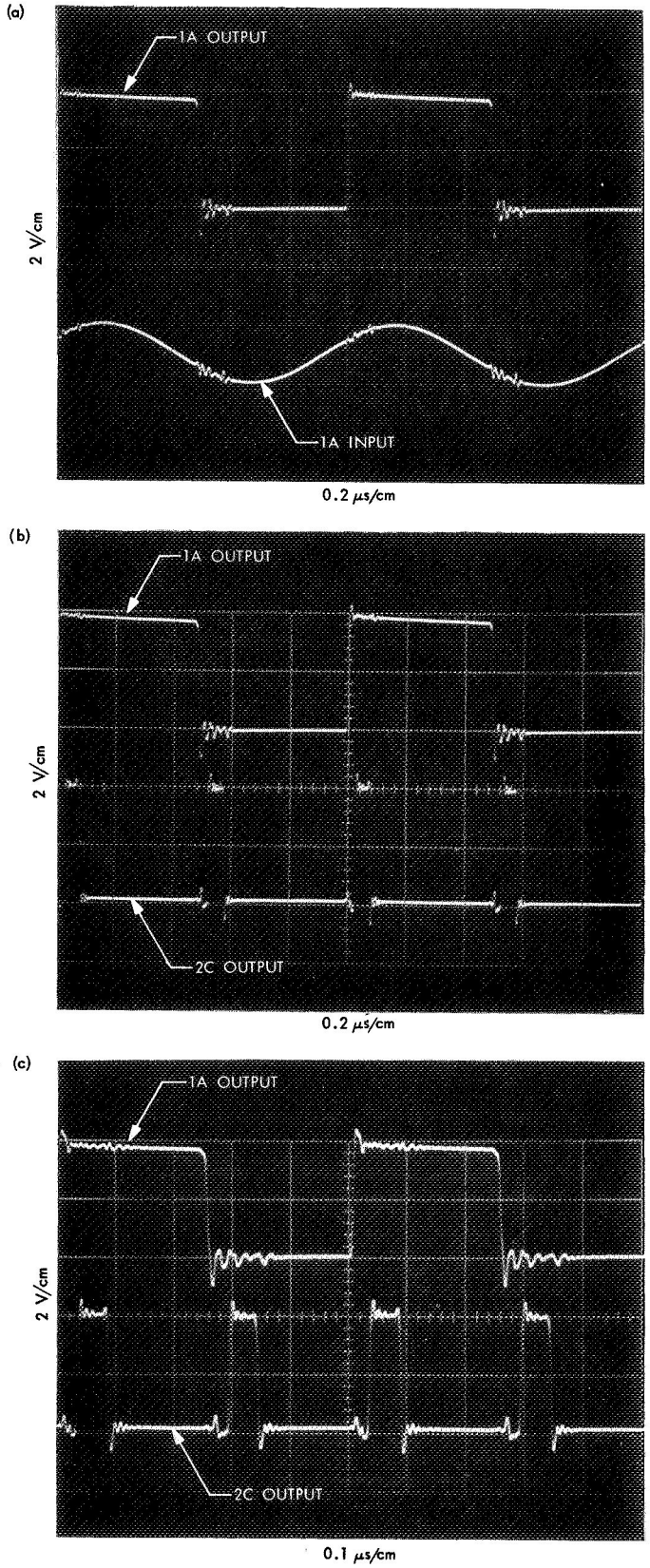


Fig. 2. Digital doubler oscilloscope waveforms

determined by capacitors C3 and C4 and are independent of the input frequency.

3. Experimental Results

Oscilloscope waveforms are shown in Fig. 2. The output of 1A is shown on the upper trace and the input to 1A is shown on the lower trace of Fig. 2a. The input signal is a 1-MHz sine wave. The output of 2C is shown on the lower trace of Fig. 2b. To demonstrate the fact that output signal pulse width is independent of frequency, the input signal is changed to 2 MHz in Fig. 2c. Note the sweep time is halved but the output pulse width remains constant.

B. Digital Telemetry and Command: A Collection of Results on Computational Complexity,

J. E. Savage[1]

1. Introduction

This article presents a variety of results obtained in our continuing effort to understand computational procedures and the effect of machine parameters on computation. The focus of this article is on a measure of "computational work" introduced earlier in the study of decoders for error correction.[2] We show that many different machines exist which compute simple and very complex functions with a computational work near the minimum required for these functions. We do not know whether the minimum can be approached when functions are neither very simple nor very complex. We consider computational procedures where the computation time depends on input data and the bound on the maximum time required for any data point. Bounds are also derived on the number of storage cells required to compute a function with an autonomous machine. In addition, we comment on the work potential of a computer which has many types of storage units attached to it and we offer a few remarks on the sorting problem. Finally, a simple quantum-mechanical bound is derived on the maximal complexity of functions which can be computed in t seconds with E ergs of energy.

All results are simply derived and follow from extensions of an inequality.[2] This inequality states that if the sequential machines $S_1, \cdots, S_k$ are interconnected to compute a function $f(\ )$, if they have the same cycle length and are constructed with logic elements from Ω and individually accessed memory cells, and if S_i has X_i logic elements and executes T_i cycles, then

$$\chi = \sum_{i=1}^{k} X_i T_i \geqq C_\Omega(f) \tag{1}$$

where $C_\Omega(f)$ is the minimum number of logic elements from Ω required to compute $f(\ )$ with a combinational machine. The quantity χ *is called computational work* and $C_\Omega(f)$ *is called the complexity of* $f(\ )$ *relative to* Ω. We note that the $C_\Omega(f)$ and $C_{\Omega'}(f)$ differ by at most a small constant factor since each element of Ω' can be replaced by a few elements from Ω and vice versa. For this reason, we let Ω be the set of 2-input binary logic elements and we assume that memory cells are binary.

2. A Bound on the Maximum Time to Compute a Function

The Inequality (1) assumes that for $1 \leqq i \leqq k$ the number of cycles for which S_i is used, namely, T_i, is fixed for the function $f(\ )$, that is, the T_i's do not depend on the arguments of the function. For many algorithms [which define finite functions $f(\ )$] the running time of the algorithm does depend on the data points. Consider, then, a single machine constructed with logic elements from Ω and individually accessed binary memory cells which computes $f(\ )$ with a variable number of cycles. Such a machine effectively has the power to turn off its own clock, since it stops at the end of its computation. If the machine is to be modified so that it always runs the same length of time on all data points [and so that Inequality (1) is applicable], the machine must be modified.

The value of the function must appear in the last state of the machine if its running time is variable. Thus, if the machine can be modified so that the value of the function

---

[1]Consultant, Division of Engineering, Brown University, Providence, R.I.

[2]Savage, J. E., "The Complexity of Decoders—Part II: Computational Work and Decoding Time," *IEEE Trans. Inform. Theory* (to be published).

computed can be saved and yet the machine is allowed to run until T_{max}, the maximum number of cycles required on any data point, then we can apply Inequality (1). Let the value of the function be stored in b_s bits of storage. Then, the binary signal which normally terminates a computation by turning off the clock can be used as an input to b_s *and* gates which transfer the value of the function to auxiliary storage. (Here we assume that these b_s bits always fall in the same positions in computer memory in the original machine.) With these modifications and with the addition of only b_s elements from Ω, we have constructed a sequential machine which has canonical form and runs T_{max} cycles on all inputs.

Proposition 1. *Let S be a sequential machine which computes a finite function $f(\ )$ which is such that each point of its range can be represented with b_s binary digits. Let S be constructed with 2-input binary logic elements and binary memory cells and let X be the number of logic elements it contains. Suppose that T_{max} is the maximum number of cycles required by S to compute $f(\ )$ on any point in its range. Then,*

$$(X + b_s)\, T_{max} \geqq C_\Omega(f) \tag{2}$$

We see that Inequality (2) is very close to Inequality (1) for a single machine when b_s is small relative to X. It should be noted that Inequality (2) only limits the size of T_{max} and does not limit the average number of cycles (which in fact may be a more important quantity). One can argue, however, that if the number of points in the domain of $f(\ )$ which require a number of cycles near T_{max} is small, then $f(\ )$ can be modified by deleting these points with a small effect on $C_\Omega(f)$. This would suggest that Inequality (2) may often hold at least approximately when T_{max} is replaced by the average number of cycles.

3. Storage Required for Autonomous Computation

Suppose that S computes $f(\ )$ and receives no external inputs, that is, all data needed to compute $f(\ )$ is part of the initial state of S. We say that S computes $f(\ )$ autonomously. Let s_0 be the initial state of S and $s_1, \cdots, s_T$ the successive states of S entered during the T cycles it executes while computing $f(\ )$ on some data point. We observe that no two of these states can be equal if $f(\ )$ is to be computed in finite time. Otherwise, the state repeated will be repeated infinitely many times.

Proposition 2. *Let S compute $f(\ )$ with no more than T_{max} cycles and let S have X logic elements from Ω and use b_s bits to specify a point in the range of $f(\ )$. Then, the number of bits of storage in S, namely, $|S|$, must satisfy*

$$|S| \geqq \log_2 (T_{max} + 1) \geqq \log_2 \left[\frac{C_\Omega(f)}{X + b_s} + 1 \right] \tag{3}$$

Proposition 3. *Let S consist of a small sequential machine which is attached to a very large random-access or tape storage which has $|S|$ words of storage each of which contains m bits. Then,*

$$(m + 1)\, |S|\, T_{max} (1 + \epsilon) \geqq C_\Omega(f) \tag{4}$$

for large $|S|$ where $\epsilon = \epsilon(|S|) > 0$.

This proposition follows from Inequality (2) and an earlier result (SPS 37-64, Vol. II, pp. 30–33) on the effective computing power of storage. The bound of Inequality (4) is stronger than that of Inequality (3) except when T_{max} is very large as could be expected since the storage assumed in Proposition 3 permits access to single m bit words at a time.

The complexity $C_\Omega(f)$ is large for most binary functions from $\{\Sigma_2\}^n$ to $\{\Sigma_2\}^q$ where $\Sigma_2 = \{0, 1\}$. In fact, it can be shown (Ref. 1, Theorem 2) that $C_\Omega(f)$ is greater than

$$C_\Omega(f) \geqq \frac{1}{2} \frac{q\, 2^n}{n + \log_2 q} (1 - \epsilon) \tag{5}$$

for most of these function for large n when $0 < \epsilon < 1$. Thus, the bound of Inequality (3) grows as $N + \log q$ when X is small and n and q are large. The bound of Inequality (4), however, has a much larger rate of growth.

4. Computation of Simple and Complex Functions

We turn next to the lower bound on computational work given by Inequality (1) and ask whether it can be reached for arbitrary functions f. We show that the bound can nearly be reached for simple and complex Boolean functions. (Every finite function is equivalent, up to isomorphism, to the direct product of Boolean functions of restricted domain.) We do not know whether the same is true for functions whose complexity is not at these extremes.

Consider the function $f_A(y) = y_1 y_2 \cdots y_n$ which is the *and* of y_1, y_2, y_n. $C_\Omega(f_A)$ is equal to $n - 1$ as we show by observing that (1) f_A is dependent on all inputs, (2) a

single logic element can reach at most two inputs, and (3) each additional element allows at most one additional input to be reached; also, f_A can be computed with exactly $n-1$ 2-input *and* gates, so $C_\Omega(f_A) = n-1$.

The work done to compute f_A with a single machine which contains X logic elements and which executes T cycles is given by $\chi = XT$. Clearly, then, the lower bound can be achieved when $T = 1$. When $T = k$, the machine shown in Fig. 1 will compute f_A and this machine contains an $l = \lceil n/k \rceil$-input[3] *and* gate, an *or* gate, and $k\lceil n/k \rceil + 1$ binary memory cells. The machine reads in the n inputs in groups of $\lceil n/k \rceil$ digits and, when n/k is not an integer, the last group will contain $k\lceil n/k \rceil - n$ entries which are 1. The number of logic elements in the machine is $X = \lceil n/k \rceil$ and the work done by it is

$$\chi = k\left\lceil \frac{n}{k} \right\rceil \geqq n-1 \tag{6}$$

with near equality for all $1 \leqq k \leqq n$, especially when k divides n.

Consider next the function $f_B(\mathbf{y}) = y_1^{c_1} y_2^{c_2} \cdots y_n^{c_n}$

[3] $\lceil x \rceil$ is the smallest integer equal to or greater than x.

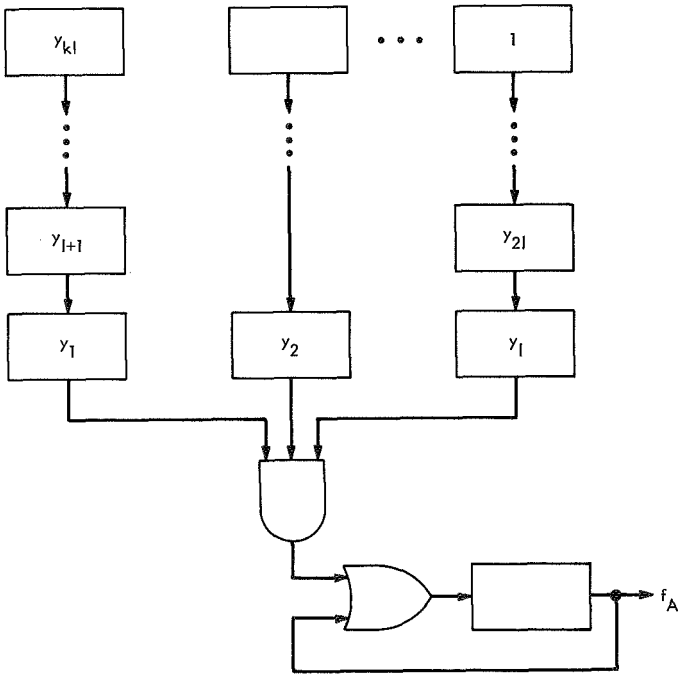


Fig. 1. Sequential machine which computes f_A

where

$$y_i^{c_i} = \begin{cases} y_i, & c_i = 0 \\ \bar{y}_i, & c_i = 0 \end{cases} \tag{7}$$

$\bar{y}_i$ is the complement of y_i and the product denotes the logical *and*. *Again* $C_\Omega(f_B) = n-1$, which follows from an argument analogous to that given above. The machine shown in Fig. 2 computes f_B in $T = k$ cycles and contains $l = \lceil n/k \rceil$ modulo-2 adders, $\lceil n/k \rceil - 1$ 2-input *and*'s, and one *or* for a total of $X = 2\lceil n/k \rceil$ logic elements. It has $2\lceil n/k \rceil + 1$ memory cells and does a work χ given by

$$\chi = 2k\left\lceil \frac{n}{k} \right\rceil \geqq n-1 \tag{8}$$

which is about a factor of 2 away from the bound when k divides n.

Clearly, the functions f_A and f_B are examples of simple functions. Now we turn to the computation of arbitrary

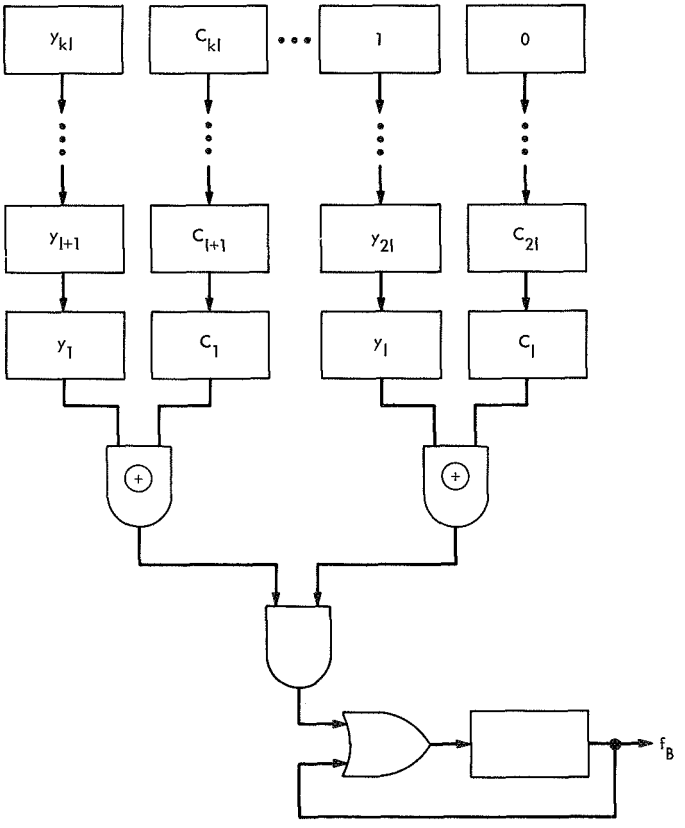


Fig. 2. Sequential machine which computes f_B

Boolean functions. (These are functions from Σ_2^n to Σ_2.) As mentioned above, the fraction of these functions for which

$$C_\Omega(f) \geqq \frac{1}{2}\frac{2^n}{n}(1-\epsilon) \tag{9}$$

where $0 < \epsilon < 1$, approaches 1 with increasing n. On the other hand, Lupanov (Ref. 2) has shown that, given $\epsilon > 0$, there exists an $N = N(\epsilon)$ such that, for $n \geqq N$, every Boolean function of n variables can be realized with

$$C_\Omega(f) \leqq \frac{2^n}{n}(1+\epsilon) \tag{10}$$

where $N(\epsilon) \to \infty$ as $\epsilon \to 0$. This means that the complexity of most Boolean functions of N variables is very large and grows nearly exponentially with N.

Any Boolean function f of n variables can be expanded as the logical *or* of minterms and the minterm $y_1^{c_1} y_2^{c_2} \cdots y_n^{c_n}$ is included if and only if f has value 1 when $y_i = \bar{c}_i$ for each $i, 1 \leqq i \leqq n$. The machine shown in Fig. 3 calculates minterms individually and successively using a stored list of all vectors $\mathbf{c} = (c_i, c_2, \cdots, c_n)$ which correspond to minterms in f. The logic unit contains n modulo-2 adders and one n-input *and* gate constructed from $n-1$ 2-input *and* gates and has one *or* gate for a total of $X = 2n$ gates. The machine executes $T = M$ cycles, where $M (\leqq 2^n)$ is the number of minterms in the function. Clearly, k copies of the minterm circuit and $k-1$ additional *or* gates could be used to compute the function in $T = \lceil m/k \rceil$ cycles. The more general machine will contain $x = 2nk + k - 1$ logic elements and do a computational work

$$\chi = [(2n+1)k - 1]\left\lceil \frac{M}{k} \right\rceil \geqq 2nM \tag{11}$$

with near equality when k divides M. Since most functions require an M near 2^{n-1}, we have that

$$\chi \simeq n2^n \tag{12}$$

for almost all Boolean functions of n variables when n is large. Comparing Inequality (12) with Inequalities (9) and (10), we find that

$$\chi \simeq n^2 C(f) \tag{13}$$

for most of these functions, and we have:

PROPOSITION 4. *Most Boolean functions f of n variables can be computed with a computational work* $\chi = XT$ *which is near the minimum required for f for all* $1 \leqq T \leqq M$ *where M is the number of minterms in the sum-of-products expansion of f.*

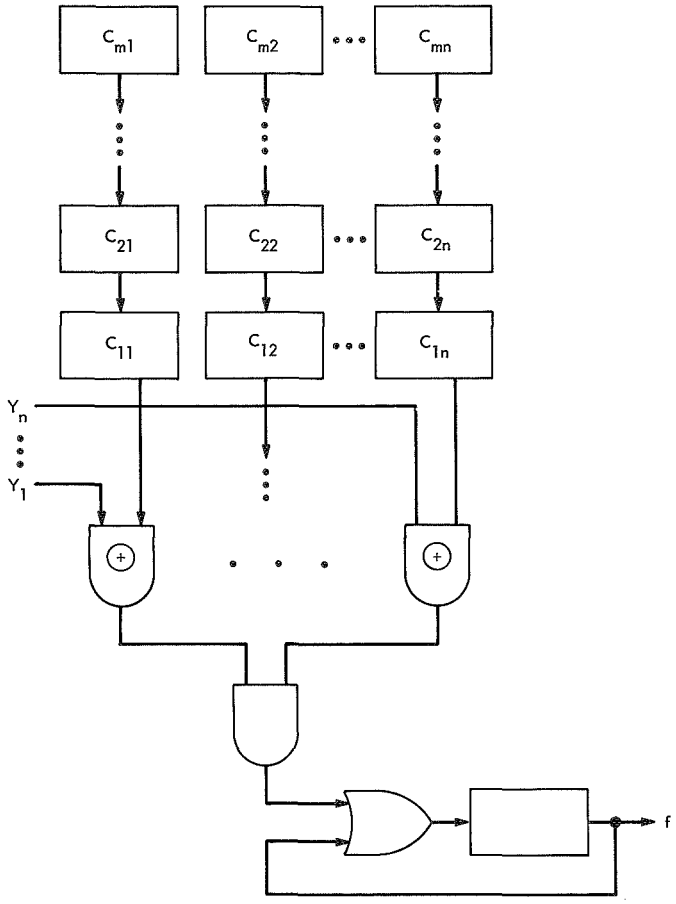


Fig. 3. Sequential machine which computes Boolean functions

It is worth noting that the number of memory cells in the above general machine is much larger than the bound of Inequality (3) but closer to Inequality (4). The storage is organized so that one word is effectively read per cycle.

5. Work Potential of a Computer

The *work potential* of a machine $\chi(t)$ is the computational work which it can do in t seconds. A machine which has X logic elements and a cycle duration D_0 has work potential

$$\chi(t) = X\frac{t}{D_0}$$

for t, a multiple of D_0. Suppose, however, that a machine has access to many types of storage with different cycle times. In particular, assume that it has a central processor

with X logic elements and cycle time D_0 and K storage units with cycle times $D_1, D_2, \cdots, D_k$ and effective computing powers $P_1, \cdots, P_k$ (the number of logic elements in an equivalent canonical machine). Then, if we order the storage units so that $D_0 \leq D_1 \leq D_2 \leq \cdots \leq D_k$, and if we assume that D_0 divides D_i, $1 \leq i \leq k$, then the work potential of the machine is given by

$$\chi(t) = t\left\{\frac{X}{D_0} + \sum_{j=1}^{J(t)} \frac{P_j}{D_j}\right\} \tag{14}$$

where $J(t)$ is such that

$$D_{J(t)} \leq t < D_{J(t)+1} \tag{15}$$

A typical plot of $\chi(t)$ is given in Fig. 4.

Many machines exist in which the slowest storage unit (such as tape) has a cycle time which is 100 to 200 times as great as the cycle time of the central processor. Since the computing power P_j is proportional to storage capacity, it is likely that P_k/D_k is much larger than X/D_0. If this is the case, the curve $\chi(t)$ will rise steeply for t near D_k and one could expect most jobs to be completed in a time near D_k if they are matched to the storage available.

The work potential rises with a slope X/D_0 if

$$P_j/D_j << X/D_0$$

so that the storage units do not improve the work potential much. On the other hand, if $P_j/D_j >> X/D_0$, the work potential may be dominated by a subset of the storage units. Since storage units are costly, a *good rule of thumb may be to choose storage units which satisfy*

$$\frac{P_j}{D_j} \simeq \frac{P_i}{D_i} \tag{16}$$

In the absence of other criteria, this rule may prove useful.

6. A Quantum-Mechanical Bound on Complexity

In this subsection, we derive a bound on the maximal complexity of any function which is computable with E ergs of energy in t seconds. Consider a sequential machine which has binary logic elements and executes T cycles to compute f. The output of each logic element can be in either of two states and the elements connected to it must make a state determination. Let the two states correspond to energy levels differing by ΔE. Then a state determination will require the consumption of at least ΔE units of energy, and the total energy E required to operate X logic elements must be at least $X\Delta E$. The duration of a cycle cannot be less than the time required for a state determination, or Δt. Then, E and the number of cycles T are bounded by

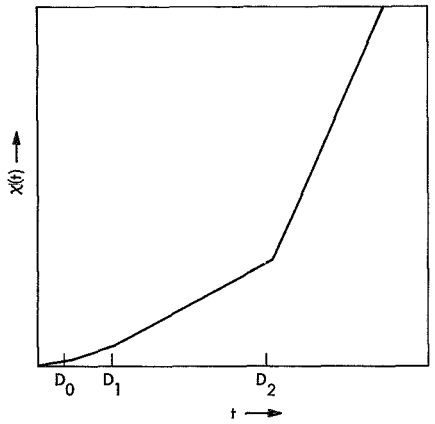


Fig. 4. Work potential

$$X \leq \frac{E}{\Delta E}\, T \leq \frac{t}{\Delta t} \tag{17}$$

where t is the number of seconds for which the machine is active. Then, the computational work of which it is capable is bounded by

$$\chi = XT \leq \frac{Et}{\Delta E \Delta t} \tag{18}$$

But the Heisenberg uncertainty relation says that measurements cannot be made unless

$$\Delta E \Delta t \geq \frac{h}{2\pi} \tag{19}$$

where h is Planck's constant with value $h = 6.624 \times 10^{-27}$ erg-second.

Therefore, function f cannot be computed by such a machine unless

$$C_\Omega(f) \leq \chi \leq \frac{(Et)\,2\pi}{h} = 0.95\,(Et)\,(\text{erg-second})^{-1}\,10^{27} \tag{20}$$

While the bound may seem to be extremely large, it is worth noting that most Boolean functions with only 110 variables cannot be computed with one erg of energy in one second. This follows from the application of the bounds of Inequalities (9) and (10).

7. A Comment on the Sorting Problem

Suppose a list of L integers from the set $\{0,1,2,\cdots,I-1\}$ is formed and we wish to order the list in the order of increasing terms. If the sorting is done in binary, a binary function with $L\lceil\log_2 I\rceil$ binary inputs and the same number of outputs will define the sorting. Consider the first sorted digit. It could be any one of the digits on the original list, so we deduce that the first output is a function of all $L\lceil\log_2 I\rceil$ inputs. Consequently,

$$C_\Omega(f) \geqq L\lceil\log_2 I\rceil - 1$$

if 2-input binary logic elements are used. This implies that the computational work required to sort must grow at least linearly with L. If no overlap exists between the circuits used to compute the $L\lceil\log_2 I\rceil$ outputs, then $C_\Omega(f)$ will grow as the square of L.

There are sorting procedures for two tape machines which sort in a number of cycles proportional to L. The effective computing power of these storage units is proportional to L, so the work done is proportional to L^2.

8. Conclusions

Six topics tied together by their relation to computational work have been discussed. One should interpret the results presented as early contributions to a nascent computer science.

References

1. Savage, J. E., "The Complexity of Decoders—Part I: Classes of Decoding Rules," *IEEE Trans. Inform. Theory*, Vol. IT-15, No. 6, Nov. 1969.
2. Lupanov, O. B., "A Method of Circuit Synthesis," *Z. zv. Vysshikh Uchebn. Zavedenii, Radiofiz*, No. 1, 1958.

C. Communication Statistics: Finite-Sample Quantile Estimation, *I. Eisenberger*

1. Introduction

Asymptotically optimum, unbiased estimators of the unknown mean μ and unknown standard deviation σ of a normal distribution using k sample quantiles are derived in Ref. 1 and given there in Tables 1 and 2 for $k = 2, 4, 6, \cdots, 20$. The asymptotic efficiency of the estimators relative to the best estimators using all the sample values are also given. However, since sample sizes are never infinite, the question arises as to the usefulness of the estimators for moderate values of the sample size n. The purpose of this article, therefore, is to give the exact moments of the sample quantiles used in estimating μ and σ for $k \leqq 6$ and $n = 50, 100, 200$. The expected value of the estimators of σ and the estimator of μ using one quantile will also be given, as well as the efficiency of all the estimators considered. Due to the symmetry of the normal density and the use of pairs of symmetric quantiles in estimating μ for $k > 1$, the estimators of μ are unbiased for $k > 1$.

2. The Asymptotic and Exact Distribution of Sample Quantiles

To define a quantile, consider a sample of n independent sample values $x_1, x_2, \cdots, x_n$ taken from a distribution of a continuous type with distribution function $F(x)$ and density function $f(x)$. The quantile of order p of the distribution or population, denoted by $\zeta(p)$, is defined as the root of the equation $F(\zeta) = p$, that is,

$$p = \int_{-\infty}^{\zeta(p)} f(x)\,dx$$

The corresponding sample quantile $z(p)$ is defined as follows. If the sample values are arranged in non-decreasing order of magnitude,

$$x_{(1)} \leqq x_{(2)} \leqq \cdots \leqq x_{(n)}$$

then $x_{(i)}$ is called the ith order statistic and

$$z(p) = x_{[np]+1}$$

where $[np]$ denotes the greatest integer $\leqq np$.

If $f(x)$ is differentiable in some neighborhood of each quantile value, it has been shown (Ref. 2) that the joint distribution of any number of quantiles is asymptotically

normal as $n \to \infty$ and that, asymptotically,

$$E(z(p)) = \zeta(p)$$

$$\mathrm{Var}(z(p)) = \frac{p(1-p)}{nf^2(\zeta(p))}$$

$$\rho_{12} = \left[\frac{p_1(1-p_2)}{p_2(1-p_1)}\right]^{1/2}$$

where ρ_{12} is the correlation between $z(p_1)$ and $z(p_2)$, $p_1 < p_2$.

Now let $g_1(x)$ denote the exact density function of the sample quantile z_1 of order p_1 and let $g_2(y)$ denote the exact density function of the sample quantile z_2 of order p_2, $p_1 < p_2$. Let $\mu_1 = [np_1]$ and let $\mu_2 = [np_2]$. Then it is also shown in Ref. 2 that $g_1(x)$ and $g_2(y)$ are given by

$$\left.\begin{aligned} g_1(x) &= \binom{n}{\mu_1}(n-\mu_1)[F(x)]^{\mu_1}[1-F(x)]^{n-\mu_1-1}f(x), \quad -\infty < x < \infty \\ g_2(y) &= \binom{n}{\mu_2}(n-\mu_2)[F(y)]^{\mu_2}[1-F(y)]^{n-\mu_2-1}f(y), \quad -\infty < y < \infty \end{aligned}\right\} \tag{1}$$

The joint density $g(x, y)$ of z_1 and z_2, which is needed in order to compute the variance of the estimators, will now be derived. The joint probability that z_1 is situated in an interval $(x, x + dx)$ and z_2 is situated in an interval $(y, y + dy)$ is identical with the probability that, among the n sample values, μ_1 are $< x$, $n - \mu_2 - 1$ are $> y + dy$, $\mu_2 - \mu_1 - 1$ are between $x + dx$ and y, one falls between x and $x + dx$, while the remaining value falls between y and $y + dy$.

Hence, since the observations are independent,

$$g(x,y) = 2\binom{n}{\mu_1}\binom{n-\mu_1}{n-\mu_2-1}\binom{\mu_2-\mu_1+1}{\mu_2-\mu_1-1}[F(x)]^{\mu_1}[1-F(y)]^{n-\mu_2-1}[F(y)-F(x)]^{\mu_2-\mu_1-1}f(x)f(y),$$

$$-\infty < x < \infty \ , \ x < y < \infty \tag{2}$$

3. The Quantile Estimators of μ and σ

The estimators of μ for $k > 1$, as given in Table 1 of Ref. 1, are of the form

$$\hat{\mu} = \sum_{i=1}^{k/2} a_i \left[z(p_i) + z(p_{k-i+1}) \right]$$

where

$$p_i < p_j \qquad \text{if } i < j$$

$$p_i = 1 - p_{k-i+1}$$

and

$$\sum_{i=1}^{k/2} a_i = 0.5$$

This means that μ is a linear combination of pairs of symmetric quantiles such that, for any choice of the p_i and coefficients a_i which satisfies the above conditions, $E(\hat{\mu}) = \mu$. However, for each value of k, the orders of the quantiles and the values of the a_i were chosen so as to minimize the asymptotic variance of $\hat{\mu}$. For $k = 1$, the asymptotically unbiased estimator of μ is given by

$$\hat{\mu} = z(0.5)$$

In all cases, the efficiency of $\hat{\mu}$ was defined as

$$\mathrm{Eff}(\hat{\mu}) = \frac{\sigma^2}{n\,\mathrm{Var}(\hat{\mu})} \tag{3}$$

The estimators of σ in Table 2 of Ref. 1 are also linear combinations of pairs of symmetric quantiles of the form

$$\hat{\sigma} = \sum_{i=1}^{k/2} b_i \left[z(p_{k-i+1}) - z(p_i) \right]$$

where the orders of the quantiles and the coefficients b_i were chosen so as to satisfy simultaneously the conditions that $\hat{\sigma}$ be asymptotically unbiased and that the asymptotic

variance of σ be a minimum. The efficiency of σ was defined as

$$\text{Eff}(\hat{\sigma}) = \frac{\sigma^2}{2n \operatorname{Var}(\hat{\sigma})} \tag{4}$$

4. Exact Moments of the Quantiles and Estimators

Using Eq. (1), the means and variances of the quantiles used in estimating μ and σ for the cases under consideration were computed for $n = 50, 100$, and 200 and are given in Table 1. Since

$$E(z(p)) - \mu = -[E(z(1-p)) - \mu]$$

and

$$\operatorname{Var}(z(p)) = \operatorname{Var}(z(1-p))$$

the results for $k > 1$ are given only for $i \leqq k/2$.

Using Eq. (2) and the results previously obtained, the exact correlations between $z(p_i)$ and $z(p_j)$ were computed and are given, along with the asymptotic values, in Table 2.

The normalized expected values of the estimators of σ were computed and are given in Table 3. The asymptotic values are, of course, equal to one for all k. The normalized bias of the estimator of μ using one quantile is shown in Table 1 as $[E(zp) - \mu]/\sigma$ for $k = 1$.

The variances of the estimators of μ and σ were computed and, by means of Eqs. (3) and (4), the efficiencies of the estimators were determined and are also given in Table 3.

The results of this investigation indicate that even for moderate values of n, the quantile estimators of the unknown parameters of a normal population, although derived on the basis of the asymptotic distribution and moments of the quantiles, can be useful.

References

1. Eisenberger, I., and Posner, E. C., "Systematic Statistics Used for Data Compression in Space Telemetry," *J. Am. Statist. Assoc.*, Vol. 60, pp. 97–133, Mar. 1965. Also, Technical Report 32-510, Jet Propulsion Laboratory, Pasadena, Calif., Oct. 1, 1963.
2. Cramer, H., *Mathematical Methods of Statistics*. Princeton University Press, Princeton, N.J., 1946.

Table 1. Asymptotic and exact moments of the quantiles used to estimate μ and σ for $k = 1, 2, 4, 6$ and $n = 50, 100, 200$

k	p	Asymptotic		$n = 200$		$n = 100$		$n = 50$	
		$\frac{E(z(p)) - \mu}{\sigma}$	$\frac{n \operatorname{Var}(z(p))}{\sigma^2}$	$\frac{E(z(p)) - \mu}{\sigma}$	$\frac{n \operatorname{Var}(z(p))}{\sigma^2}$	$\frac{E(z(p)) - \mu}{\sigma}$	$\frac{n \operatorname{Var}(z(p))}{\sigma^2}$	$\frac{E z(p)) - \mu}{\sigma}$	$\frac{n \operatorname{Var}(z(p))}{\sigma^2}$
For estimating μ									
1	0.5	0	1.5708	0.00625	1.5657	0.0125	1.5622	0.0249	1.5556
2	0.2709	−0.610	1.8007	−0.605	1.7945	−0.596	1.7816	−0.610	1.7827
4	0.1068	−1.244	2.8170	−1.238	2.8080	−1.249	2.8255	−1.218	2.7136
	0.3512	−0.382	1.6567	−0.378	1.6512	−0.371	1.6434	−0.384	1.6412
6	0.0540	−1.607	4.2465	−1.617	4.3560	−1.591	4.1423	−1.628	4.2930
	0.1915	−0.872	2.0824	−0.868	2.0779	−0.857	2.0539	−0.873	2.0621
	0.3898	−0.280	1.6163	−0.285	1.6138	−0.292	1.6117	−0.278	1.6000
For estimating σ									
2	0.0694	−1.483	3.6599	−1.492	3.7205	−1.507	3.7641	−1.463	3.5044
4	0.0230	−1.995	7.5562	−1.997	7.7199	−1.946	6.8769	−1.855	5.8242
	0.1271	−1.140	2.5568	−1.137	2.5578	−1.147	2.5669	−1.119	2.4833
6	0.0104	−2.312	13.5558	−2.229	11.5449	−2.148	9.5767	−2.248	10.7783
	0.0548	−1.600	4.2100	−1.617	4.3800	−1.591	4.1470	−1.627	4.2964
	0.1696	−0.956	2.2062	−0.963	2.2240	−0.971	2.2256	−0.049	2.1719

Table 2. Asymptotic and exact values of the correlation between the quantiles used to estimate μ and σ for $k = 2, 4, 6$ and $n = 50, 100, 200$

k	Correlation	$\hat{\mu}$				$\hat{\sigma}$			
		Asymptotic values	Exact values			Asymptotic values	Exact values		
			$n = 200$	$n = 100$	$n = 50$		$n = 200$	$n = 100$	$n = 50$
2	ρ_{12}	0.3716	0.3746	0.3812	0.3743	0.0748	0.0714	0.0695	0.0787
4	ρ_{14}	0.1196	0.1203	0.1180	0.1274	0.0235	0.0228	0.0270	0.0346
	ρ_{23}	0.5413	0.5446	0.5522	0.5427	0.1456	0.1457	0.1437	0.1533
	$\rho_{12} = \rho_{34}$	0.4700	0.4696	0.4627	0.4855	0.4021	0.3920	0.4301	0.4757
	$\rho_{13} = \rho_{24}$	0.2544	0.2559	0.2554	0.2631	0.0585	0.0577	0.0623	0.0728
6	ρ_{16}	0.0571	0.0547	0.0594	0.0554	0.0105	0.0130	0.0169	0.0141
	ρ_{25}	0.2367	0.2381	0.2439	0.2386	0.0580	0.0542	0.0592	0.0553
	ρ_{34}	0.6383	0.6324	0.6271	0.6432	0.2042	0.2001	0.1982	0.2089
	$\rho_{12} = \rho_{56}$	0.4909	0.4768	0.4927	0.4820	0.4258	0.4753	0.5271	0.4922
	$\rho_{13} = \rho_{46}$	0.2989	0.2938	0.3076	0.2945	0.2268	0.2549	0.2899	0.2610
	$\rho_{14} = \rho_{36}$	0.1920	0.1862	0.1930	0.1890	0.0463	0.0515	0.0578	0.0434
	$\rho_{15} = \rho_{26}$	0.1163	0.1141	0.1203	0.1150	0.0247	0.0267	0.0316	0.0280
	$\rho_{23} = \rho_{45}$	0.6089	0.6126	0.6231	0.6094	0.5328	0.5148	0.5423	0.5146
	$\rho_{24} = \rho_{35}$	0.3890	0.3880	0.3912	0.3919	0.1088	0.1042	0.1003	0.1077

Table 3. The normalized expected value of the estimators of σ and the asymptotic and exact efficiencies of the estimators of μ and σ for $n = 50, 100, 200$

k	$E(\hat{\sigma})/\sigma$			Asymptotic		$n = 200$		$n = 100$		$n = 50$	
	$n = 200$	$n = 100$	$n = 50$	Eff $(\hat{\mu})$	Eff $(\hat{\sigma})$	Eff $(\hat{\mu})$	Eff $(\hat{\sigma})$	Eff $(\hat{\mu})$	Eff $(\hat{\sigma})$	Eff $(\hat{\mu})$	Eff $(\hat{\sigma})$
1				0.6366		0.6387		0.6401		0.6428	
2	1.0068	1.0169	0.9872	0.8098	0.6522	0.8108	0.6388	0.8126	0.6270	0.8164	0.6801
4	1.0000	0.9929	0.9585	0.9201	0.8244	0.9215	0.8208	0.9234	0.8415	0.9263	0.9035
6	0.9985	0.9861	0.9980	0.9560	0.8943	0.9582	0.8925	0.9596	0.9184	0.9615	0.9165

D. Frequency Generation and Control: Analysis of Random Modulation in Amplifier Circuits,

A. Sward and G. Thompson

1. Introduction

The growing demand for more stable frequency sources has been answered in part with the advent of the hydrogen maser frequency standard. The requirements for the frequency standard and some preliminary data are described in SPS 37-59, Vol. II, pp. 40–43. This demand for stability has imposed a number of stringent requirements upon the phase noise and phase stability of receivers for the Deep Space Network. Therefore, much effort has been put into analyzing and minimizing phase noise in RF communication equipment.

Until recently, it was difficult to make low-level phase noise measurements. Phase jitter measurements were restricted to voltage-controlled oscillators, frequency synthesizers, etc., where phase jitter was high. However, the availability of more sensitive measuring equipment has enabled better measurements to be made (SPS 37-64, Vol. II, pp. 55–59).

In general, a sinusoidal carrier may be subjected to either of two kinds of modulation, angle modulation and

amplitude modulation. Consider a carrier signal $x(t)$ modulated by the low-frequency processes $A(t)$ and $\phi(t)$.

$$x(t) = [1 + A(t)] \cos [\omega_0 t + \phi(t)] \tag{1}$$

Since $A(t)$ and $\phi(t)$ are low-frequency random processes, the effect upon $x(t)$ is a broadening or degradation of the spectrum. The question, however, is to what extent each of the two types of noise contributes to the broadening of the spectrum.

In this article, a theoretical basis for meaningful amplitude-modulation (AM) and phase-modulation (PM) measurements is developed. In the circuits encountered in a receiver (amplifiers, frequency multipliers, dividers, etc.), the dominant AM and PM noise processes have a power spectral density with a $1/f$ behavior. To compare the performance of one circuit relative to another, a measure $\mathcal{J}$ of the power in the AM and PM processes is defined. The goals for the circuit design can be seen from the functional behavior of $\mathcal{J}$.

2. General Theory

Consider the complex signal $x_c(t)$ of the form

$$x_c(t) = [1 + A(t)] \exp [j\omega_0 t + j\phi(t)] \tag{2}$$

where ω_0 is the frequency and where $A(t)$ and $\phi(t)$ represent random amplitude and phase modulation, respectively. The functions $A(t)$ and $\phi(t)$ are assumed to be independent, zero-mean, baseband random processes.

The autocorrelation function $R_{x_c}(\tau)$ of $x_c(t)$ is defined to be (Ref. 1, p. 37)

$$R_{x_c}(\tau) = \langle x_c(t + \tau) x_c^*(t) \rangle \tag{3}$$

where the asterisk (*) denotes the complex conjugate and the brackets $\langle\ \rangle$ denote an ensemble average. Substituting Eq. (2) into this expression and using the fact that $A(t)$ and $\phi(t)$ are independent, one finds that

$$R_{x_c}(\tau) = \langle [1 + A(t + \tau)] [1 + A(t)] \rangle \times \langle \exp [j\phi(t + \tau) - j\phi(t)] \rangle \exp (j\omega_0 \tau) \tag{4}$$

The first ensemble average in Eq. (4) can be written in terms of the autocorrelation $R_A(\tau)$ of $A(t)$ as

$$\langle [1 + A(t + \tau)] [1 + A(t)] \rangle = 1 + R_A(\tau) \tag{5}$$

By defining the low-frequency process $m(t)$ as

$$m(t) = \exp [j\phi(t)] \tag{6}$$

the second ensemble average in Eq. (4) can be written

$$R_m(\tau) = \langle \exp [j\phi(t + \tau) - j\phi(t)] \rangle \tag{7}$$

where $R_m(\tau)$ is the autocorrelation function of $m(t)$. One can now express $R_{x_c}(\tau)$ in terms of $R_A(\tau)$ and $R_m(\tau)$ as

$$R_{x_c}(\tau) = [1 + R_A(\tau)] R_m(\tau) \exp (j\omega_0 \tau) \tag{8}$$

In amplifiers, the phase-modulation process $\phi(t)$ has been observed experimentally (SPS 37-64, Vol. II) to have a small magnitude ($|\phi(t)| << 1$). For processes with finite power, such as the model for $1/f$ noise given in *Subsection 3*, Eq. (6) can be approximated by (Ref. 1, p. 123)

$$m(t) \approx 1 + j\phi(t) - \frac{\phi^2(t)}{2}$$

The autocorrelation $R_m(\tau)$ is then approximately

$$R_m(\tau) \approx 1 - R_\phi(0) + R_\phi(\tau)$$

where $R_\phi(\tau)$ is the autocorrelation function of $\phi(t)$. Hence, Eq. (8) can be written

$$R_{x_c}(\tau) \approx [1 + R_A(\tau)] [1 - R_\phi(0) + R_\phi(\tau)] \exp (j\omega_0 \tau) \tag{9}$$

The narrowband function of primary interest is the real part of $x_c(t)$ given by $x(t)$, as in Eq. (1),

$$\left.\begin{aligned} x(t) &= Re [x_c(t)] \\ x(t) &= [1 + A(t)] \cos [\omega_0 t + \phi(t)] \end{aligned}\right\} \tag{10}$$

which has the autocorrelation function $R_x(\tau)$ given by (Ref. 1, p. 110)

$$R_x(\tau) = \frac{1}{2} Re [R_{x_c}(\tau)]$$

From Eq. (9),

$$R_x(\tau) = \frac{1}{2} [1 + R_A(\tau)] [1 - R_\phi(0) + R_\phi(\tau)] \cos \omega_0 \tau \tag{11}$$

This is the time-domain result that was to be obtained. The second-order statistics of $x(t)$ are related to the second-order statistics of $A(t)$ and $\phi(t)$ through Eq. (11).

The Fourier transform of $R_x(\tau)$ is the power spectral density $S_x(\omega)$ of $x(t)$ given by (Ref. 2, p. 1)

$$S_x(\omega) = \int_{-\infty}^{\infty} R_x(\tau)\, e^{-j\omega\tau}\, d\tau$$

Applying this transformation to Eq. (11) and using well-known properties of the Fourier integral (Ref. 2, Chap. 2 and 3), one obtains

$$S_x(\omega) = \frac{1}{4}\{2\pi\,[1 - R_\phi(0)]\,\delta(\omega) + S_\phi(\omega) + [1 - R_\phi(0)]\,S_A(\omega) + \frac{1}{2\pi}\,S_A(\omega) \circledast S_\phi(\omega)\} \times \circledast\,[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)] \tag{12}$$

where $S_A(\omega)$ and $S_\phi(\omega)$ are the spectral densities of $A(t)$ and $\phi(t)$, respectively, and where $\circledast$ denotes the convolution operator [Ref. 2, p. 27]. From Eq. (12), one can see that the signal $x(t)$ is the sum of a discrete carrier component with power P_c given by

$$P_c = \frac{1}{2}\,[1 - R_\phi(0)] \tag{13}$$

and three sideband components which are the result of the amplitude and phase modulation. Note that if either $S_A(\omega)$ or $S_\phi(\omega)$ is identically zero, Eq. (12) reduces to the form of linear-phase or amplitude modulation, respectively. The primary restriction on the validity of Eq. (12) is that it applies to the case of low-power phase modulation ($|\phi(t)| << 1$).

3. Application to 1/f Noise

In the frequency synthesis and distribution circuits considered here, the power spectra of $A(t)$ and $\phi(t)$ have been found empirically (SPS 37-64, Vol. II) to have the form of $1/f$ noise. The model for these spectra is taken to be

$$S_A(\omega) = \begin{cases} \dfrac{b_A^2}{|\omega| + a} & \text{for } |\omega| \leq c \\ 0 & \text{for } |\omega| > c \end{cases} \tag{14}$$

and

$$S_\phi(\omega) = \begin{cases} \dfrac{b_\phi^2}{|\omega| + a} & \text{for } |\omega| \leq c \\ 0 & \text{for } |\omega| > c \end{cases} \tag{15}$$

where c is the effective 3-dB frequency of the narrowband circuit, where a $<<1$ is a constant, and where b_A and b_ϕ are empirically determined constants. The spectra are arbitrarily assumed to be zero for $|\omega| > c$, but this should have negligible effect on the results as long as the response of the circuit falls off faster than $1/f$ as $f \to \infty$. The parameter $a/2\pi$ is a very small number less than one cycle per year (Ref. 3) which has yet to be measured experimentally. For this reason, a is the same in both $S_A(\omega)$ and $S_\phi(\omega)$, and its actual value has little effect on the results to be obtained.

In Eq. (12), the convolution of $S_A(\omega)$ and $S_\phi(\omega)$ is one of the components of sideband modulation. Let $S_{AX\phi}(\omega)$ be the cross term between $A(t)$ and $\phi(t)$ so that (Ref. 2, p. 27)

$$S_{AX\phi}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_A(\xi)\, S_\phi(\omega - \xi)\, d\xi \tag{16}$$

Using the particular spectra in Eqs. (14) and (15), one can find that

$$S_{AX\phi}(0) = \frac{b_A^2\, b_\phi^2\, c}{\pi\, a\,(c + a)} \tag{17a}$$

$$S_{AX\phi}(\omega) = \frac{b_A^2\, b_\phi^2}{\pi}\,\frac{1}{\omega}\ln\frac{(\omega + a)\,(c - \omega + a)}{a\,(c + a)} + \frac{1}{\omega + 2a}\ln\frac{(\omega + a)}{a}, \qquad 0 < |\omega| \leq c \tag{17b}$$

$$S_{AX\phi}(\omega) = \frac{b_A^2\, b_\phi^2}{\pi}\,\frac{1}{\omega + 2a}\ln\frac{(c + a)}{(\omega - c + a)}, \qquad c < |\omega| \leq 2c \tag{17c}$$

$$S_{AX\phi}(\omega) = 0, \qquad 2c < |\omega| < \infty \tag{17d}$$

For the case of simultaneous amplitude and phase modulation with $1/f$ noise, the spectrum of the carrier signal is given by Eqs. (12), (14), (15), and (17).

4. The Comparative Measure $\mathcal{J}(\omega_1, \omega_2)$

In evaluating the amount of signal degradation produced by a circuit, some measure of the energy loss from the carrier to the sidebands is needed. Often spectral broadening of the carrier is important, but for the low-index modulation above, the broadening is of the order of 10^{-8} Hz, which cannot be resolved. Contrary to the random walk of the phase of an oscillator, the phase deviation of an amplifier cannot increase without bound because it is limited by the phase transfer characteristic of the amplifier. Of greater interest is a measure of the amount of sideband power to carrier power. For phase modulation alone, such a measure has been defined (Ref. 4). Another measure $\mathcal{J}(\omega_1, \omega_2)$ can be defined for the more general case of both amplitude and phase modulation as

$$\mathcal{J}(\omega_1, \omega_2) \equiv \frac{\text{single sideband power for } |\omega| \in [\omega_1, \omega_2]}{\text{total carrier power}} \tag{18}$$

per device, referred to the input. Since $1/f$ noise is the predominant spectral shape considered here, then, for $a << \omega_1 < \omega_2 << c$, $\mathcal{J}(\omega_1, \omega_2)$ will indicate the relative level of the modulation spectrum independently of a and c. This property can be seen by evaluating $\mathcal{J}(\omega_1, \omega_2)$ for the spectra of Eqs. (14) and (15).

Substituting the sideband and carrier components from Eq. (12) into Eq. (18) yields

$$\mathcal{J}(\omega_1, \omega_2) = \frac{\frac{1}{4\pi}\int_{\omega_1}^{\omega_2} \{S_\phi(\omega) + [1 - R_\phi(0)]\, S_A(\omega) + S_{AX\phi}(\omega)\}\, d\omega}{P_c} \tag{19}$$

where P_c is given by Eq. (13). From Eq. (17b), in the range $a << |\omega| << c$, $S_{AX\phi}(\omega)$ can be approximated by

$$S_{AX\phi}(\omega) \approx \frac{2b_A^2 b_\phi^2}{\pi}\frac{1}{\omega}\ln\frac{\omega}{a}, \qquad a << |\omega| << c$$

By using this approximation in Eq. (19), one can find that

$$\mathcal{J}(\omega_1, \omega_2) = \frac{(b_A^2 + b_\phi^2)\ln\frac{(\omega_2 + a)}{(\omega_1 + a)} + \frac{b_A^2 b_\phi^2}{\pi}\left[\ln\frac{\omega_1\omega_2}{a^2}\ln\frac{\omega_2}{\omega_1} - \ln\frac{(c+a)}{a}\ln\frac{(\omega_2+a)}{(\omega_1+a)}\right]}{2\pi\left[1 - \frac{b_\phi^2}{\pi}\ln\frac{(c+a)}{a}\right]}$$

The approximation that $a << \omega_1 < \omega_2 << c$ still holds so that

$$\mathcal{J}(\omega_1, \omega_2) \approx \ln\frac{\omega_2}{\omega_1}\,\frac{b_A^2 + b_\phi^2 + \frac{b_A^2 b_\phi^2}{\pi}\ln\frac{\omega_1\omega_2}{ca}}{2\pi\left(1 - \frac{b_\phi^2}{\pi}\ln\frac{c}{a}\right)} \tag{20}$$

The parameter a is not known so that one must estimate the relative sizes of the terms in Eq. (20). Experimentally (SPS 37-64, Vol. II), b_A^2 and b_ϕ^2 in Eq. (20) are on the order of 10^{-6}. For an amplifier bandwidth $c/2\pi = 10^6$Hz, one would have to observe the amplifier output for 10^{10^6} yr before the last terms in the numerator and denominator of Eq. (20) would be comparable to the first terms. Hence, Eq. (20) can be approximated as

$$\mathcal{J}(\omega_1, \omega_2) \approx \frac{1}{2\pi}(b_A^2 + b_\phi^2)\ln\frac{\omega_2}{\omega_1} \tag{21}$$

The approximation made in obtaining Eq. (21) from Eq. (20) should be checked when applying these results to a particular circuit.

The measure $\mathcal{J}(\omega_1, \omega_2)$ includes the effects of both amplitude and phase modulation. In practice, one often wishes to know the relative contributions of the two types of modulation. Hence, additional measures, $\mathcal{A}(\omega_1, \omega_2)$ for

amplitude modulation and $\mathcal{P}(\omega_1, \omega_2)$ for phase modulation (Ref. 4), are useful. For amplitude modulation alone define

$$\mathcal{A}(\omega_1, \omega_2) \equiv \frac{\text{single sideband AM power for } |\omega| \in [\omega_1, \omega_2]}{\text{total carrier power}} \tag{22}$$

per device, referred to the input. The carrier function that is amplitude-modulated only is given by $x(t)$ in Eq. (10) with $\phi(t) \equiv 0$. In this case, the spectral density is given by Eq. (12) with $S_\phi(\omega) \equiv 0$ so that $\mathcal{A}(\omega_1, \omega_2)$ can be expressed as

$$\mathcal{A}(\omega_1, \omega_2) = \frac{\frac{1}{4\pi}\int_{\omega_1}^{\omega_2} S_A(\omega)\, d\omega}{P_c}$$

where the carrier power P_c is given by Eq. (13) with $R_\phi(0) = 0$. Evaluating the above function for the $1/f$ spectrum of Eq. (14), one obtains

$$\mathcal{A}(\omega_1, \omega_2) = \frac{b_A^2}{2\pi} \ln \frac{(\omega_2 + a)}{(\omega_1 + a)} \tag{23}$$

Similarly, for phase modulation alone define (Ref. 4)

$$\mathcal{P}(\omega_1, \omega_2) \equiv \frac{\text{single sideband PM power for } |\omega| \in [\omega_1, \omega_2]}{\text{total carrier power}} \tag{24}$$

per device, referred to the input. When $A(t) \equiv 0$ in Eq. (10), then $S_A(\omega) \equiv 0$ in Eq. (12) so that

$$\mathcal{P}(\omega_1, \omega_2) = \frac{\frac{1}{4\pi}\int_{\omega_1}^{\omega_2} S_\phi(\omega)\, d\omega}{\frac{1}{2}[1 - R_\phi(0)]}$$

Evaluating this expression for the $S_\phi(\omega)$ in Eq. (15) and assuming $R_\phi(0) << 1$, one finds

$$\mathcal{P}(\omega_1, \omega_2) = \frac{b_\phi^2}{2\pi} \ln \frac{(\omega_2 + a)}{(\omega_1 + a)} \tag{25}$$

Comparing Eqs. (21), (23), and (25) and assuming $a << \omega_1 < \omega_2$, then one can write

$$\mathcal{J}(\omega_1, \omega_2) \approx \mathcal{A}(\omega_1, \omega_2) + \mathcal{P}(\omega_1, \omega_2) \tag{26}$$

This relationship implies that when the AM and the PM are both small, then the modulation processes can be handled separately.

5. Application to Circuit Design

A plot of Eq. (26) is given in Fig. 1. The design goal for a circuit should be clear from this plot. For a given value of $\mathcal{A}$, one should make $\mathcal{P} \approx \mathcal{A}$ to obtain nearly the lowest amount of signal degradation due to amplitude and phase modulation. Further reductions of $\mathcal{P}$ can realize at most a 3-dB improvement in $\mathcal{J}$. Obviously both $\mathcal{A}$ and $\mathcal{P}$ should be made as small as possible.

In practice, comparisons using the above ideas are very simple. If one standardizes on a 1-Hz bandwidth at 10 Hz, then a measurement of the power or rms voltage in this bandwidth can serve as a basis for the comparison of circuit performance. Detailed measurements and measurement techniques are presented in SPS 37-64, Vol. II.

6. Conclusions

This article has shown that random modulation of the amplitude and phase of a sinusoidal carrier produces sideband components which degrade the spectral characteristics of the ideal carrier. A measure of this degradation is given by the sideband to carrier power ratio $\mathcal{J}$ defined in Eq. (20). Evaluation of $\mathcal{J}$ in Eq. (23) for AM and PM by $1/f$ noise spectra shows that the two noise components

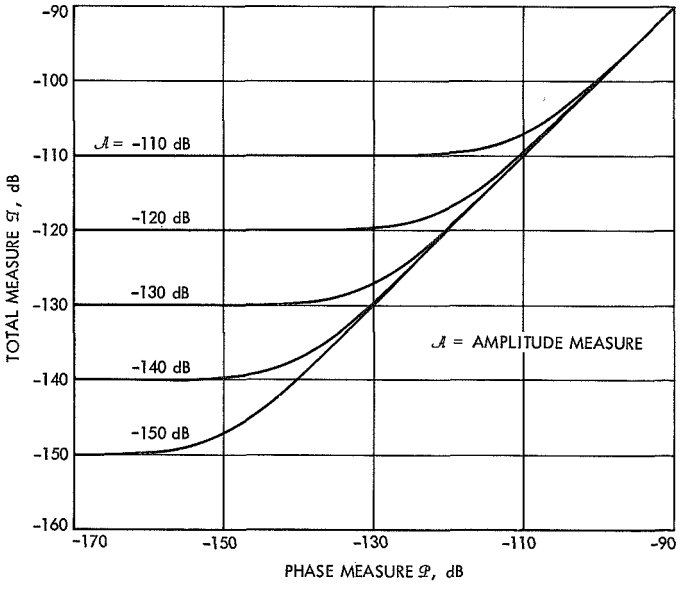


Fig. 1. Relative contributions of amplitude and phase noise to single sideband power

affect $\mathcal{J}$ symmetrically. The value of $\mathcal{J}$ is effectively minimized by making the AM and PM contributions equal and as small as possible.

References

1. Rowe, H. E., *Signals and Noise in Communication Systems.* D. Van Nostrand Co., Inc., Princeton, N.J., 1965.
2. Papoulis, A., *The Fourier Integral and Its Applications.* McGraw-Hill Book Co., Inc., New York, 1962.
3. Atkinson, W. R., Fey, R. L., and Newman, J., "Spectrum Analysis of Extremely Low Frequency Variations of Quartz Oscillators," *Proceedings of the IEEE,* Vol. 51, p. 379, Feb. 1963.
4. Halford, D., *Phase Noise of RF Amplifiers and Frequency Multipliers,* U.S. Government Memorandum 253.04. National Bureau of Standards, Boulder, Colorado, Oct. 25, 1967.

E. Frequency Generation and Control: Angle Demodulation Using State-Variable Techniques, *G. Thompson*

1. Introduction

Space communication systems typically use a sinusoidal carrier that is phase-modulated by a multitude of signals such as engineering and scientific data, command and control information, and integrated doppler frequency shift. Analyses of such systems generally require a large number of simplifying assumptions in order to obtain a tractable receiver design (SPS 37-58, Vol. III, pp. 42–50, and SPS 37-60, Vol. III, pp. 46–56). The purpose of this article is to formulate the problem using the state-variable approach, which enables one to design a receiver system and to predict its performance in a straightforward manner.

A fairly general form for the signal $g(t)$ detected by a space communications receiver is (SPS 37-58, Vol. III, pp. 42–50)

$$g(t) = (2P_s)^{1/2} \times \sin\left[\omega_0 t + \int_0^t w_d(\tau)\, d\tau + \sum_{k=1}^{K} s_k(t) \cos \omega_k t + \theta_0\right] \quad (1)$$

where

P_s = total signal power

ω_0 = nominal carrier frequency in sec^{-1}

$w_d(\tau)$ = residual doppler frequency shift

$s_k(t)$ = data signal which biphase-modulates the subcarrier $\cos \omega_k t$

θ_0 = unknown initial phase angle

The signal in Eq. (1) can be represented in the form

$$g(t) = (2P_s)^{1/2} \sin\left[\omega_0 t + H(t)\, X(t)\right] \quad (2)$$

where $H(t)$ is a $1 \times N$ row vector, and where $X(t)$ is an $N \times 1$ column vector of random processes or state variables $x_i(t)$, $i = 1, 2, \cdots, N$. The statistics of the $x_i(t)$ are chosen to model the unknown parameters and/or the random processes of the angle modulation in Eq. (1).

For signals of the form of Eq. (2), the receiver structure resulting from the state-variable approach (Ref. 1) can be significantly simplified. In *Subsection 2,* the receiver, called a nonlinear, recursive estimator (NLRE), is derived for the case where the state vector $X(t)$ can be modeled as the output of a linear system driven by white noise. The assumption that the time variations of $X(t)$ are slow compared to the carrier frequency ω_0(narrowband) allows some of the system parameters to be decoupled from the estimate of the signal. Considerable simplification of the system structure results.

At large values of signal-to-noise ratio (SNR), a linear equivalent system for the NLRE is known to be optimum in the minimum-mean-square-error (MMSE) sense. The MMSE covariance matrix for the linear system is given by the solution to a matrix Riccati equation, which also applies to the NLRE under the narrowband assumption mentioned above. Hence, the state-variable approach becomes optimum at high SNR and is sub-optimum in the vicinity of threshold. Since the state-variable approach is intended to handle large systems of random processes, the matrix formulation makes specification of system structure and evaluation of system performance straightforward.

Two examples of the application of this method to phase-modulated carrier signals are given.

2. The Nonlinear-Recursive Estimator

The $N \times 1$ vector random process $X(t)$ is assumed to be a vector Markov process (Ref. 1, p. 8) generated by the output of a linear system driven by an $R \times 1$ white-noise vector $U(t)$. The stochastic system can be represented by

$$\dot{X}(t) = A(t)X(t) + B(t)U(t) \tag{3}$$

where $A(t)$ and $B(t)$ are $N \times N$ and $N \times R$ matrices, respectively. The observable $y(t)$ is the sum of the signal $g(t)$ in Eq. (2) corrupted by additive noise $v(t)$ so that

$$y(t) = (2P_s)^{1/2} \sin[\omega_0 t + H(t)X(t)] + v(t) \tag{4}$$

The properties of the noise are given by the equations

$$E[U(t)] = 0 \tag{5a}$$

$$E[v(t)] = 0 \tag{5b}$$

$$E[U(t_1)U^T(t_2)] = K\delta(t_2 - t_1) \tag{5c}$$

$$E[v(t_1)v(t_2)] = L\delta(t_2 - t_1) \tag{5d}$$

$$E[U(t_1)v(t_2)] \equiv 0 \tag{5e}$$

where K and L are appropriately dimensioned matrices, where $\delta(\cdot)$ is the Dirac delta function, and where the superscript T indicates the matrix transpose.

The estimate $\hat{X}(t)$ of $X(t)$ is determined by a NLRE which has the form (Ref. 2, p. 180)

$$\left.\begin{aligned} \dot{\hat{X}}(t) &= A(t)\hat{X}(t) + P(t)C^T(t)L^{-1}\{y(t) - g[\hat{X}(t), t]\} \\ \hat{X}(0) &= E[X(0)] \end{aligned}\right\} \tag{6}$$

where

$$\left.\begin{aligned} \dot{P}(t) &= A(t)P(t) + P(t)A^T(t) - P(t)C^T(t)L^{-1}C(t)P(t) \\ &\quad + B(t)KB^T(t) \\ P(0) &= E\{[X(0) - X(0)][\hat{X}(0) - \hat{X}(0)]^T\} \end{aligned}\right\} \tag{7}$$

and where

$$C(t) = \left[\frac{\partial g[X(t), t]}{\partial x_i}\right]_{X=\hat{X}} \tag{8}$$

A block diagram representation of this NLRE appears in Fig. 1.

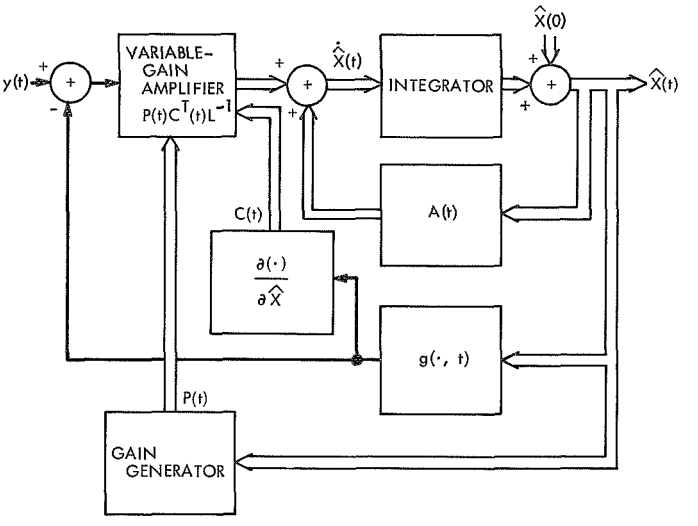


Fig. 1. Nonlinear, recursive estimator for a scalar observation

The NLRE is the result of linearizing the signal about the estimate and applying the Kalman–Bucy filtering algorithm (Ref. 2, Chap. 7). Other approximations are possible, but in practice they probably will perform similarly (Ref. 3). In this regard, further theoretical work is necessary to predict the performance of a NLRE, especially in the threshold region.

For sinusoidal signals, a narrowband assumption is often invoked (Ref. 1, p. 54), which in this case gives significant results. For $g(t)$ given by Eq. (2) then $C(t)$ in Eq. (8) is given by

$$C(t) = (2P_s)^{1/2} H(t) \cos[\omega_0 t + H(t)\hat{X}(t)] \tag{9}$$

The use of this equation in Eqs. (6) and (7) yields terms at a frequency of $2\omega_0$ which may be neglected since the remainder of the estimator is low-pass (Ref. 1, p. 54). Hence, with the narrowband assumption the NLRE has the form

$$\left.\begin{aligned} \dot{\hat{X}} &= A\hat{X} + \frac{(2P_s)^{1/2}}{L} P H^T y(t) \cos(\omega_0 t + H\hat{X}) \\ \hat{X}(0) &= E[X(0)] \end{aligned}\right\} \tag{10}$$

where

$$\left.\begin{aligned} \dot{P} &= AP + PA^T - \frac{P_s}{L} PH^T HP + BKB^T \\ P(0) &= E\{[X(0) - \hat{X}(0)][X(0) - \hat{X}(0)]^T\} \end{aligned}\right\} \tag{11}$$

The time argument has been dropped except where necessary to preserve clarity.

The simplification resulting from the narrowband assumption is evident in Fig. 2. The differencing operation $y(t) - g[\hat{X}(t), t]$ is eliminated and the system that generates the P matrix is now decoupled from the estimate $\hat{X}(t)$, which enables one to generate the P matrix before any observations are made. The validity of this approximation has been verified by simulation results (Ref. 1, p. 95). This decoupling occurs only for the case of angle modulation, which is, however, of primary interest for space communication systems. As shown in *Subsection 3*, the P matrix now represents the true error covariance matrix at large values of the SNR.

3. Linearized Model

To find a linear equivalent system for a NLRE, consider the behavior of the system without any input noise so that

$$y(t) = g(t) = (2P_s)^{1/2} \sin(\omega_0 t + HX)$$

For this case, the last term in Eq. (10) becomes

$$\frac{PC^T}{L} y(t) = \frac{2P_s}{L} PH^T \cos(\omega_0 t + H\hat{X}) \sin(\omega_0 t + HX)$$

or

$$\frac{PC^T}{L} y(t) = \frac{P_s}{L} PH^T \times \{\sin[2\omega_0 t + H(X + \hat{X})] + \sin[H(X - \hat{X})]\}$$

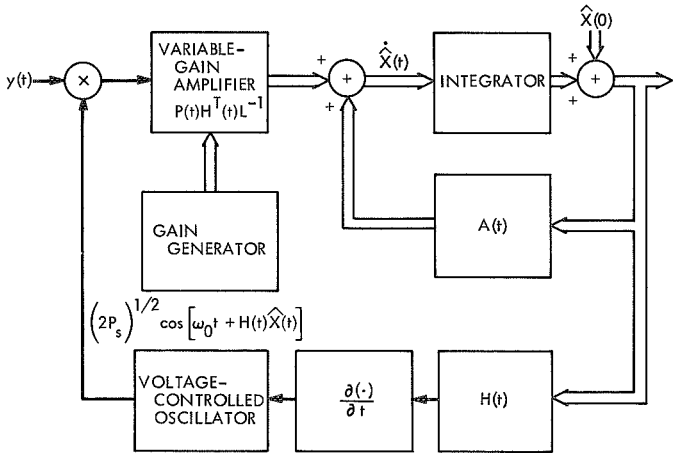


Fig. 2. NLRE using narrowband assumption for sinusoidal signals

The double-frequency term can be neglected because of the narrowband assumption used above. If the estimation error is small, then a linearizing assumption similar to that used in phase-locked loop theory (Ref. 6, p. 18) is that

$$\frac{PC^T}{L} y(t) \approx \frac{P_s}{L} PH^T H(X - \hat{X}) \tag{12}$$

Substitution of this expression into Eq. (10) leads to the baseband linear system shown in Fig. 3. Since the system is linear, one can consider the response of the system to noise separately. Hence, the total input to the linear system is then

$$y(t) = H(t) X(t) + v'(t) \tag{13}$$

where the double-sided spectral density of $v'(t)$ is L/P_s W/Hz. The MMSE estimator (Kalman–Bucy filter) for the observation in Eq. (13) and the stochastic system of Eq. (3) are (Ref. 2, p. 119)

$$\dot{\hat{X}}(t) = \left(A - \frac{P_s}{L} P H^T H\right) \hat{X}(t) + \frac{P_s}{L} P H^T y(t) \tag{14a}$$

$$\hat{X}(0) = E[X(0)] \tag{14b}$$

The NLRE of Eq. (10) becomes the Kalman–Bucy filter of Eq. (14) at high values of SNR.

The error covariance matrix for the above Kalman–Bucy filter is given by the solution to Eq. (11) (Ref. 2, p. 119) which applies to the NLRE of *Subsection 2*. Hence, for large values of the SNR where the linear model is valid, the NLRE becomes a MMSE estimator. For the particular case of rational spectra for $X(t)$, i.e., $A(t)$ is a stability matrix, the steady-state solution of Eq. (11) equals the error covariance of the Wiener optimum linear filter (Ref. 1, p. 42). A steady-state solution of

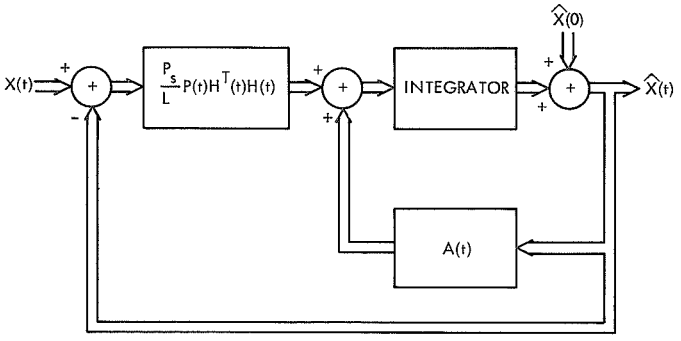


Fig. 3. Equivalent linear baseband estimator

Eq. (11) exists and is unique if A, H, B, and K are continuous functions of time and if A and H form a completely observable system (Ref. 4, p. 62).

The advantage of the state-variable approach is that the specification of the structure and performance of the estimator are easily formulated in matrix notation. The matrix Riccati equation for P in Eq. (11) is especially suited to numerical solution by a digital computer. Generalizations to systems of greater dimensionality require no new conceptual developments. The theory gives exact results for large SNRs and approximate results at low SNRs. Performance evaluation and error bounds in the vicinity of threshold require computer simulation of the algorithm and/or further analytical work.

4. Applications

a. Scalar phase modulation. Consider the stochastic system of Eq. (3) when $A(t) = -a$ and $B(t) = b$ so that $X(t) = x$, a scalar.

$$\dot{x}(t) = -a\,x(t) + b\,u(t) \tag{15}$$

The noise driver $u(t)$ has zero mean with $E[u(t)\,u(t')] = \delta(t - t')$. If $x(t)$ is stationary, then it has a first-order Butterworth spectrum $S_x(\omega)$ with autocorrelation function $R_x(\tau)$, which are given by

$$S_x(\omega) = \frac{b^2}{\omega^2 + a^2}$$

$$R_x(\tau) = \frac{b^2}{2a}\,e^{-a|\tau|}$$

Let the observation be

$$y(t) = g(t) + v(t)$$

or

$$y(t) = (2P_s)^{1/2} \sin[\omega_0 t + x(t)] + v(t) \tag{16}$$

so that the problem is to estimate the phase modulation $x(t)$, assuming that the carrier frequency is known. The noise $v(t)$ has the properties of Eqs. (5b), (5d), and (5e).

Using the narrowband assumption $a << \omega_0$, one obtains the NLRE from Eq. (10),

$$\left.\begin{aligned} \dot{\hat{x}}(t) &= -a\hat{x}(t) + \frac{(2P_s)^{1/2}}{L}\,p(t)\,y(t)\cos[\omega_0 t + \hat{x}(t)] \\ \hat{x}(0) &= 0 \end{aligned}\right\} \tag{17}$$

with $p(t)$ from Eq. (11),

$$\left.\begin{aligned} \dot{p}(t) &= -2ap(t) - \frac{P_s}{L}\,p^2(t) + b^2 \\ p(0) &= \frac{b^2}{2a} \end{aligned}\right\} \tag{18}$$

This system can be implemented as shown in Fig. 4a.

For the scalar case, one can solve Eq. (18) to yield (Ref. 5, p. 64)

$$p(t) = \frac{p(0)}{1 + \dfrac{P_s p(0)}{aL}\,\dfrac{\tanh \gamma a t}{\tanh \gamma a t + \gamma}} \tag{19}$$

where

$$\gamma = \left(\frac{1 + P_s b^2}{a^2 L}\right)^{1/2}$$

For comparison with the Wiener optimum linear filter, the steady-state value $\tilde{p}$ of $p(t)$ is needed. As $t \to \infty$,

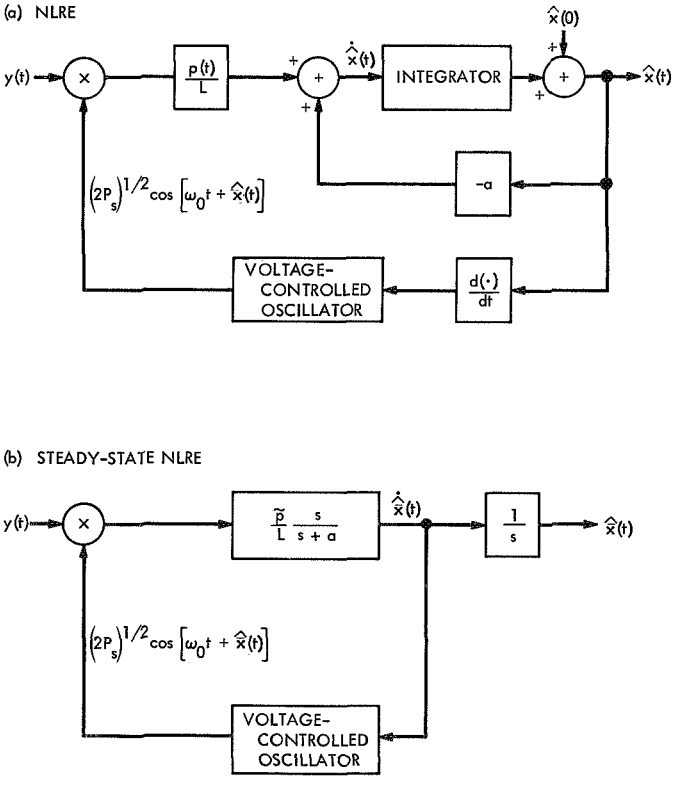


Fig. 4. Systems for scalar phase modulation

Eq. (19) becomes $\tilde{p}$ given by

$$\tilde{p} = \lim_{t\to\infty} p(t) = \frac{aL}{P_s}(\gamma - 1) \tag{20}$$

The steady-state NLRE is given by Eq. (17) with $\tilde{p}$ substituted for $p(t)$. The system can be represented by the block diagram in Fig. 4b, which is a first-order phase-locked loop. This is the Wiener optimum filter in the linear region for tracking phase modulation with no loop detuning (Ref. 6, p. 137). The phase error variance $\sigma^2_{x-\hat{x}}$ can be found by using the Yovits–Jackson formula (Ref. 6, p. 143) and is given by $\tilde{p}$ in Eq. (20).

b. Subcarrier demodulation. An application which demonstrates the extension of the state-variable approach to more complicated systems is given by the addition to the previous example of a subcarrier data channel of known synchronization. Assume the observation to have the form

$$y(t) = (2P_s)^{1/2} \sin[\omega_0 t + x_1(t) + x_2(t) \sin \omega_s t] + v(t) \tag{21}$$

where ω_s is the subcarrier frequency. For simplicity, assume that the statistics of $x_1(t)$ and $x_2(t)$ are given by Eq. (3) with

$$A = \begin{bmatrix} -a_1 & 0 \\ 0 & -a_2 \end{bmatrix}$$

and

$$B = \begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix}$$

The NLRE for this example is given by Eq. (10) with

$$H = [1 \quad \sin \omega_s t] \tag{22}$$

and with the P matrix given by the solution to the set of equations

$$\dot{p}_{11} = -2a_1 p_{11} - \frac{P_s}{L} \times (p_{11}^2 + 2p_{11}p_{21} \sin \omega_s t + p_{21}^2 \sin^2 \omega_s t) + b_1^2 \tag{23a}$$

$$\dot{p}_{12} = -(a_1 + a_2) p_{12} - \frac{P_s}{L} \times [p_{11}p_{12} + (p_{11}p_{22} + p_{12}^2) \sin \omega_s t + p_{12}p_{22} \sin^2 \omega_s t] \tag{23b}$$

$$p_{21} = p_{12} \tag{23c}$$

$$\dot{p}_{22} = -2a_2 p_{22} - \frac{P_s}{L} \times (p_{12}^2 + 2p_{12}p_{22} \sin \omega_s t + p_{22}^2 \sin^2 \omega_s t) + b_2^2 \tag{23d}$$

In order to gain insight into the operation of the NLRE, assume that the crosscorrelation term p_{12} in Eq. (23) is zero. Then, from Eq. (10) with H given by Eq. (22), one obtains

$$\begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \end{bmatrix} = \begin{bmatrix} -a_1 \hat{x}_1 \\ -a_2 \hat{x}_2 \end{bmatrix} + \frac{(2P_s)^{1/2}}{L} \begin{bmatrix} p_{11} \\ p_{22} \sin \omega_s t \end{bmatrix} y(t) \cos(\omega_0 t + H\hat{X}) \tag{24}$$

The steady-state solution of Eq. (24) can be implemented as shown in Fig. 5.

By neglecting the crosscorrelation between $x_1 - \hat{x}_1$ and $x_2 - \hat{x}_2$, the steady-state NLRE in Fig. 5 is equivalent to

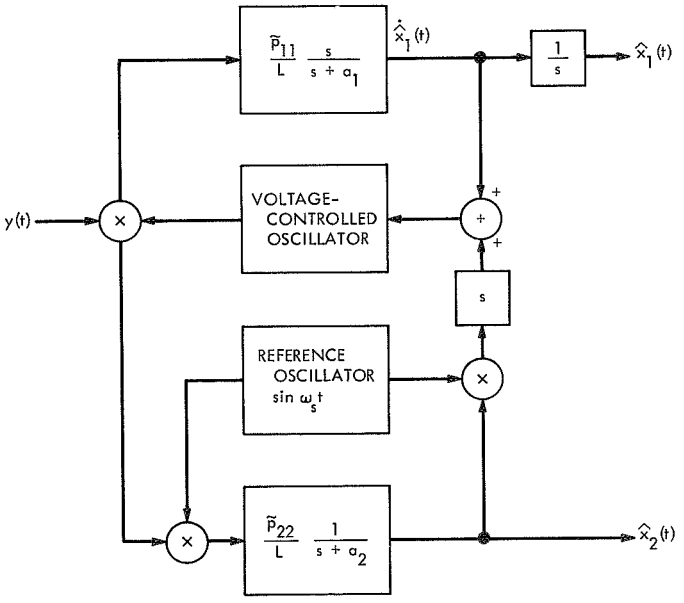


Fig. 5. Steady-state NLRE for subcarrier modulation

the linear MMSE filter obtained by considering x_1 and x_2 as separate modulation processes (Ref. 6, p. 131). The resulting error variances can be found by using the Yovits–Jackson formula (Ref. 6, p. 143) or alternatively by finding the steady-state solution to Eq. (23) with $p_{12} \equiv 0$. Following the latter procedure, one finds the steady-state values $\tilde{p}_{11}$ and $\tilde{p}_{22}$,

$$\tilde{p}_{11} = \frac{\gamma_1 - 1}{2R_1} \tag{25}$$

$$\tilde{p}_{22} = \frac{\gamma_2 - 1}{R_2} \tag{26}$$

where

$$R_i = \frac{P_s}{2La_i}, \qquad i = 1, 2 \tag{27}$$

$$\gamma_1 = [1 + 4R_1 p_{11}(0)]^{1/2} \tag{28}$$

$$\gamma_2 = [1 + 2R_2 p_{22}(0)]^{1/2} \tag{29}$$

where

$$p_{ii}(0) = \frac{b_i^2}{2a_i}, \qquad i = 1, 2 \tag{30}$$

Equations (25) and (26) are plotted in Figs. 6 and 7, respectively; these figures show the variances of each component of phase modulation as a function of the SNR for that component. Following Snyder (Ref. 1, p. 83), a threshold is shown at a level of $\tilde{p}_{ii} = 0.25$ above which the curves give an accurate description of the receiver performance when $p_{12} \equiv 0$.

One would expect that the NLRE would perform better when Eq. (23) is solved without assuming $p_{12} \equiv 0$. However, a numerical solution of this equation has shown that, for this example, $|p_{12}| << p_{11}, p_{22}$ for the range of parameter variations within the narrowband assumption above. Hence, the values of $\tilde{p}_{11}$ and $\tilde{p}_{22}$ given by Eqs. (25) and (26) are applicable in general, and the performance of the NLRE in estimating one state-variable component is essentially independent of the modulation by the other component. For example, the value of $\tilde{p}_{11}$ is invariant with respect to changes of R_2, a_2, or $p_{22}(0)$.

5. Conclusion

The state-variable approach to phase demodulation is particularly suited to the solution of many problems involving a large number of modulation sidebands. In most cases of interest, the method probably does not offer greatly improved performance, but each case should be evaluated against alternative receiver designs. The method does include second-order statistics in the receiver design, and at large values of SNR it does specify the MMSE receiver both for acquisition and steady-state operation. State-variable design techniques can be considered as a means of obtaining realizable system configurations and error variance information. The results presented here provide an evaluation of performance in the linear region of operation, and performance near threshold can be obtained by computer simulation of the

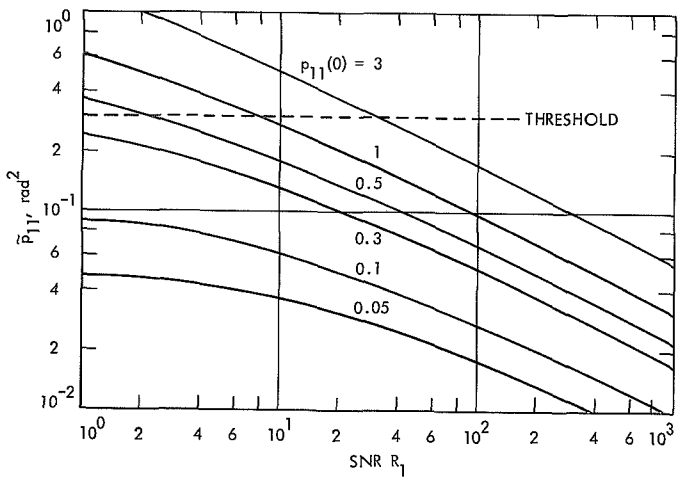


Fig. 6. Variance of $x_1 - \hat{x}_1$ as a function of SNR in bandwidth of $S_{x_1}(\omega)$

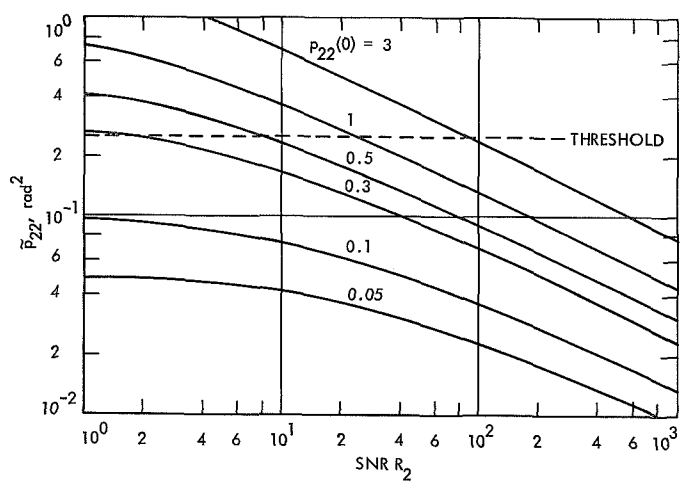


Fig. 7. Variance of $x_2 - \hat{x}_2$ as a function of SNR in bandwidth of $S_{x_2}(\omega)$

NLRE. State-variable analysis gives straightforward results without invoking numerous simplifying assumptions; this advantage becomes more important as the complexity of the system increases.

References

1. Snyder, D. L., *The State-Variable Approach to Continuous Estimation.* The MIT Press, Cambridge, Mass., 1969.

2. Nahi, N. E., *Estimation Theory and Applications.* John Wiley & Sons, Inc., New York, 1969.

3. Schwartz, L., and Stear, E. B., "A Computational Comparison of Several Nonlinear Filters," *IEEE Trans. Auto. Contr.*, Vol. AC-13, No. 1, pp. 83–86, Feb. 1968.

4. Bucy, R. S., and Joseph, P. D., *Filtering for Stochastic Processes with Applications to Guidance.* Interscience Publishers, Inc., New York, 1968.

5. Brand, L., *Differential and Difference Equations.* John Wiley & Sons, Inc., New York, 1966.

6. Viterbi, A. J., *Principles of Coherent Communication.* McGraw-Hill Book Co., Inc., New York, 1966.

VI. Communications Elements Research

ADVANCED ENGINEERING

A. Low Noise Receivers: Microwave Maser Development, *E. R. Wiebe*

1. Introduction

An instrument for automatically recording the reserve heat capacity of traveling-wave maser (TWM)/closed cycle refrigerators (CCR) has been built and tested at JPL. Using this device, the overall health of CCR systems in the field can be easily monitored between tracking periods. The measurement of reserve heat capacity is performed with higher accuracy and greater speed than possible with the manual methods generally in use.

2. Instrument Description

The instrument consists of a simple servo system which applies sufficient dc power to the liquid helium stage to maintain its temperature at about 0.1°K above normal. This power level is then displayed as reserve heat capacity on either a meter or strip chart recorder.

Figure 1 shows a schematic diagram of the unit. It can be divided into three parts:

(1) The power supply.

(2) The control unit.

(3) The indicator unit.

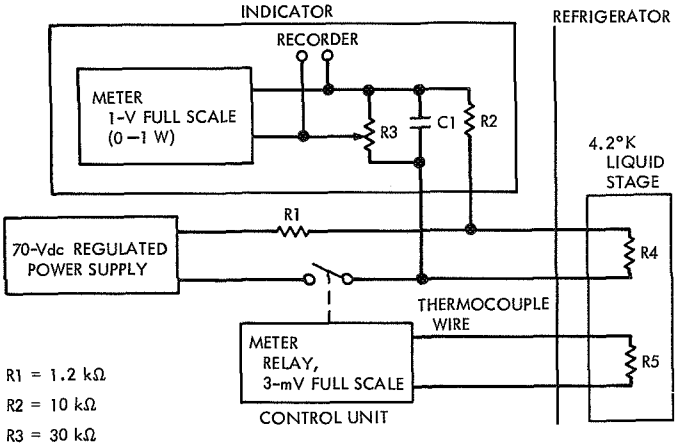


Fig. 1. Instrument for recording reserve heat capacity of TWM/CCR system

The power supply is adjusted to provide 1 W of power to heater resistor R4 on the liquid helium stage. R1 minimizes inaccuracies due to different values of heater resistance from one refrigerator to another. The control unit is a meter relay connected in series with a thermocouple and temperature-sensitive resistor R5. The meter relay is adjusted to switch the power supply on-and-off at a rate which maintains the liquid helium stage temperature at about 0.1°K above normal. The indicator circuit is a meter with a 0- –1-W linear scale connected to a voltage divider R3 and a resistance/capacitance (R/C) integrator R2 and C1. R3 is adjusted for a full-scale meter reading when a continuous 1 W is applied to the CCR heater R4. The R/C integrator provides a voltage to the meter which is proportional to the percentage of time the heater is switched on, therefore, giving a direct reading of dc watts reserve heat capacity.

3. Instrument Operation

Figure 2 shows a typical recording of reserve heat capacity from a CCR used to cool a TWM. At $T = 0$ h, after the refrigerator has cooled to operating temperature, the instrument described in *Subsection 2* was energized. Supply pressure (SP) and Joule-Thompson (J-T) pressure were both 300 psi. At these pressure conditions, the refrigerator had a J-T flow of 1.18 standard ft³/min (scfm) and a reserve heat capacity of about 880 mW.

To demonstrate how the instrument works, the above pressures were reduced to 250 and 225 psi, respectively. At $T = 2$ h, the capacity is seen to drop to around 600 mW. At about $T = 4$ h, the maser pump klystron was turned on and capacity was dropped to about 540 mW, indicating about 60 mW of pump energy directly absorbed by the maser.

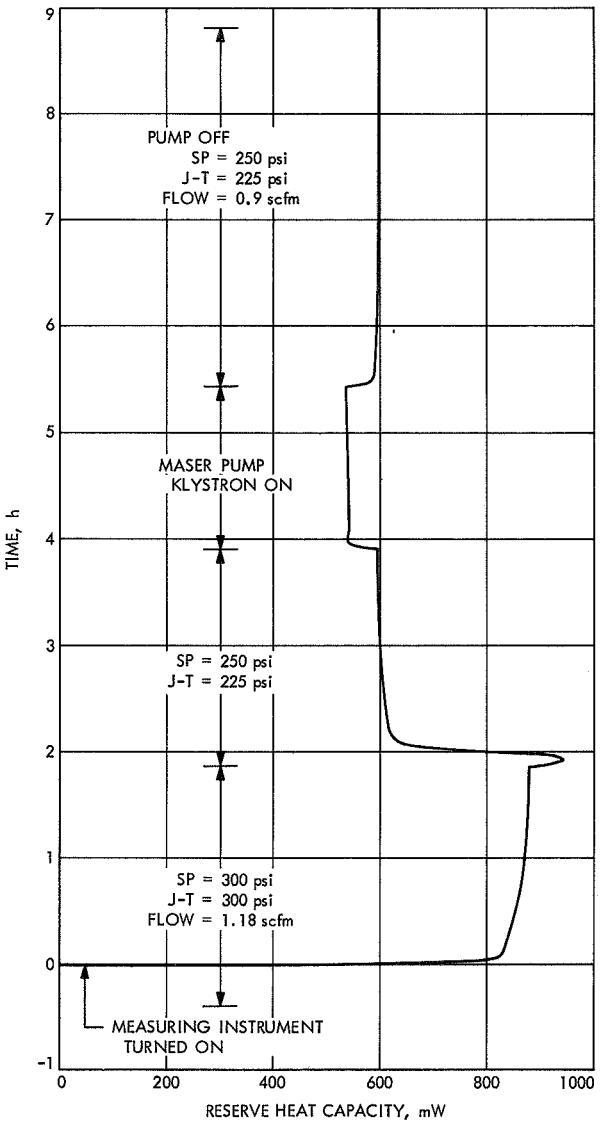


Fig. 2. Recording of reserve heat capacity of TWM/CCR system

B. Improved Calibration Techniques: Realizability Conditions on Reflection Coefficients of Unsymmetrical, Passive, Reciprocal 2-Port Networks, *T. Y. Otoshi*

1. Introduction

It is generally known that if a 2-port reciprocal microwave network is lossless, then

$$|S_{11}| = |S_{22}|$$

where $|S_{11}|$ and $|S_{22}|$ are the magnitudes of the voltage reflection coefficients as seen looking into the input and output ports, respectively, when the opposite port is terminated in a non-reflecting load.

In practice, however, microwave networks are not lossless. For a practical network that is lossy and reciprocal (but not symmetrical), it is often desired that an estimate of $|S_{22}|$ be obtained from knowledge of $|S_{11}|$. Heretofore, a simple formula for obtaining this estimate has not been widely known or available in published form. Recently, the necessary formulas have been derived by Dr. Glenn F. Engen of the National Bureau of Standards, and it is the purpose of this article to discuss these formulas and point out their potential usefulness to microwave engineers.

2. Formulas for Limits of $|S_{22}|$

During a technical discussion with Dr. Glenn F. Engen, it was discovered that bounds on $|S_{22}|$ in terms of $|S_{11}|$ could be derived from the following realizability relationship, which appeared in a previous paper by Engen (Ref. 1).

$$1 - \frac{|S_{12}|^2}{1 - |S_{11}|^2} \geqq \left| S_{22} + \frac{S_{11}^* S_{12}^2}{1 - |S_{11}|^2} \right| \tag{1}$$

where S_{11}, S_{12}, and S_{22} are scattering coefficients and (*) denotes complex conjugate.

If reciprocity is assumed, i.e., $S_{21} = S_{12}$, then from Eq. (1), the following limits of $|S_{22}|$ can be derived in terms of $|S_{11}|$.

$$|S_{22}|_{\max} = \frac{1}{L}|S_{11}| + \left(1 - \frac{1}{L}\right) \tag{2}$$

If $|S_{11}| > L - 1$, then

$$|S_{22}|_{\min} = \frac{1}{L}|S_{11}| - \left(1 - \frac{1}{L}\right) \tag{3a}$$

If $|S_{11}| \leqq L - 1$, then by additional physical considerations

$$|S_{22}|_{\min} = 0 \tag{3b}$$

where

$$L = \frac{1 - |S_{11}|^2}{|S_{21}|^2} \tag{4}$$

The symbol L is a power ratio $\geqq 1$ and is often referred to as the dissipative loss factor when the direction of energy flow is from port 1 to port 2. Letting $L_{\text{dB}} = 10 \log_{10} L$, one can use the following low-loss case approximations:

$$\left.\begin{aligned} \frac{1}{L} &\approx 1 - 0.23\, L_{\text{dB}} \\ L - 1 &\approx 0.23\, L_{\text{dB}} \\ 1 - \frac{1}{L} &\approx 0.23\, L_{\text{dB}} \end{aligned}\right\} \quad \text{for } L_{\text{dB}} \leqq 0.1 \text{ dB} \tag{5}$$

In cases where the attenuation (or the matched system insertion loss) of the 2-port is known instead of L_{dB}, the dissipative loss factor can be calculated from

$$L = (1 - |S_{11}|^2)\, 10^{A_{\text{dB}}/10} \tag{6}$$

where A_{dB} is the network attenuation in decibels.

3. Applications

Three practical situations where the $|S_{22}|$ limit formulas can be useful are described below.

(1) In the microwave calibrations of a reciprocal 2-port network where the waveguide types on each port are different, it may not be economical to assemble reflectometer systems for each waveguide type. If the loss of the 2-port device is known to be small, then from Eqs. (2) and (3), it might be found that the difference between $|S_{22}|$ and $|S_{11}|$ will be negligibly small. If such is the case, it is only required that reflection coefficient measurement be made at one of the ports.

(2) It is sometimes the case where reflection coefficients have been measured at both ports of a reciprocal 2-port network, but due to a possible measurement error, the data obtained for one of the ports appears to be questionable. Since the measured values must be bounded by the physically realizable limits, the limit formulas can be used to check the validity of the measurement data.

(3) In conjunction with the design of a new microwave system, a theoretical study is often made of the effect of such sources of errors as mismatches which can cause significant deviations from predicted system performance. In such studies, it becomes necessary to study hypothetical cases and synthesize worst-case system parameter values. Equations (2) and (3) can be used to insure that physically realizable values of voltage reflection coefficients are used in the hypothetical test cases.

4. Sample Cases

The following are a few sample cases which demonstrate the spreading that occurs between maximum and minimum values of $|S_{22}|$ as the lossiness of the 2-port is increased. Comparisons are made with actual measured values of $|S_{22}|$.

(*1*) *Lossless case.* In the lossless case, the 2-port has no dissipative losses so that

$$L = 1$$

and the limits from Eqs. (2) and (3) become

$$|S_{22}|_{max} = |S_{22}|_{min} = |S_{11}|$$

(2) *Sample low-loss case.* Given that

$$L_{dB} = 0.0016\,dB$$

$$|S_{11}| = 0.00798$$

then from use of Eqs. (2), (3), and (5),

$$|S_{22}|_{max} = 0.00835$$

$$|S_{22}|_{min} = 0.00761$$

The input data of this case was obtained from measurements made on a WC504/WR 430 copper broadband transition. The actual measured value of $|S_{22}|$ was 0.00847 $\pm$0.00023 pe_T (Ref. 2).

(3) *Sample medium-loss case.* Given that

$$A_{dB} = 0.172\,dB$$

$$|S_{11}| = 0.0004$$

then from Eqs. (2), (3), and (6),

$$|S_{22}|_{max} = 0.0392$$

$$|S_{22}|_{min} = 0$$

The input data of this case was obtained from measurements made on a cryogenic load transmission line of an *X*-band Cassegrain cone system. The actual measured value of $|S_{22}|$ was 0.021 (SPS 37-43, Vol. III, p. 59).

(*4*) *Sample high-loss case.* Given that

$$A_{dB} = 1.17\,dB$$

$$|S_{11}| = 0.0075$$

then from Eqs. (2), (3), and (6),

$$|S_{22}|_{max} = 0.242$$

$$|S_{22}|_{min} = 0$$

The input data of this case was obtained from measurements made on a compact *H*-band rotary vane attenuator whose vane was set at 20 deg. The actual measured value of $|S_{22}|$ at this setting was 0.015 (SPS 37-62, Vol. II, pp. 81–87).

5. Conclusions

It has been shown that when the loss of a 2-port network is known and the reflection coefficient is measured at only one of the ports, the limits of the magnitude of reflection coefficient at the other port can be easily calculated. It is important to remember that the limit formulas presented in this article are restricted to passive 2-port networks that are reciprocal (non-ferrous).

References

1. Engen, G. F., "An Evaluation of the Back-to-Back Method of Measuring Adaptor Efficiency," *IEEE Trans. Instr. Meas.*, Vol. IM-19, No. 1, pp. 18–22, Eq. (10), Feb. 1970.

2. Levy, G. S., et al., "The Ultra-Cone: An Ultra-Low Noise Space Communication Ground Radio Frequency System," *IEEE Trans. Micro. Theory Tech.*, Vol. MTT-16, No. 9, pp. 596–602, Sep. 1968.

C. A Noise-Adding Radiometer for Use in the DSN,

P. D. Batelaan, R. M. Goldstein, and C. T. Stelzried

1. Introduction

The communication requirements for deep space missions necessitate the use of high-gain antennas and sensitive receivers on the ground. Because of their unusually good sensitivity, these receiving systems are also well suited to radio astronomy applications. However, most gain compensation techniques normally employed for radio astronomy, such as the Dicke scheme, would degrade the communication capability of a DSN receiver. A radiometer which is compatible with the DSN communication requirement is described here.

2. Discussion

The radiometer scheme chosen is based on a technique due to Ohm and Snell (Ref. 1) except that a digital computer is used to handle the output data rather than the analog device of Ohm and Snell. A block diagram of the noise-adding radiometer (NAR) is shown in Fig. 1. Figure 2 depicts the portable NAR rack and its contents.

A solid-state noise diode (noise temperature ~400,000 K) is driven by a constant current source. The input current to the diode is square-wave-shorted by computer control at a low-frequency keying rate (11 Hz, typically). The resulting square-wave noise is coupled (−40 dB, typically) into the receiver input. This is a noise temperature T_N increase of ~ 40 K added into the receiver input at an 11-Hz square-wave rate. This is depicted in Fig. 3. The noise temperature of the receiver is multiplied by the overall gain G and it is the instability of this gain term which is eliminated to achieve increased radiometer sensitivity.

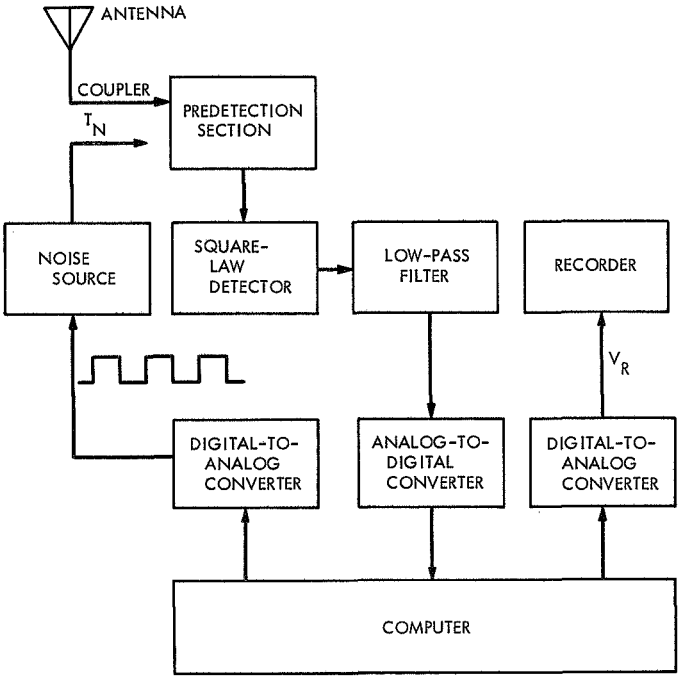


Fig. 1. Noise-adding radiometer

Salient to the operation of this NAR is a good, wideband square-law detector. The computer periodically (i.e., typically 7 kHz) samples the output of the square-law detector and forms a voltage ratio (Fig. 3)

$$V_R = \frac{v_1}{v_2 - v_1} = \frac{G(T_{op} + T_S)}{GT_N} \tag{1}$$

where

v_1 = averaged voltage at the square-law detector output during noise diode off cycle

v_2 = averaged voltage at the square-law detector output during noise diode on cycle

T_{op} = system operating noise temperature

T_S = radio source temperature

Rearranging, we have

$$\frac{T_S}{T_N} = V_R - \frac{T_{op}}{T_N} \tag{2}$$

which is gain independent. An initial zeroing operation is made with the antenna off source. A term $(T_{op}/T_N)_0$, the value of V_R when $T_S/T_N = 0$, is stored in the computer and subtracted from subsequent outputs (i.e., typically every 0.6 s). The computer output is then given by

$$V_R = \frac{T_S}{T_N} + \left[\frac{T_{op}}{T_N} - \left(\frac{T_{op}}{T_N}\right)_0\right] \tag{3}$$

This output is linearly related to the source temperature. The second term is zero immediately after the zeroing operations and drift is negligible during operation. Instability in the noise diode is reflected as corresponding instability in the radiometer.

Noise calibration of the system is accomplished by injecting a separate noise-source signal into the receiver input. The value of the calibration source is 1–4 K (typical) referenced to the antenna.

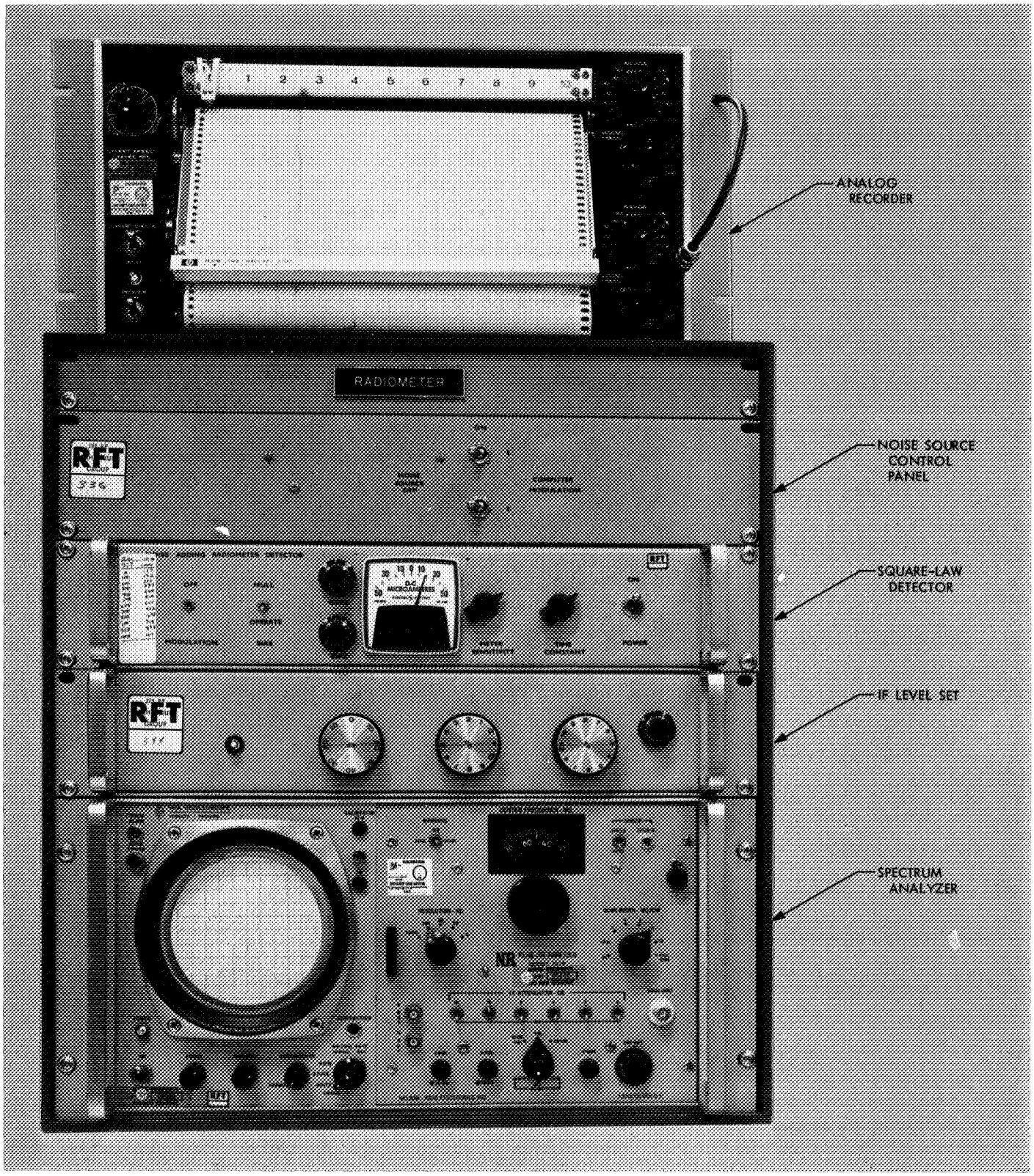


Fig. 2. Noise-adding radiometer rack

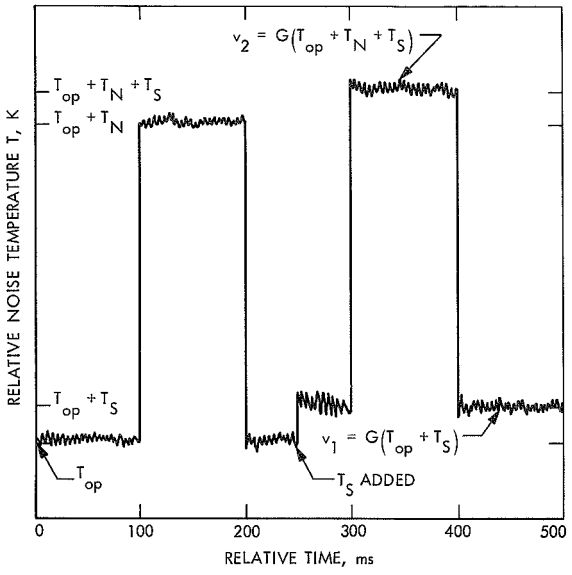


Fig. 3. Representation of the signal levels from the square-law detector

A comparison of the NAR with a total power radiometer (TPR) will serve to clarify their relative characteristics. The TPR sensitivity is given by (Ref. 2)

$$(\Delta T_{\min})_{\mathrm{TPR}} = T_{op}\left[\left(\frac{1}{\Delta\nu_{HF}\tau_{LF}}\right) + \left(\frac{\Delta G}{G}\right)^2\right]^{1/2} \quad (4)$$

where

$\Delta\nu_{HF}$ = detector bandwidth, Hz

τ_{LF} = equivalent integration time of the output integrator

G = system gain

and

ΔG = change (instability) in G

Without gain instability (the ideal radiometer),

$$\Delta T_{\min} = T_{op}\left(\frac{1}{\Delta\nu_{HF}\tau_{LF}}\right)^{1/2} \quad (5)$$

The NAR sensitivity is (*Subsection 4*)

$$(\Delta T_{\min})_{\mathrm{NAR}} = T_{op}\,2\left(\frac{1}{\Delta\nu_{HF}\tau_{LF}}\right)^{1/2}\left(1 + \frac{T_{op}}{T_N}\right) \quad (6)$$

Two observations can be made: (1) if the gain is constant for the TPR and injected noise is high for the NAR, then

$$(\Delta T_{\min})_{\mathrm{NAR}} = 2\,(\Delta T_{\min})_{\mathrm{TPR}} \quad (7)$$

and (2) the NAR exhibits (theoretically) no sensitivity to radiometer gain instability.

The NAR is not yet fully operational but some routine star tracks have been executed in the course of system development. Results have been excellent. Figure 4 shows a simultaneous TPR and NAR recording of source 3C10.

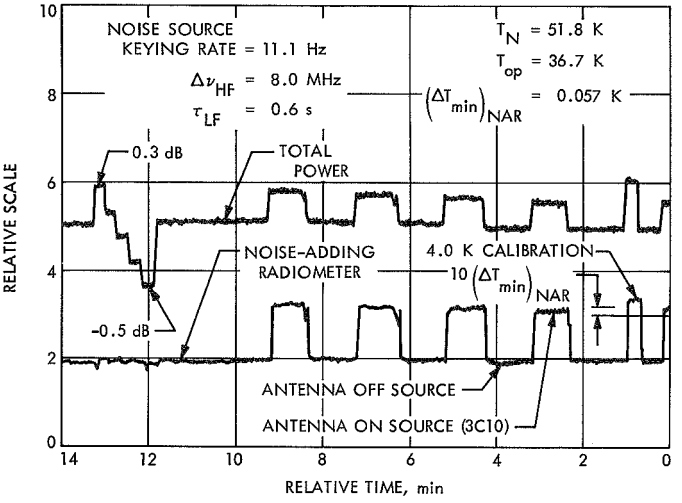


Fig. 4. Noise-adding radiometer recording (DSS 11, Feb. 11, 1970)

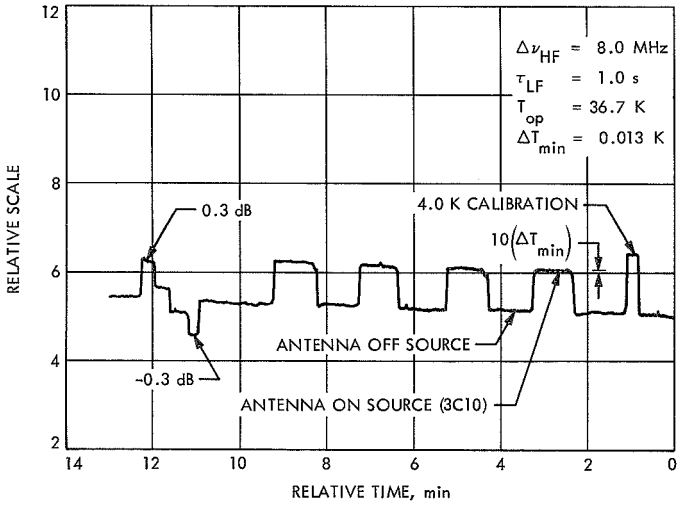


Fig. 5. Total power radiometer recording (DSS 11, Feb. 11, 1970)

While the antenna was pointing on the background sky, system gain was deliberately changed. Notice that the NAR was relatively unaffected while the TPR output was radically changed. The total power radiometer used for recording Fig. 4 has degraded sensitivity since it was operating simultaneously with the NAR and its associated noise injecting source. Figure 5 shows a recording of the normal (NAR not in operation) operation of the TPR. Again, notice the gain sensitivity due to deliberate gain changes and also due to drift.

3. Conclusion

Examination of Fig. 4 shows that the NAR was operating with a minimum sensitivity of 0.066 K, which is about 1.2 times the theoretical minimum of 0.057 K determined from Eq. (6). Similarly, Fig. 5 shows that the TPR was operating with a minimum sensitivity of 0.019 K, which is about 1.5 times its theoretical (ideal) minimum of 0.013 K determined from Eq. (5).

The conclusion to be drawn from Figs. 4 and 5 is that the total power radiometer has lower short-term jitter and is most useful for short-term measurements (seconds) and that the noise-adding radiometer has less drift and is most useful for long-term measurements (minutes).

The noise-adding technique of achieving a gain insensitive radiometer has been demonstrated to be a powerful technique which is ideally suited to application at a DSN station. Care must be taken in fabrication of a true square-law detector and the usual problem of gain linearity must be controlled throughout the system. The keyed noise-source stability is also critical.

4. Calculation of $\Delta T_{\min}$

From Fig. 3, the average values of v_1 and v_2 are seen to be

$$v_1 = G(T_{op} + T_S) \tag{8}$$

$$v_2 = G(T_{op} + T_N + T_S) \tag{9}$$

For errors Δv_1 and Δv_2, we have from Eq. (1), since $|\Delta v_1| << v_1$ and $|\Delta v_2| << v_2$,

$$\Delta V_R = \left(\frac{\partial V_R}{\partial v_1}\right)\Delta v_1 + \left(\frac{\partial V_R}{\partial v_2}\right)\Delta v_2 \tag{10}$$

Since v_1 and v_2 are independent (bar denotes average),

$$\sigma^2_{V_R} = (\Delta V_R)^2 = \left(\frac{\partial V_R}{\partial v_1}\right)^2 \overline{(\Delta v_1)^2} + \left(\frac{\partial V_R}{\partial v_2}\right)^2 \overline{(\Delta v_2)^2}$$

$$= \frac{v_2^2\,\overline{(\Delta v_1)^2} + v_1^2\,\overline{(\Delta v_2)^2}}{(v_2 - v_1)^4} \tag{11}$$

Let P be the switching cycle period. Then we have a total power radiometer during each half switching cycle and from Ref. 2, p. 245,

$$\overline{(\Delta v_i)^2} = \frac{v_i^2}{(\Delta\nu_{HF})\dfrac{P}{2}}, \qquad i = 1, 2 \tag{12}$$

Using Eqs. (11) and (12), we obtain

$$\sigma_{V_R} = \frac{2v_1v_2}{(v_2 - v_1)^2\,(\Delta\nu_{HF}P)^{1/2}} \tag{13}$$

The minimum detectable temperature $\Delta T_{\min}$ may be defined as that T_S which causes V_R to change by σ_{V_R}. From Eq. (2), the ΔV_R due to $T_S = \Delta T_{\min}$ is

$$\Delta V_R = \frac{\Delta T_{\min}}{T_N} \tag{14}$$

Upon equating Eqs. (13) and (14) and using Eqs. (8) and (9) with T_S suppressed because it is small compared to T_{op},

$$\Delta T_{\min} = \frac{2\,T_{op}\left(1 + \dfrac{T_{op}}{T_N}\right)}{(\Delta\nu_{HF}P)^{1/2}} \tag{15}$$

This was derived for one switching period. Clearly, if τ_{LF} is any integer number of switching periods, we may replace P by τ_{LF} because we simply have an average over a longer time. Thus,

$$(\Delta T_{\min})_{\mathrm{NAR}} = \frac{2\,T_{op}\left(1 + \dfrac{T_{op}}{T_N}\right)}{(\Delta\nu_{HF}\tau_{LF})^{1/2}} \tag{16}$$

References

1. Ohm, E. A., and Snell, W. W., "A Radiometer for a Space Communications Receiver," *Bell System Technical Journal*, pp. 2047–2080, Sep. 1963.

2. Kraus, J. D., *Radio Astronomy*. McGraw-Hill Book Co., Inc., New York, 1966.

VII. Supporting Research and Technology

ADVANCED ENGINEERING

A. 30-ft-diam Reflector Upgrade Study,

M. S. Katow

1. Introduction

As described previously (SPS 37-52, Vol. II, p. 86), the distortions of an az-el type of antenna resulting from the change in direction of the gravity force due to elevation angle rotation can be resolved from two loading cases. The first is the gravity loading along the symmetric axis, and the other loading is at right angle to the symmetric axis or the antisymmetric direction. The final distortion number is the vector sum of the deflections resulting from the change in the two loadings components.

For an arrangement using radial trusses, the minimum distortion of the reflector is attained if the trusses are arranged so that (1) in the symmetric loading case the deflections are symmetrical about the axis of symmetry, and the ribs deflect to a parabolic curve of longer focal length, and (2) for the antisymmetric case the rib trusses must be deep enough to minimize the deflections from the antisymmetric loading.

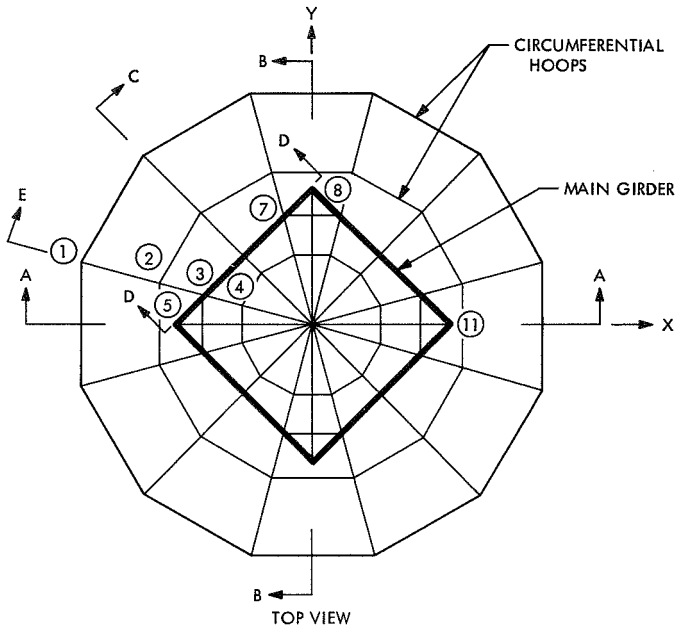
To satisfy requirements in (1) easily, without regard to other considerations (such as the effects by wind loading on the driving torques) the truly axially symmetric rigid truss arrangement, with structure to mask out the reaction effects of the unsymmetric elevation bearings, is the ideal solution.

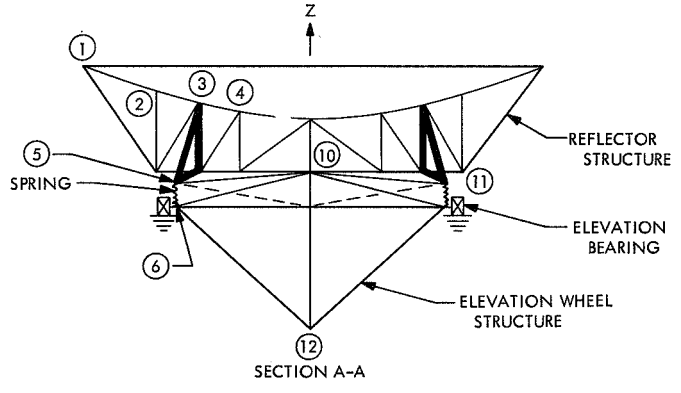
The two elevation bearings cause non-symmetric deflections, and there is a need for passive compensating devices, such as springs, to reduce the symmetric distortion without interacting and increasing the distortion in the antisymmetric loading case.

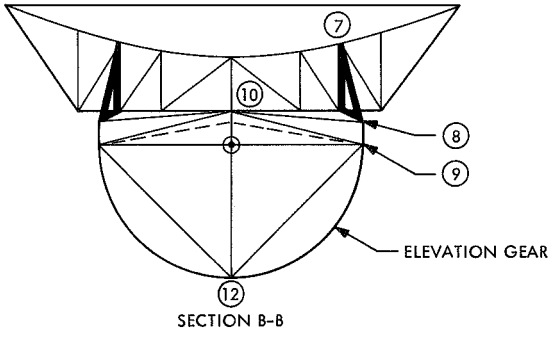
This article describes the result of adding a spring-type passive compensating device to the non-axial symmetric 30-ft reflector structure at the Venus DSS. Only the results of the analytical study using a structural computing program (NASTRAN) are presented.

2. Discussion

Figure 1 shows in schematic the arrangement of the trusses in the tipping assembly. The reflector structure







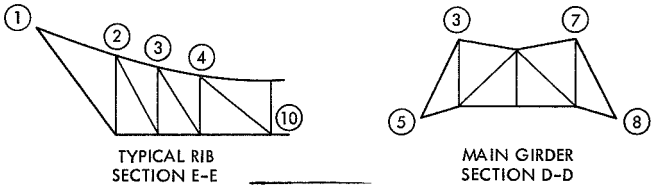


Fig. 1. Truss arrangements of tipping assembly in 30-ft-diam-antenna

consists of radial rib trusses and circumferential loop trusses supported by the main girder and the central connection to the elevation wheel structure at node 10. The main girder support points are, typically, nodes 5 and 8 of the elevation wheel structure.

For clarity purposes, the diagonals in the upper surface of the reflector structure are not shown. The quadripod is also deleted, since it is now supported from the elevation wheel structure and has little or no effect on the reflector distortion. The tipping assembly is supported on the two elevation bearings, and the pinion gear is anchored to the yoke, which is also not shown because of the unrestrained supporting arrangement.

Presently, the main girder is rigidly supported in the Z direction by solid bar members between nodes 5 and 6 plus 8 and 9 as well as by bars on the other reflector half. Because of the high stiffness of node 6 as compared to node 8, the main girder, and consequently the reflector

CONTOUR DEFINITIONS

LABEL	NORMAL ERROR, in.
A	-0.035
B	-0.030
C	-0.025
D	-0.020
E	-0.015
F	-0.010
G	-0.005
H	0.000
I	0.005
J	0.010
K	0.015
L	0.020
M	0.025
N	0.030
O	0.035

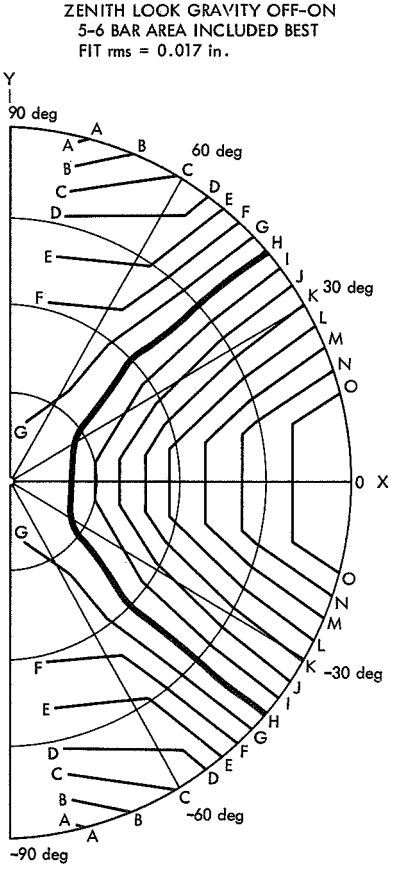


Fig. 2. Deflection contour map for zenith look for half dish, 30-ft-diam-antenna, best fit rms = 0.017 in.

structure, deflects non-uniformly about the plane of symmetry (Y-Z) for zenith-look gravity load. The resulting best-fit deflections are shown in Fig. 2, with the rms value of 0.017 in.

By replacing the 5–6 bar (17$\frac{3}{16}$ in. long) with a spring that deflects 0.026 in. for a 900-lb load, the Z deflection of node 5 can approach that of node 8 resulting in a more symmetrically shaped deflection pattern. After best-fit, the deflection pattern is shown in Fig. 3 with the resulting reduction of the rms to 0.002 in.

Although the addition of the spring improved the gravity-deflection patterns, the reflector structure became more sensitive to distortion from wind loading. Additional wind bracing connecting nodes 5, 10, 11 and a sliding connection in Z direction to node 12 may be required. This bracing will greatly increase the stiffness of the reflector structure in the X–Z plane and a sliding connection to node 12 will be necessary only if RF boresight pointing accuracy is involved.

3. Conclusions

The 30-ft reflector structure is a practical design, using rigid bars that can be improved for gravity distortion by the addition of passive compensators in a form of springs. However, additional bracing for minimizing distortions due to wind loadings may be desirable.

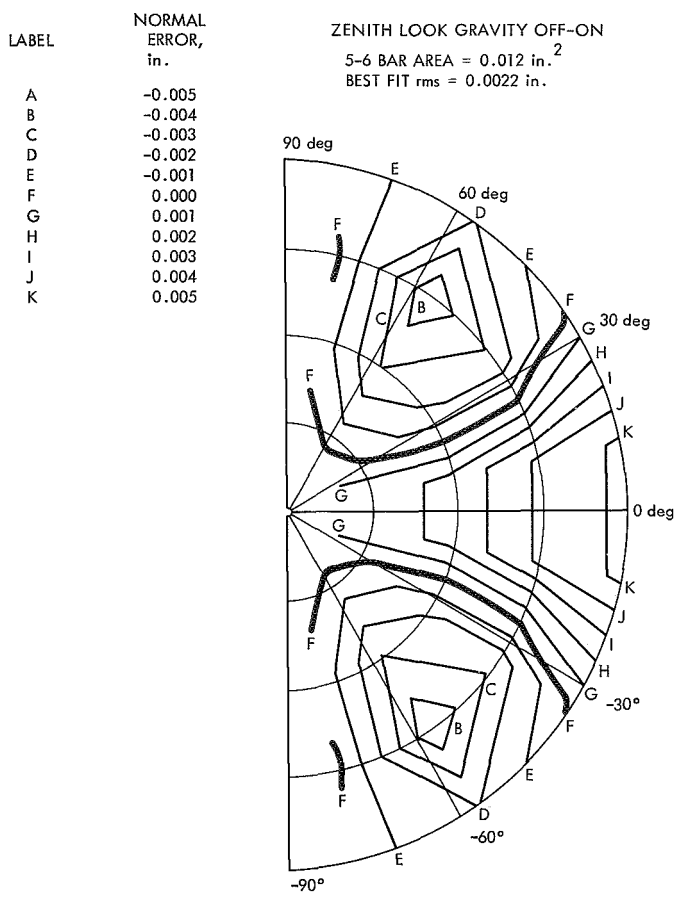


Fig. 3. Deflection contour map for zenith look, half dish, best fit rms = 0.002 in.

B. A Method for Selecting Antenna Rigging Angles to Improve Performance, *R. Levy*

1. Introduction

Methods to reduce root-mean-square (rms) deviations for antenna reflector surfaces are of great interest because of the unfavorable influence of large deviations upon the antenna's RF performance. By convention, these deviations are determined with respect to the best-fitting paraboloid by means of a mathematical procedure (Ref. 1) that permits rigid body motions of the reflector and a change in focal length. From an RF standpoint, the change in the paraboloidal configuration that is implied by the best-fit parameters[1] is either not objectionable or can be compensated with physical adjustment.

One major source of reflector distortion is the change in deflections caused by gravity as the antenna is moved over a range of elevation positions. This is the result of the relative change in direction of gravity loading (weight of structure and components) with respect to the structure.

[1]There can be up to six parameters consisting of three translations, two rotations, and one change in focal length.

2. Antenna Rigging to Reduce rms Deviations

A common field procedure that is used to reduce the extreme rms deviations is to adjust the surface as closely as possible to the exact paraboloid with the antenna positioned at some intermediate (rigging) angle. When this is done, the gravity distortions at the horizon and zenith positions are substantially reduced because these now are the result of only the change in loading from the rigging angle position. A good choice is to pick the rigging angle that causes the horizon and zenith rms deviations to be equal. It appears that this will provide the smallest peak rms deviation over the entire attitude range. Another, and possibly superior, choice is to select the rigging angle that minimizes an average weighted rms deviation over the entire elevation range. A logical weighting function for this would be based upon the relative importance of the surface accuracy at each elevation angle.

Usually, rigging angles have been selected by either of two procedures:

(1) By an arbitrary choice, say of 45 deg. This has the disadvantage that a poor guess will result in higher-than-necessary surface deviations.

(2) As the result of a parameter study using the analytic predictions for deviations. The study is performed by first assuming a trial rigging angle and then computing the best-fit rms deviations at the horizon attitude and at the zenith attitude. The trial rigging angle is incremented and the process is repeated until a sequence of extreme rms values is generated for a discrete set of trial angles. The selection is made from inspection of these values. This method is cumbersome and has the disadvantage of a large number of computations needed to best-fit each of the distortion vectors for every trial rigging angle. If weighted-average minimization is to be performed, the number of best-fit computations becomes excessive because of the necessity to compute the rms at elevations intermediate to the zenith and horizon attitudes.

3. Equations for Closed-Form Computations

A convenient method can be used to determine the best-fit rms distortions at any elevation attitude. This will facilitate a parameter study for determination of the rigging angle that minimizes the weighted-average rms. If alternatively, the objective is to make the rms at zenith equal to the rms at the horizon, the rigging angle can be obtained immediately by direct computation, thereby replacing the need to perform a parameter study. A major advantage of the approach to be described is that instead of performing the best-fit computations repeatedly with each incremental choice of rigging angle, these computations are executed once only for each of the two following vectors consisting of the cartesian components of the displacements at a prescribed set of reflector surface nodes:

$\{u_y\}$ = displacement vector for gravity loading applied parallel to the reflector y-axis (in a vertical plane when the antenna points to horizon)

$\{u_z\}$ = displacement vector for gravity loading applied parallel to the reflector z-axis (in a vertical plane when the antenna points to the zenith)

The assumption of the linearity of loading and displacements frequently permits great simplification in the conventional parameter study. The assumption will be made here also, and it follows from this that the displacements at any elevation attitude angle α are linear combinations of $\{u_y\}$ and $\{u_z\}$.

Thus,

$$\{u\} = \eta\{u_y\} + \zeta\{u_z\} \tag{1}$$

where

$$\left.\begin{aligned} \eta &= \cos\gamma - \cos\alpha \\ \zeta &= \sin\gamma - \sin\alpha \end{aligned}\right\} \tag{2}$$

and γ is the rigging angle.

Since the coefficients in Eq. (2) reflect the changes from the rigging position, Eq. (1) represents the change in distortion from the rigging position. Figure 1 shows the orientation of the coordinate system at the elevation angle α and the components of the gravity loading w.

For the computation of the rms best-fit rms surface deviations (SPS 37-40, Vol. IV, pp. 176–186), the vector of half pathlength RF deviations could be expressed as

$$\{\rho\} = [A]\{u\} + [B]\{h\} \tag{3}$$

where $[A]$ and $[B]$ are geometry matrices containing invariant data that depends only upon the coordinates of the reflector surfaces nodes, and $\{h\}$ is the 6×1 vector of the best-fit parameters.

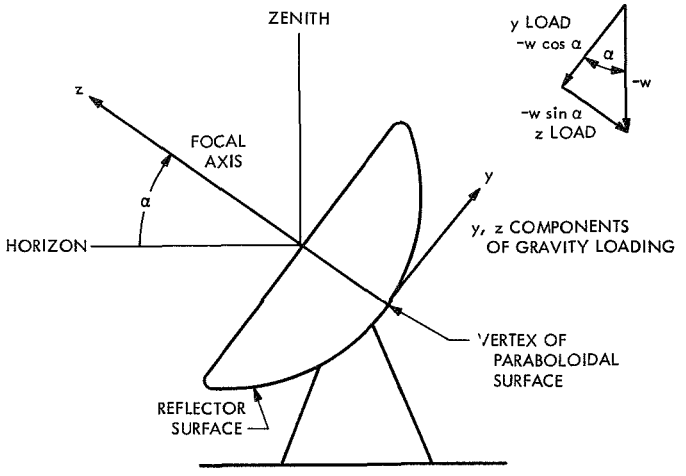


Fig. 1. Coordinate system for reflector

Forming the sums of squares, e.g., $\{\rho\}^t\{\rho\}$, and taking the partial derivatives with respect to $\{h\}$ results in the following solution for the best-fit parameters:

$$\{h\} = -[B^tB]^{-1}[B]^t[A]\{u\} \tag{4}$$

Substituting Eq. (4) in Eq. (3),

$$\{\rho\} = [R]\{u\} \tag{5}$$

in which

$$[R] = ([A] - [B][B^tB]^{-1}[B]^t[A]) \tag{6}$$

Then, substituting Eq. (1) in Eq. (5),

$$\{\rho\} = [R][u_y u_z]\begin{Bmatrix}\eta\\ \zeta\end{Bmatrix} \tag{7}$$

Forming the sums of squares,

$$\{\rho\}^t\{\rho\} = \eta^2(u_y^tR^tRu_y) + \zeta^2(u_z^tR^tRu_z) + 2\eta\zeta(u_y^tR^tRu_z) \tag{8}$$

Let

SS_y = sums of squares of half pathlength deviation for y-loading

SS_z = sums of squares of half pathlength deviation for z-loading

$\{\rho_y\}, \{\rho_z\}$ = corresponding half pathlength deviation vectors

Then

$$\{\rho\}^t\{\rho\} = \eta^2SS_y + \zeta^2SS_z + 2\eta\zeta\{\rho_y\}^t\{\rho_z\} \tag{9}$$

To make the sums of squares equal in the zenith and horizon attitudes, substitute the following in Eq. (9):

$$\left.\begin{array}{lll}\text{Zenith:} & \eta = \cos\gamma, & \zeta = \sin\gamma - 1\\ \text{Horizon:} & \eta = \cos\gamma - 1, & \zeta = \sin\gamma\end{array}\right\} \tag{10}$$

and then equate the zenith and horizon sums of squares. After canceling and rearranging terms, the following equation is obtained:

$$(2SS_y - 2\{\rho_y\}^t\{\rho_z\})\cos\gamma - (2SS_z - 2\{\rho_y\}^t\{\rho_z\})\sin\gamma = SS_y - SS_z \tag{11}$$

Equation (11) is of the form

$$A\cos\gamma - B\sin\gamma = C \tag{12}$$

where

$$A = 2SS_y - 2\{\rho_y\}^t\{\rho_z\} \tag{12a}$$

$$B = 2SS_z - 2\{\rho_y\}^t\{\rho_z\} \tag{12b}$$

$$C = SS_y - SS_z \tag{12c}$$

The solution of Eq. (11) for the desired rigging angle is

$$\gamma = \phi - \epsilon \tag{13}$$

where

$$\epsilon = \tan^{-1}\left(\frac{B}{A}\right) \tag{14a}$$

$$\phi = \cos^{-1}\left(\frac{C}{[A^2 + B^2]^{1/2}}\right) \tag{14b}$$

Hence, to compute the optimum rigging angle, it is necessary only to know the sums of the squares of the deviation vectors (which can be computed from the best-fit rms deviations) for both y-loads and z-loads acting alone and also to save the corresponding deviation vectors to compute the sum of the cross products. All of these quantities are readily available from the best-fit computation program. After the rigging angle has been

found, Eq. (9) in conjunction with Eq. (2) can be used to compute the sums of squares at any attitude, which leads immediately to the corresponding rms deviations.

4. Example 1

The rigging angle will be computed to obtain equal rms distortions at zenith and horizon for the 210-ft-diameter antenna. The following data is available (SPS 37-52, Vol. II, pp. 86–92) from existing computations based upon an arbitrarily chosen rigging angle of 45 deg:

Attitude	RMS deviation, in.
Zenith	0.075
Horizon	0.040

The analysis of this antenna assumed symmetry about the y–z and the x–z planes. As a result, it follows that

$$\{\rho_y\}^t \{\rho_z\} = 0$$

Although SS_y and SS_z were not given, Eq. (9) can be used to find SS_y/N and SS_z/N (N = number of surface points) as follows:

$$\text{Zenith:} \quad (0.075)^2 = (0.7071)^2 \frac{SS_y}{N} + (-0.2929)^2 \frac{SS_z}{N}$$

$$\text{Horizon:} \quad (0.040)^2 = (-0.2929)^2 \frac{SS_y}{N} + (0.7071)^2 \frac{SS_z}{N}$$

Solving these equations simultaneously provides

$$\frac{SS_y}{N} = 110.81 \times 10^{-4}$$

$$\frac{SS_z}{N} = 12.98 \times 10^{-4}$$

From Eqs. (12a), (12b), (12c),

$$\frac{A}{N} = 221{,}62 \times 10^{-4}$$

$$\frac{B}{N} = 25.96 \times 10^{-4}$$

$$\frac{C}{N} = 97.83 \times 10^{-4}$$

From Eqs. (14a), (14b),

$$\epsilon = 6°41'$$

$$\phi = 64°00'$$

and from Eq. (13), the desired rigging angle is

$$\gamma = 57°19'$$

Finally, straightforward substitution in Eq. (10) and then Eq. (9) yields, for either zenith or horizon attitudes,

$$\frac{SS}{N} = 32.61 \times 10^{-4}$$

or

$$\text{rms deviation} = 0.0571 \text{ in.}$$

The relative reduction in the maximum peak rms at the computed rigging angle with respect to the previous maximum, which occurred at the zenith position, for the arbitrary 45-deg rigging angle is

$$\text{reduction in maximum peak} = \frac{0.075 - 0.057}{0.075} \times 100\%$$

$$= 24\%$$

On the other hand, the increase of the previous minimum peak rms, which occurred at the horizon position, is

$$\text{increase in minimum peak} = \frac{0.040 - 0.057}{0.040} \times 100\%$$

$$= 42\%$$

This example illustrates the effectiveness of the present approach when the constraint is with respect to the maximum rms at any elevation attitude. Here the maximum peak that occurred at the zenith end of the elevation range was reduced at the expense of an increase at the horizon end of the range. The design for an alternative constraint, which entails a weighting function of the elevation attitude, is illustrated in the following example.

5. Example 2

Surface distortions for an 85-ft-diameter antenna that had been determined previously (SPS 37-52, Vol. II,

pp. 99–105) were used again as input to the best-fit rms computer program. This time, new instructions were provided to develop the necessary additional terms and perform the computations to find the rigging angle for equal rms at zenith and horizon.

The results obtained for this antenna, which were based upon 115 surface nodes distributed over one-half of the (symmetrical) structure, are given in Table 1. The maximum values of the best-fit rms surface distortions are compared for: (1) an arbitrary choice of rigging angle; (2) the rigging angle that is presently used for the antenna (found by an approximated parameter study); and (3) the rigging angle computed by the methods presented here. Table 1 shows that the minimum error is achieved with the rigging angle computed according to the proposed new method.

Table 1. Comparison of best-fit rms deviations for various rigging angles

Rigging angle, deg	Method	Best-fit rms half pathlength error, in.		Extreme rms error, in.	Reduction of extreme rms error, %
		Horizon	Zenith		
45.0	Arbitrary choice	0.0247	0.0146	0.0247	Basic
40.0	Parameter study	0.0228	0.0161	0.0228	7.7
31.3	Eq. (13)	0.0189	0.0189	0.0189	23.5

Finally, a parameter study was performed for this antenna to find the rigging angle for a weighted-average rms minimization. The two sums of squares and the cross product term [see Eq. (9)] developed by the best-fit program were used in a separate *Fortran* program that performed the computations. The weighting function was equal to the cosine of the elevation attitude angle. This provides the minimum expected value of the rms deviation for targets uniformly distributed with respect to elevation angles between zenith and horizon. Rigging angles and elevation angles in the study were incremented by 5 deg from horizon to zenith. For each rigging angle, the weighted rms deviations were computed for the complete set of elevation attitude angles. It was found to within four significant figures that rigging angles of either 25 or 30 deg were optimal. The corresponding weighted rms deviations are plotted in Fig. 2. The computer program to perform the parameter study required less than 7 sec of computer time on the Univac 1108.

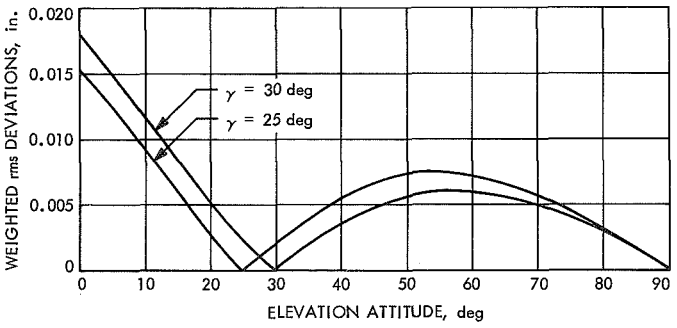


Fig. 2. RMS deviations weighted by cos α

6. Conclusions

A new procedure is described to compute the reflector surface rms gravity distortions from the best-fitting paraboloid at any arbitrary antenna elevation attitude. To apply the procedure, it is necessary to know the vectors of half pathlength deviations from the best-fitting paraboloids for two independent gravity loadings. Once these two vectors are known, no further best-fitting computations are performed for any arbitrary elevation attitudes where the best-fit rms distortion is to be determined. The simplicity with which the rms distortions can be computed for arbitrary elevation angles facilitates the selection of the optimum rigging angle to minimize gravity distortions. In the case where it is desired to achieve the minimum rms distortion over the entire elevation range, the optimum rigging angle can be obtained by a closed-form computation from the equation presented. In the case where it is desired to minimize the weighted-average rms distortion over a given elevation attitude range, the parameter study necessary to choose the best rigging angle is facilitated by the rapidity with which the rms distortions can be computed.

Reference

1. Katow, M. S., and Schmele, L., *UTKU/Schmele Parabaloid RMS Best-Fit Program, Volume I*, Technical Report 32-979. Jet Propulsion Laboratory, Pasadena, Calif., Apr. 15, 1970.

C. Low-Frequency Low-Level Stress Reversals on an Assembly of Bolted Joint Specimens,

V. B. Lobb

1. Introduction

JPL is currently using an inorganic zinc coating as an effective corrosion barrier on the 210-ft-diam antennas. This coating exhibits excellent corrosion protection and a high frictional coefficient for static loads. However, on previous joint integrity research (SPS 37-52, Vol. II) JPL found that the inorganic zinc coating broke down on the contact surfaces when undergoing stress reversal. Once this breakdown occurs, there is an appreciable loss in bolt preload and joint slip results. This breakdown occurred at a 20,000-psi (design allowable) net section stress level.

To ensure that this desirable coating could be used without joint slippage, JPL needed to know the exact stress level at which this coating breakdown would occur. A research program was initiated to further study bolted structural assemblies with plain mill-scale faying surfaces and inorganic zinc paint faying surfaces. Plain mill scale was assigned the role of a control surface by which the inorganic zinc painted surface was to be compared. A series of different design stress levels was used to determine or bracket the desired breakdown point.

Previous tests had already demonstrated the ability of the inorganic zinc paint surface to resist slip displacement under static loading. This research was directed toward short-term, relatively low-level repetitive load reversals. This research investigated the effect of surface conditions, bolt design, bolt strength level, and applied stress versus joint-slip displacement. The objective was to determine which design parameter and what stress level would affect inorganic zinc-painted coatings on bolted connections.

2. Research Program

A series of low-frequency, low-level stress reversal tests was conducted on assemblies of bolted-joint specimens. The test series consisted of 116 separate scheduled tests conducted on 58 specimens. The tests were conducted in a special H-frame test stand which alternately reversed specimen load to a given percent of design allowable. Load was maintained within an accuracy of $\pm$ 5%. Slip displacement was measured on both sides of the specimen to a sensitivity of 0.0001 of an inch. Dial-indicating gauges mounted on the specimens for the duration of the test were employed.

The test evaluated three parameters as they influenced joint behavior under complete reversal of load. These parameters are as follows:

(1) Four bolt types: ASTM A325, ASTM A490, hexagon-structural, and interference-body-bearing bolts.

(2) Two faying surface conditions: dry mill scale and inorganic zinc paint.

(3) Plate stresses: 25%, 37½%, 50%, 62½%, 75%, and 87½% of design allowable net section stresses.

Parameters which remained constant for all tests included the following:

(1) All joints were two-bolt single shear connections.

(2) All joint material was ASTM A36 steel.

(3) Each joint was made of a hot-rolled steel angle which was connected to a hot-rolled steel gusset plate.

(4) All bolts were ¾ 10 UNC.

(5) All bolts were tightened in accordance with a turn-of-the-nut installation procedure.

(6) All A325 bolts were assembled with one washer under the nut.

(7) All A490 bolts were assembled with washers under both the nut and the head.

(8) All bolts, nuts, and washers were in the as-received black-finished condition.

(9) All the joint geometry, allowable stresses, and design procedures followed the recommendations of the American Institute of Steel Construction.

3. Test Procedure

Testing was performed in a 200,000-lb capacity Enerpac H-frame hydraulically operated press (Fig. 1). Special coupling fixtures were employed, which permitted both tension and compression loads to be applied to the specimen without removing it from the press for the duration of the test.

Each specimen was installed in a series with a strain gauge load transducer used to monitor load for the duration of the test. Two sets of dual dial-indicating gauges

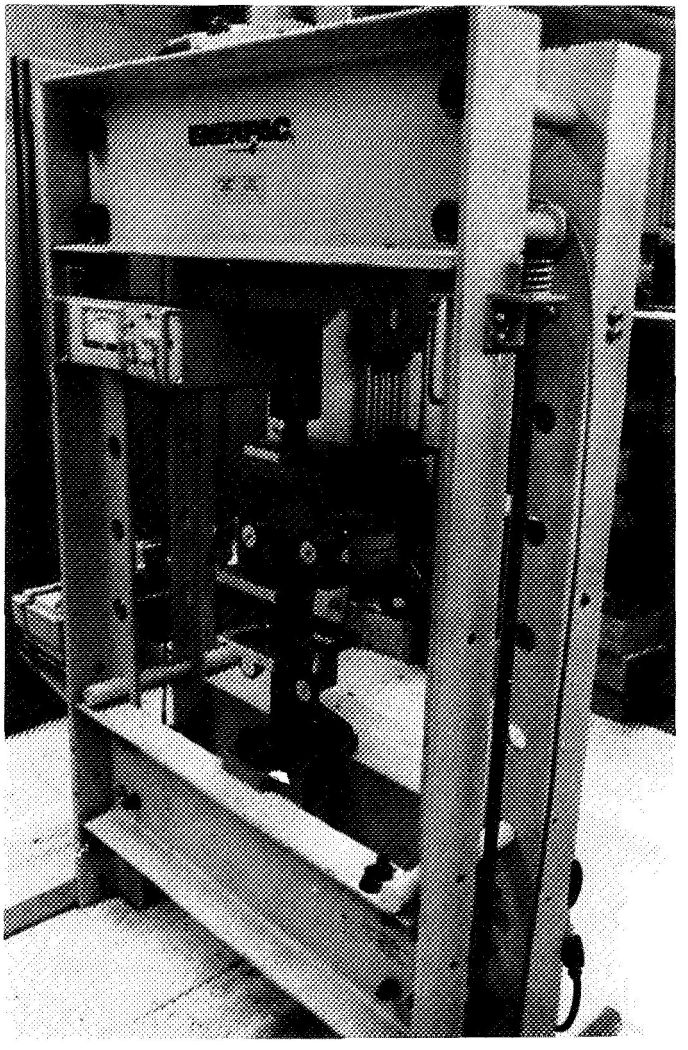


Fig. 1. Test stand with specimen in position

were mounted on the specimen for the duration of the test. All dial gauges were positioned to read positive when the specimen was tensioned and negative when the specimen was compressed. Readings were taken prior to loading, at 15-min increments for a total of 60 min., and at unloading. This procedure was repeated for reverse loading and continued for the duration of the test. Initial loadings always tensioned the specimen. Complete tests included five tension loadings of 1-h duration each and five compression loadings of 1-h duration each for a total of 10 h per test. Test loads were generated and maintained manually with a hydraulic hand pump. Temperature fluctuations never exceeded 4°F.

4. Results and Discussion

To compare the effect of specimen faying surfaces coating and bolt-type to slip displacement, an averaging procedure was employed. Since an individual test consisted of 51 readings, a point-by-point comparison of the value is almost impossible. Therefore, average values were calculated for each hour, even though slight variations were evident due to displacement with respect to time, or other inherent variations due to gage performance and the accuracy employed in reading the dial indicators. A typical example of this is shown in Fig. 2, which shows the results of three different tests averaged out. In general, the test results show no significant difference in slip displacement between hex and bearing-type bolts, when used with plain mill-scale-finish members. This is also true for the A325 and A490 strength levels of each type of bolts.

Inorganic-zinc-coated specimens were analyzed and are shown in Fig. 3. It can be seen in this case that bolt type does have an influence on slip-displacement behavior. With A325 bolts, the effect of bolt-type is not so pronounced, as slip displacements are somewhat the same until a net section stress of 17 psi is achieved. Beyond this point, the ability of the hex bolt to resist slip displacement is impaired, and a significant difference exists between it and the bearing bolt. With inorganic-coated specimens, at the same test loading, the slip displacement for both A490 hex and bearing bolts was found to be greater than for the A325 bolts. The net sectional stress of the A490 connection is lower than that of the A325 connection, at a given load level; the bolt shear stresses are identical. The only variables which could possibly explain the greater A490 joint slip over the A325 joint is

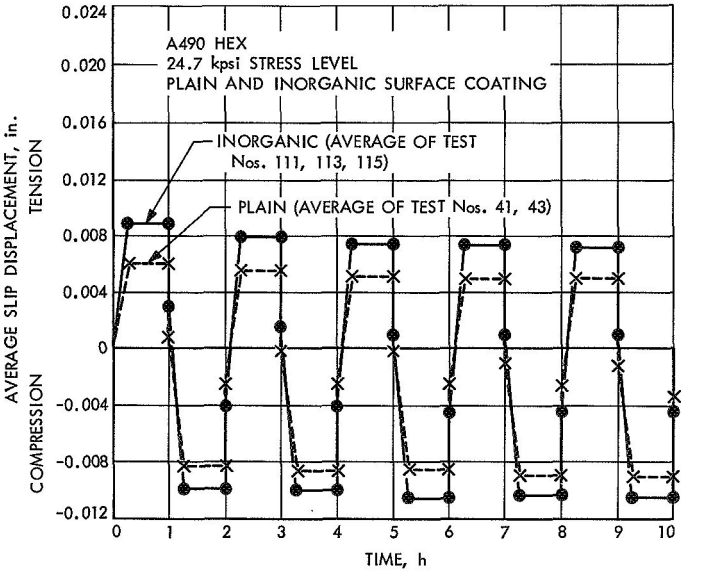


Fig. 2. Average hourly slip displacement versus time

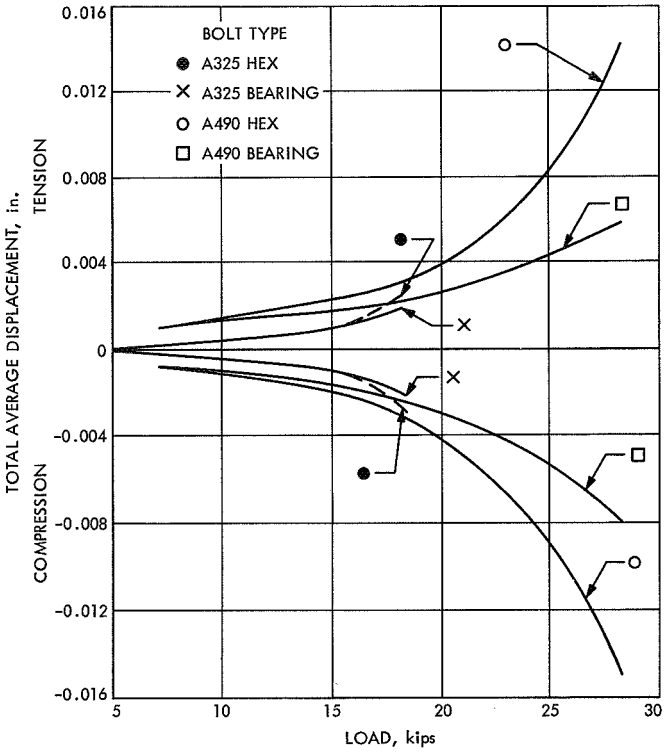


Fig. 3. Load versus average total slip displacement for inorganic-zinc-paint-coated specimens for different bolt types

narrowed to those of bolt preload and the effect of the inorganic coating itself.

A490 bolts with inorganic-coated specimens demonstrate that bolt type (hex versus bearing) does have a pronounced effect on slip displacement, as shown in Fig. 3.

The A490 hex bolt allowed over 100% greater slip (at 100% design allowable) than the A490 bearing bolt. Therefore, it can be concluded that, however detrimental the effect of the inorganic coating may be with hex bolts, this can be negated through the use of bearing bolts with the same strength level.

The effect which inorganic faying surfaces coating has on joint slip displacement is illustrated by the behavior of A490 hex bolts as shown in Fig. 4. Slip displacement of inorganic-coated specimens is greater than those with plain faying surfaces at all test loads. At the lowest load level of 7.1 kips, or 25% of the design allowable, the inorganic-coated specimens averaged 0.004 and 0.003 in. greater slip displacement in tension and compression, respectively, than the plain mill-scale specimens.

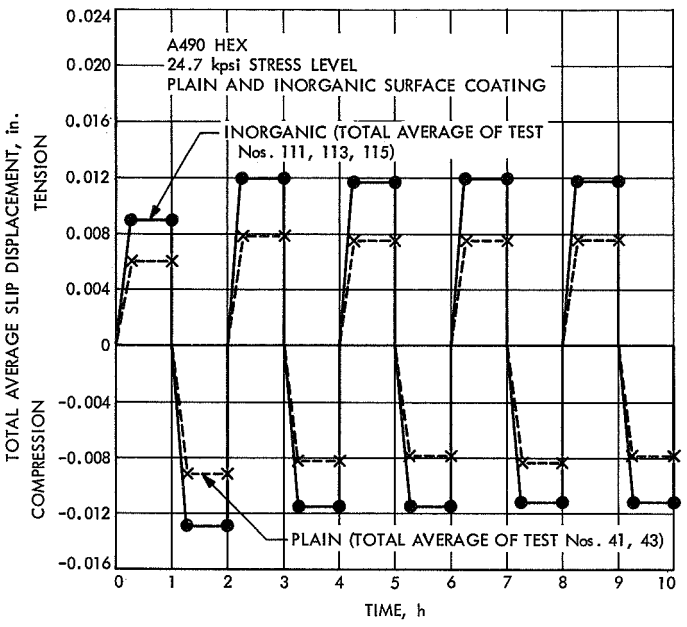


Fig. 4. Total average hourly slip displacement versus time

At the highest scheduled test loading of 24.7 kips, or 87½% of design allowable, inorganic-coated specimens averaged 0.004 in. greater slip in tension, and 0.003 in. greater slip in compression, than the plain mill-scale specimens. At all loads, inorganic specimens averaged 0.002 in. greater slip in tension and compression than mill-scale specimens.

Although the net section stresses for all A490 connections as compared to A325 connections, at a given test load, were 32% lower, and the bolt shear stresses were identical, A490 connections with inorganic zinc-paint faying surfaces had greater slip displacements than A325 connections. This behavior could be due to the greater clamping pressures induced by the A490 bolts with respect to the A325 bolts (A490 bolts have 20% greater clamping pressure than A325 bolts). This increased clamping pressure apparently causes a premature microscope physical beakdown in inorganic-zinc coating. This breakdown in combination with the abrading action of the stress reversal results in a loss of bolt elongation, which leads to a loss in the original bolt clamping pressure and, consequently, a reduction in the resistance to slip displacement.

Spot checks of the coating indicate an average coating thickness of 4 mils (0.004 in.). The normal grip length, including washers for A490 bolts, was approximately one diameter, which is ¾ in. The A490 bolt elongation with

approximately the same grip is 0.005 in. when properly loaded. The initial tensioning of the bolts will cause some degree of coating compression. However, the combining effect of short bolt grips (low length-to-diameter ratio), thick relatively soft inorganic zinc painting coating, and the abrading action of stress reversal will cause a significant loss in bolt tension.

The above behavior may be true only for connection with relatively low bolt-length-to-diameter ratio. Thicker connections (larger length-to-diameter ratio) may not experience this phenomenon. Because of increased bolt elongation, due to longer grip length, (with identical coatings thicknesses) these connections would be less sensitive to losses in bolt elongation due to coating breakdown and may provide improved slip resistance.

5. Summary and Conclusions

The principal purpose of the research was to determine which design parameter and stress level would warrant use of inorganic zinc-painted coatings on specimens undergoing low-cycle low-load complete-stress reversals. In addition, part of the research program was to confirm previous stress-reversal studies concerning the effect of load level and type of fastener. The conclusions which were drawn from this research program follow:

(1) Slip displacement under complete stress reversal is a function of both applied external load and bolt shear stress. Slip behavior is not directly related to net sectional stresses.

(2) The effect which inorganic zinc paint faying surface coating has on joint slip displacement (all slip displacement refers to average total slip displacements) is illustrated by the behavior of A490 hex head structural bolts. Slip displacement is considerably greater than plain faying-surface specimens at all test loads. At lowest load levels (7.1 kips, 25% of design allowable), inorganic-coated specimens averaged 0.0004 in. greater slip than plain faying surfaces. At the highest scheduled test loading (24.7 kips, 87½% of design allowable) inorganic-coated specimens averaged 0.004 in. greater slip than plain faying surface specimens. Overall, inorganic coating averaged, at all loadings, 0.002 in. greater slip displacement than plain surface specimens.

(3) A significant fact resulting from the program is illustrated by the ability of the high-strength bearing bolt to negate the effects of faying surface coatings. Test results indicate that no difference can be attributed to faying surface coating when using either A325 or A490 high-strength bearing bolts.

(4) When comparing the effect of bolt types, relative to slip displacement, it is evident that no significant difference exists between A490 hex-head and A490 high-strength bearing bolts with plain faying-surface specimens. When the same comparison is made with inorganic-coated faying surface, the effect of the coating comes into play, and the hex structural bolts indicate inferior performance. This leads to the same conclusion stated previously; namely, that inorganic-coated faying surfaces permit greater slip displacements than plain surfaces.

(5) In respect to A325 strength level bolts, plain or inorganic coated specimens did not influence the slip displacement until 87½% of design allowable net section stress was achieved for coated specimens. Net section stress is greater than the aforementioned result in greater slip displacement for A325 hex bolts.

(6) Supplementary tests conducted at 100% design allowable on organic-coated specimens were comparable to previous tests done on research reported in SPS 37-52, Vol. II, pp. 92–97.

(7) A logical explanation for the greater slip displacement experienced by A490 connections with organic zinc paint faying surfaces than the A325 connections is two-fold: the A490 bolts have 20% greater bolt clamping pressure than A325 bolts, and this apparently causes premature microscopic physical breakdown of the inorganic-zinc coating; and this breakdown, in combination with the abrading action of the stress reversal results in a loss of bolt elongation and, hence, a loss of original bolt clamping pressure, and consequently increases slip displacement.

(8) Previous research (SPS 37-52, Vol. II) showed that faying surface coatings affected a joint's resistance to slip. From this previous research, a ranking of increased resistance to slip is as follows:

(a) Faying surfaces painted with red oxide primer give the least resistance to slip displacement.

(b) Zinc (hot-dipped galvanized) faying surfaces increase resistance to slip.

(c) Plain mill-scale faying surfaces further increase resistance.

(d) The most resistant were inorganic-zinc-painted faying surfaces.

This ranking of joint resistance does not hold true when the joint is subjected to load reversals. These tests showed different results when the specimens were loaded at 100% of design allowable stresses (20,000 psi) and 88% allowable design stress (17,000 psi). At these stress levels the inorganic zinc coating broke down and showed a greater slip than that of the mill-scale faying surface.

For the 210-ft deep space tracking antenna, where the stress levels generally run from 4,000 to 12,000 psi, the inorganic zinc silicate coating will not cause excessive slip under these low-level stress reversals. For stress levels of these magnitudes, the inorganic zinc silicate paint still affords an acceptable coating. However, the thickness of the inorganic zinc coating should be controlled in the area of the faying surface and should not exceed a thickness of 4 mils, and this coating should not be employed in any area subjected to stress reversal which could exceed the 17,000-psi stress level.

D. Switched-Carrier Experiments, *R. B. Kolbly*

The DSIF has a requirement, in support of the *Viking* mission, for simultaneous transmission of two carriers from one antenna. These carriers would be separated by approximately 5 MHz and carry ranging modulation. The power level required is a minimum of 40 kW in each carrier and its associated sidebands. The 400-kW transmitter can supply this easily if excitation is supplied at the two frequencies, but past experience and test data have shown that intermodulation products in the receiver passband well above threshold are generated by this approach. These products are generated whenever two carriers exist in the nonlinear klystron; there is even evidence that they may be generated in the waveguide equipment or the antenna structure.

If the excitation could be switched, so that the klystron and antenna see only one carrier at a time, these products would not be generated. Such an approach has been proposed by M. Easterling and the Goldstone Development Support Group of the RF Systems Development Section has started an experiment to determine the feasibility of this method.

The basic idea is to time-share the klystron between the two carriers by switching back and forth between them. If the switching is arranged so that each carrier is transmitted 50% of the time, each carrier will have 25% of the power and the other 50% of the power will be in the sidebands about each carrier at the odd harmonics of the switching frequency. If the switching frequency is chosen to be high compared to the bandwidth of any tracking loops and high compared to the command data rate, but low compared to the ranging modulation and different from the command subcarrier, there will be no interference due to the switching sidebands. The problem of choosing a switching rate relative to the command subcarrier is similar to the problem of choosing two noninterfering telemetry subcarriers.

There is a possible problem during the acquisition of a carrier in that the carrier tracking loop in the spacecraft might lock to a switching sideband. This is very unlikely if any reasonable switching frequency is used. If it should be considered to be of concern, the problem could be avoided by switching with a pseudonoise (PN) sequence which would spread the spectrum of the switching sidebands and leave no strong signal for a possible spurious lock. The spread spectrum might cause some degradation of the tracking loop performance, however, by putting some of the sideband energy into the loop bandwidth. This suggests that a proper strategy might be to use PN switching during acquisition and square-wave switching during tracking. Changing from one to the other would have no effect on the signal being tracked by either spacecraft.

The proposed approach has a 50% efficiency independent of the modulation on either carrier. Thus, it would provide two 100-kW carriers from the 400-kW transmitter or two 5-kW carriers from a 20-kW transmitter. The equipment in Fig. 1 is being utilized to investigate the feasibility of the above approach. The diplexer assembly is installed in such a manner that

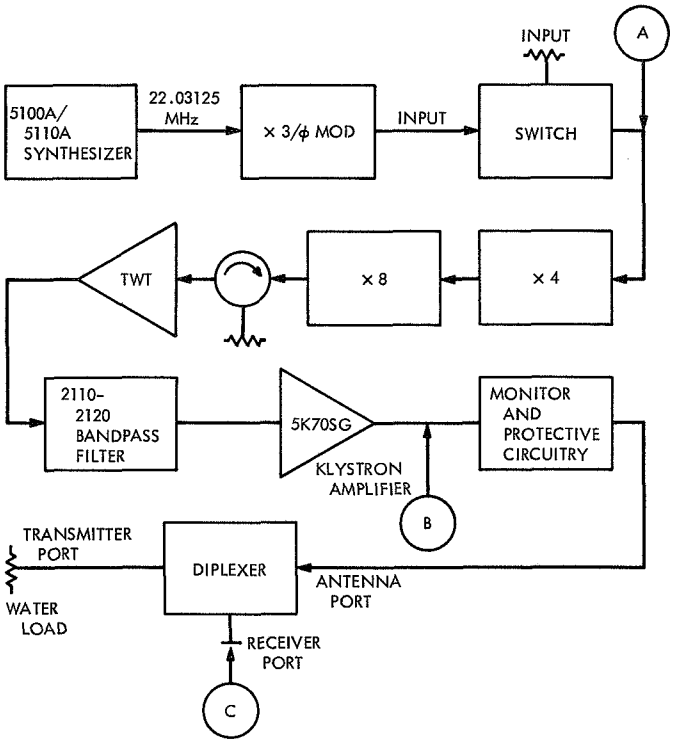


Fig. 1. Switched-carrier test configuration

there is very low loss (less than 1 dB) between the output of the klystron and the detector in the receiver (2290–2300 MHz) band. Since we will be measuring very high order sidebands (i.e., when switching 2115 MHz at 500 kHz, the 360th sideband falls at 2295 MHz), the 500-kHz switching frequency is derived from a stable source, the same synthesizer used to furnish the 22.03125 MHz for the exciter.

At this time, experiments are being run using one carrier switched on and off at a 500-kHz rate. Since the 500 kHz is coherent with the exciter frequency, we are able to determine the source of a particular sideband or spurious product. By using one frequency rather than two, we are concerned with only one-half of the total spectrum, simplifying the analysis of data. Since we can treat each carrier as being keyed on and off, we will suffer no loss of data by this approach. Also, the use of the 2115.000000-MHz carrier and the 500.000-kHz switch rate allows modulation sidebands to be separated from any other spurious products.

A diode switch has been built to perform the 66-MHz switching function. Figure 2 illustrates the final configuration of the switch. The diode switch is driven with a biphase 0 to −16-V 500-kHz square wave which is supplied from modified 3C S-PAC digital logic cards.

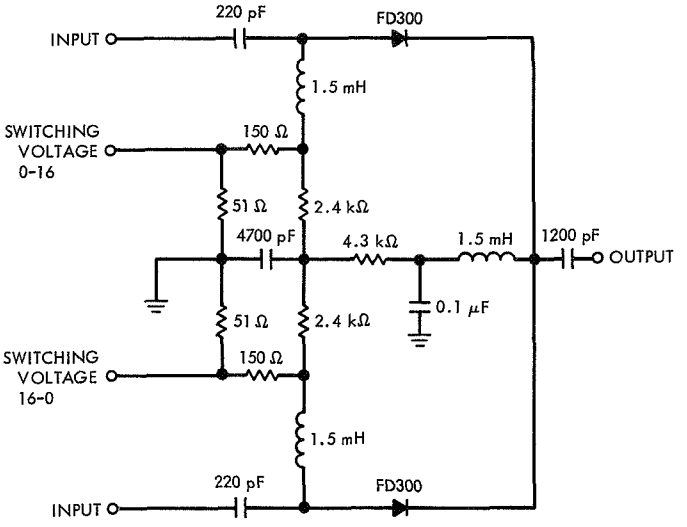


Fig. 2. Diode switch schematic

Figure 3 is the spectrum measured at point A in Fig. 1. This indicates that the diode switch is performing as predicted. Figure 4 is the spectrum at the output of the klystron amplifier. The sidebands have been filtered by several bandwidth limited items at this point, including the klystron amplifier. It will be noted that because the klystron is operating in a saturated mode the rise time of the output pulses is sharpened.

This investigation will continue with the following experiments:

(1) Investigation of spurious and modulation products in the output of the klystron in the DSIF receiver band (2290–2300 MHz). This will be done with a sensitive phase-lock receiver and a low-noise preamplifier.

(2) Investigation of spurious and modulation products close to the transmitted carriers.

(3) Measurement of system temperature in a standard DSIF diplex configuration with switching.

The results of these tests and the feasibility of using this system to support dual carrier commitments will be examined.

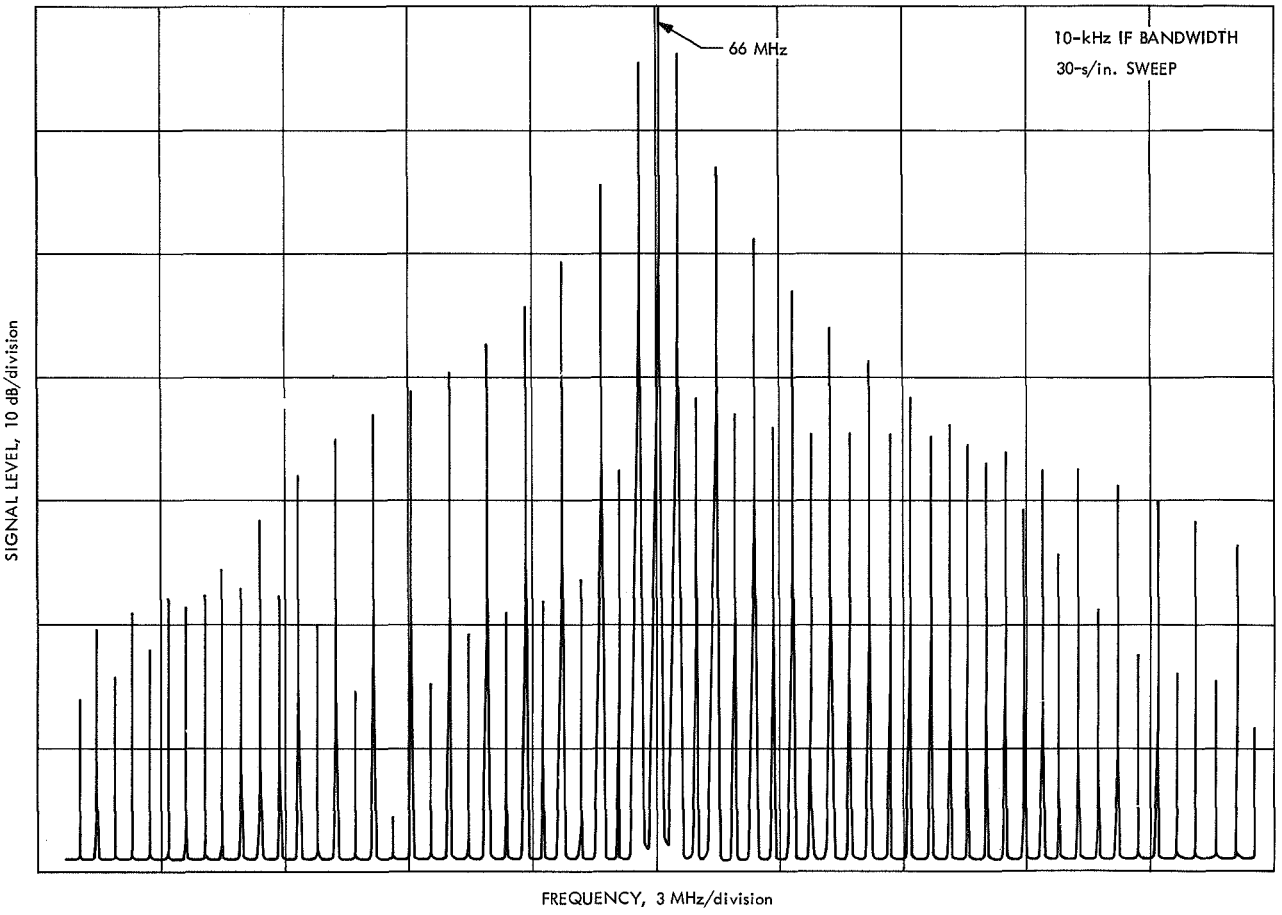


Fig. 3. Spectrum at output of diode switch

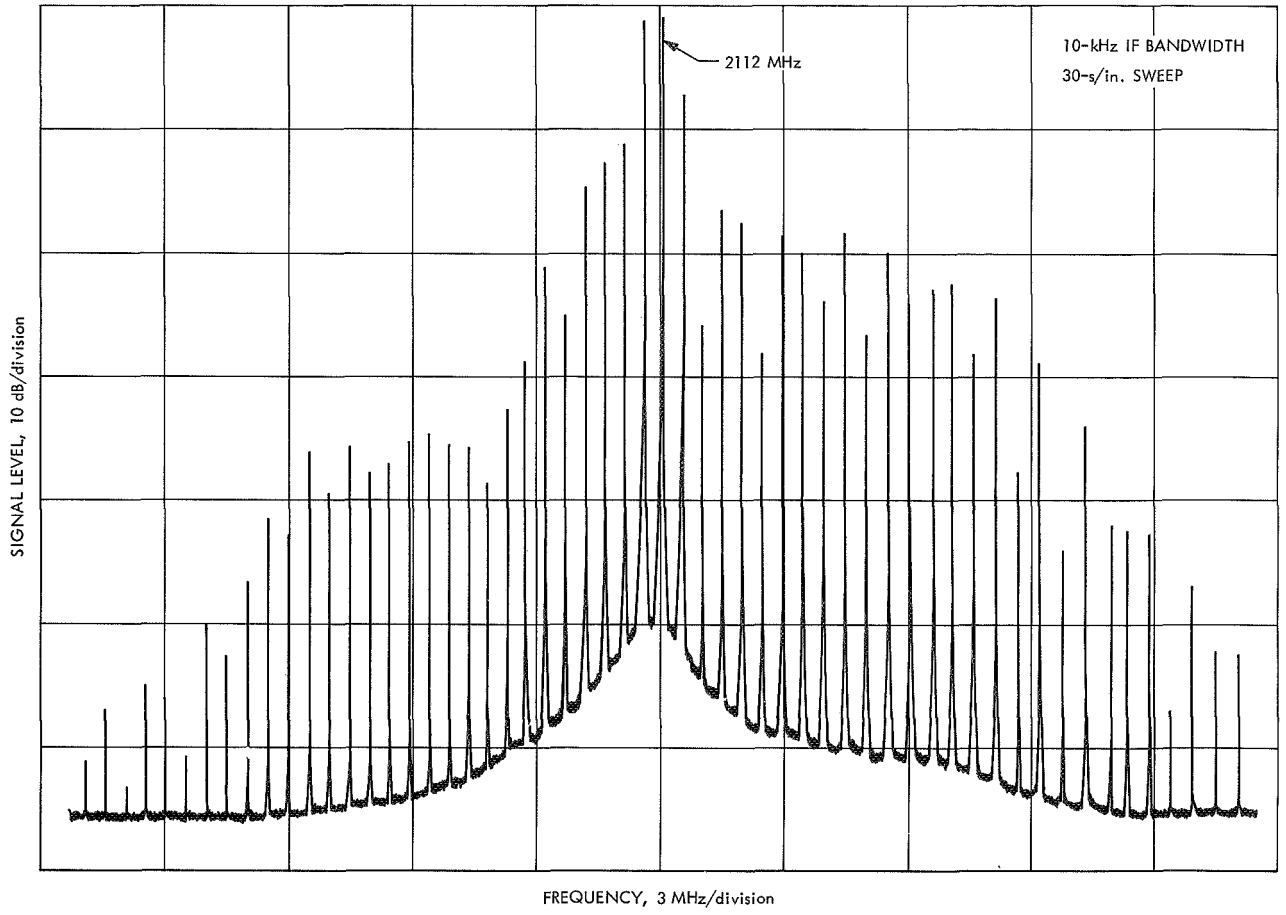


Fig. 4. Spectrum at output of klystron amplifier

E. DSS 13 Operations, *E. B. Jackson*

During the period June 16 through August 15, 1970, the major activities at DSS 13 included a special clock synchronization experiment, routine pulsar observations for 24 h each week, cooperative planetary radar activities and continued development of the ephemeris update tracking program (SPS 37-60, Vol. II, pp. 70–71). Installation of the Model IV receiver/exciter was completed and long-term testing begun; installation of the 100-kW, 2115-MHz transmitter was also completed.

Continuing with the clock synchronization experiment which was begun on June 1, 1970, daily transmissions were made to DSSs 14, 41, 42, 51, and 62, the National Bureau of Standards (NBS) and the U.S. Naval Observatory (USNO). This experiment is to determine the resolution and accuracy possible with the existing system, and explore the reliability of both the transmitting and receiving stations with regard to capability to support daily long-term operation.

The experiment worked very well, and some system weaknesses were uncovered; for example, the narrowband receiving filter had a tendency to drift in center frequency. Additionally, it was shown that the system resolution is such that the effects of the Moon's topography cause a periodic variation in the timing of the received data. The experiment was completed on July 31, 1970, and transmissions continue only to DSS 14, the NBS, and the USNO on a regularly scheduled basis.

An additional cooperative experiment was performed with DSS 14 and the Haystack Facility of the Massachusetts Institute of Technology. This was an effort to make precision ranging measurements of the planet Mercury when it was in appropriate position with respect to the Sun. Although good data were received from the test targets (Venus and the Moon), data from Mercury were too noisy to provide the precision range points desired. The amount of signal returned from Mercury, at 7840 MHz, using a 10-μs range gate, is insufficient to provide the required signal-to-noise ratio.

A continuing program of making once-weekly precision range measurements on the planet Venus was begun during the last week in July. Using a 10-μs range gate, good data have been collected for three weeks. This experiment will continue through the end of 1970.

The ephemeris update tracking program (scan and correct, using receiver–SCOUR) has been developed to the point where successful automatic tracking of radio stars and spacecraft has been demonstrated. In this scheme, an output from the receiver proportional to the received signal is monitored by a digital computer while the computer commands the antenna to perform a scan. The variations in the received signal level while scanning are used by the computer to develop error signals to drive the antenna until the radio star or the spacecraft are maintained in the center of the antenna beam. Preliminary indications are that the signal degradation is quite small, but further testing and optimization is underway.

Long-term testing of the Model IV receiver/exciter has commenced. This receiver/exciter, constructed primarily of integrated and hybrid circuits, has real-time digital-computer-controlled first local oscillator frequency so that automatic acquisition of a spacecraft can be effected. Several successful spacecraft tracks have already been accomplished and more will be performed as the various operational parameters (sweep rate, phase lock loop bandwidth, sweep direction, etc.) undergo optimization. Spacecraft tracked to date include *Mariners VI* and *VII* and *Pioneers VIII* and *IX*. Observed signal levels vary from −168.8 dBm (*Pioneer VIII*) to −173.2 dBm (*Pioneer IX*).

VIII. SFOF Development

DEVELOPMENT AND IMPLEMENTATION

A. SFOF Digital Television Assembly, *F. L. Singleton*

1. Introduction

The digital television assembly (DTV) is a part of the user terminal and display subsystem in the SFOF. The purpose of the DTV is to provide the SFOF with a flexible real-time and near-real-time data display system. This display system will provide a multichannel, computer-generated display of alphanumeric and graphic information to support the various space flight missions starting with *Mariner* Mars 1971. The displayed information will be used for data analysis, decision making, monitoring, and management information.

The DTV will provide 60 channels of digital television displays with both alphanumeric and graphic display capability. These displays will be shown on various TV monitors throughout the SFOF, as required. When a permanent copy of a display is required, six hard copy printers will be available to print a hard copy of the selected display.

This article is a summary of the requirements for such a data display system and a description of that system.

2. Requirements for a Display System

The need for a display system with capabilities of the DTV has become increasingly apparent with each new mission. Improvements in data transmission between the spacecraft and the DSIF and between the DSIF and the SFOF have increased the volume of data available during any given time in a mission. Also, the use of larger and faster computers within the SFOF has made it possible to provide larger amounts of processed data to the end user in real time. To output this data to the user requires a display system with a high data output rate and multiple output devices.

Previous displays within the SFOF consisted mainly of line printers and X–Y plotters distributed throughout the various user areas of the SFOF. The limitations with the use of these previous displays were:

(1) Low information output rates.

(2) Nonexistence or serious limitations of displays containing both alphanumeric and graphic information.

(3) Inflexibility of display once printed (a new display must be written to show changes in data).

(4) Limited distribution of displays because of the bulk and expense of the individual display devices.

(5) Considerable maintenance time involved in keeping electromechanical display devices operating.

(6) Operational requirements to keep devices supplied with paper and toner (ink).

(7) Noticeable delay in outputting a complete display once that outputting has begun.

3. Design Approach

In order to overcome previous limitations, the design of a new display system was approached with the following basic requirements:

(1) Real-time data display.

(2) Multiple data displays.

(3) Wide distribution of the displays.

(4) Simultaneous display of alphanumeric and graphic data.

(5) Use of volatile displays to allow easy updating without complete display regeneration.

(6) Access to nonvolatile prints of displays when needed.

The method determined to meet these requirements included the following:

(1) Accepting data input from the SFOF data processing computers.

(2) Converting digital inputs from the data processing computers into an output suitable for driving television displays.

(3) Providing real-time displays on television monitors.

(4) Providing a multiple channel system with the ability to store and regenerate video (television) information for each channel.

(5) Writing over the video (television) data stored in the system with new information from the data processing computers.

(6) Providing wide distribution of the displays by use of the GCF television assembly (TVA) distribution system within the SFOF.

(7) Using an alphanumeric and graphic data generator to combine data on the video storage memory.

(8) Providing adequate resolution for graphic data.

(9) Providing a hard copy printing system that can convert the video storage signal of any display channel into a permanent copy when requested.

4. DTV Description

The DTV consists of a computer subassembly and a display subassembly as shown in Fig. 1. The computer subassembly interfaces with two IBM 360/75 computers in the SFOF central processing system and accepts digital messages intended for output to the display subassembly. It also controls the hard copy request and print sequences. The computer subassembly acts to buffer display messages and distribute them to the correct locations in the display subassembly in the most efficient manner possible.

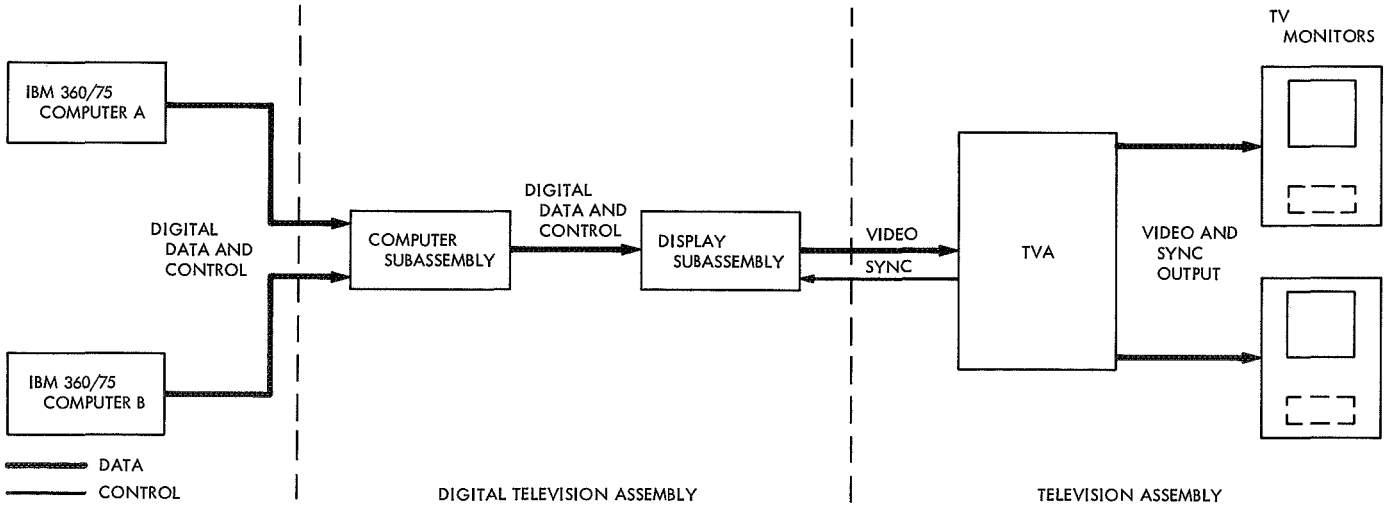


Fig. 1. DTV block diagram

The display subassembly contains the electronics necessary for the conversion of digital messages into video signals (display generation) and also for storage and regeneration of a complete video display pattern for each DTV channel. The video signals are outputted to the TVA for distribution throughout the SFOF on TV monitors.

Also, contained in the display subassembly is the equipment necessary for providing a printed hard copy of any DTV channel's display.

5. DTV Computer Subassembly

Equipment within the computer subassembly consists of two computer interface controls (CICs) and two Control Data Corporation CDC 3100 computers. Each computer includes the following peripherals: three digital magnetic tape units, a card reader-punch, a line printer, and a shared disk memory storage device. This subassembly is shown in Fig. 2.

Each CIC provides an interface between one IBM 360/75 computer and either of the two CDC 3100 computers. Each 360/75 input can be independently switched to either CDC 3100 computer by the CIC.

Each CDC 3100 computer has a 16k-word memory, a 24-bit word size, and a cycle time of 1.75 μs. One CDC 3100 will be used on-line to accept input data via the CIC and to control the outputting of this data to the display subassembly. Depending on mission support requirements, the second CDC 3100 will be used as a back-up computer; it will also be used as a display development computer for continuing development of the DTV.

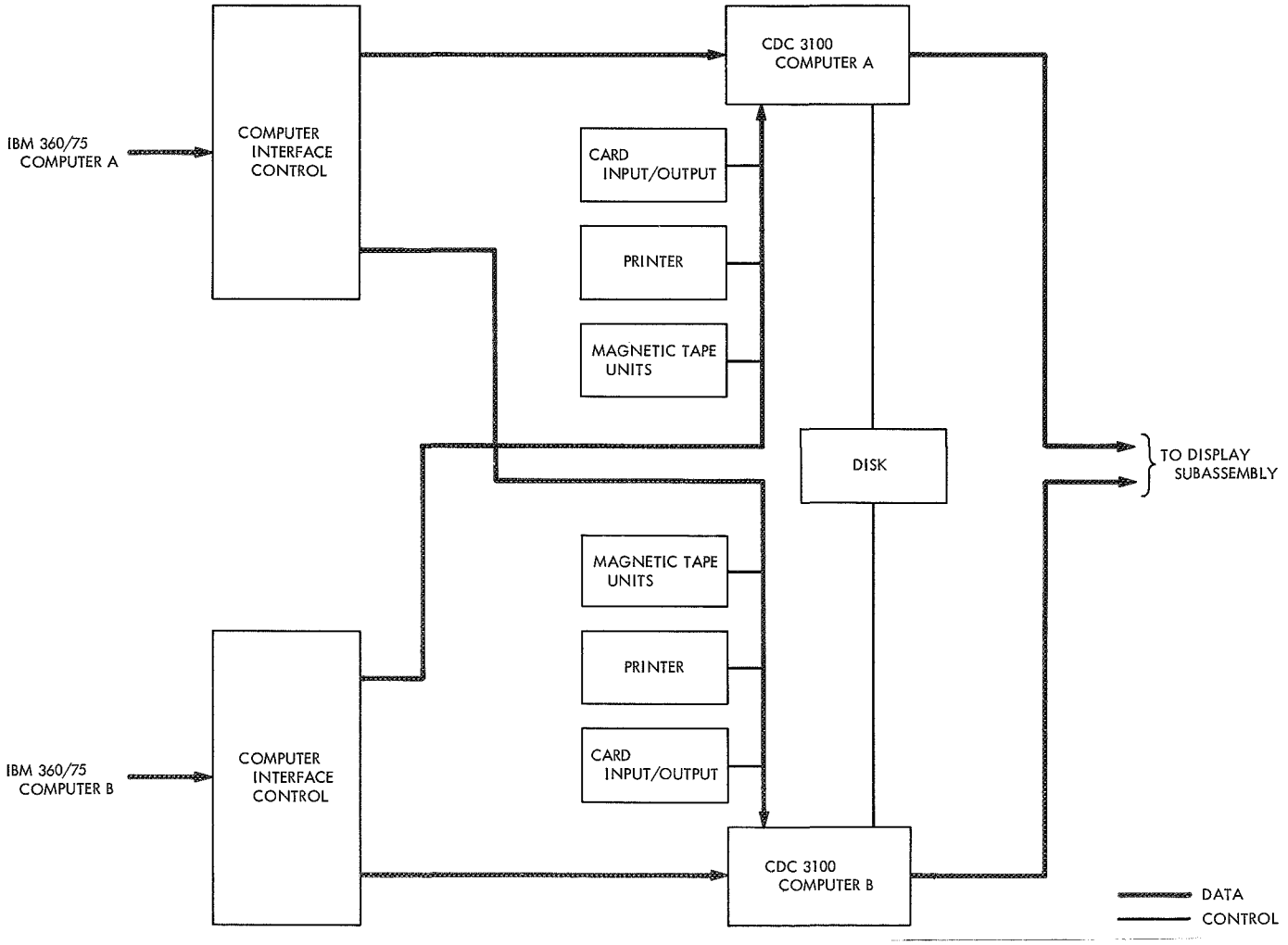


Fig. 2. DTV computer subassembly

The peripherals of each CDC 3100 are used to support operation of the computer subassembly and also assist in the further development of the DTV.

Within the computer subassembly, one CDC 3100 will normally handle all incoming DTV messages and hard copy requests. It will buffer these messages in queues slated for the display channels in the display subassembly and output them to the display channels as they are ready to accept them.

Also, requests for hard copy printouts of DTV channels are accepted by the on-line CDC 3100. A request causes the input messages to the requested DTV channel to be held in their queue until the display channel information can be transferred. The information is transferred to an alternate hard copy storage in a matter of milliseconds. Display messages are then permitted to continue to the requested display channel.

6. DTV Software

Operation of the computer subassembly will be internally controlled by the DTV operating system (DTVOS) software which is resident in the on-line CDC 3100 computer. The DTVOS will provide management of display messages, hard copy requests, and hard copy printouts.

The DTVOS will accept formatted display data generated in the IBM 360/75s and route it to the appropriate DTV channel in the display subassembly. Additionally, the DTVOS will service requests for a hard copy of data in any DTV channel by the issuance of a hard copy print instruction when a valid request is received. To insure integrity of the hard copy printout, the DTVOS will inhibit updates to the channel to be copied until printout of a valid copy is assured.

Supporting software within the DTVOS will include routines for task management, core management, and input/output handling. Other features will include a DTV diagnostic pattern generator, a basic statistics gathering routine, system initialization and operator communication with the software, as well as supervisory functions and peripheral control routines.

7. DTV Display Subassembly

This subassembly consists of the following units as shown in Fig. 3: one system control unit, three display generator units, one display image buffer, six copy request units, and six hard copy printers.

The system control unit interfaces with the computer subassembly and provides interface and control for the display generators, display image buffer, and copy request units.

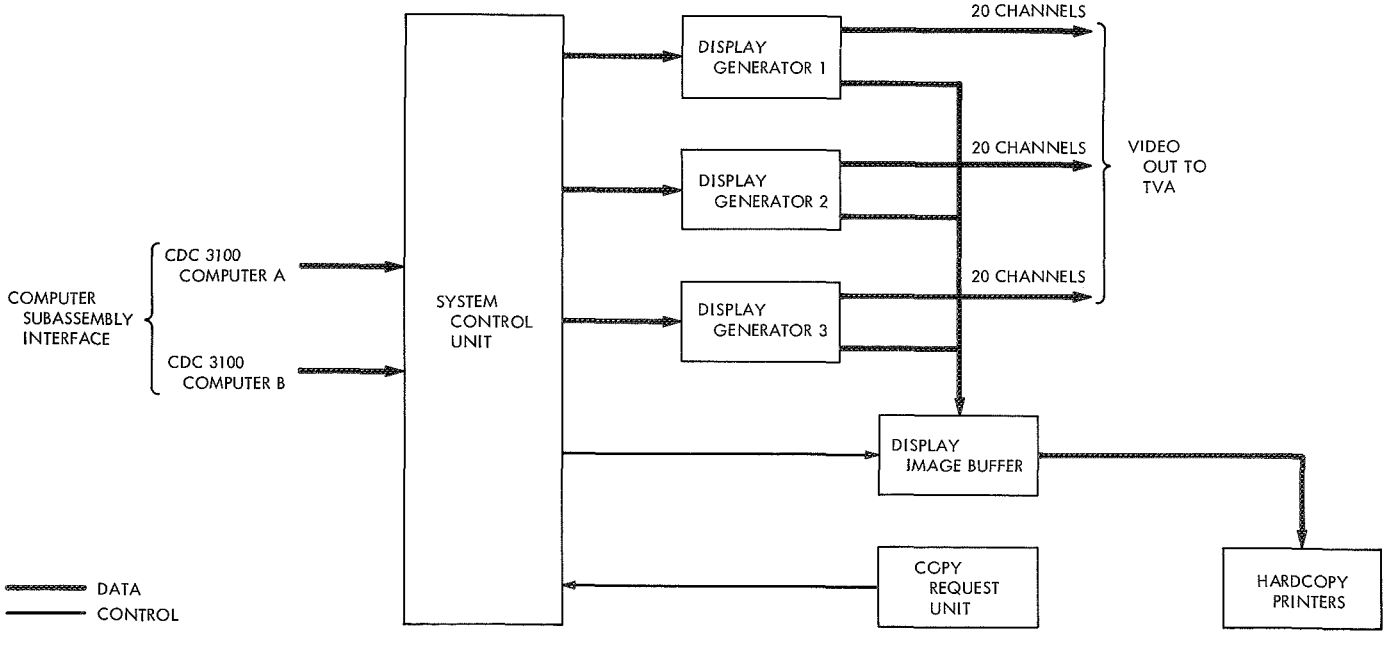


Fig. 3. DTV display subassembly

Each display generator contains a digital buffer for storing incoming digital display messages, logic for character and graphic data generation, and a disk memory for video storage with automatic signal regeneration capability for 20 channels. The incoming digital messages are converted into video information and written on the correct channel of the disk memory by the character and graphic data generation logic. These channels are synchronized to the TVA so that their outputs can be directly transferred to the TVA for distribution in the SFOF.

The display image buffer contains disk memory storage of hard copy output data and the selection and interface logic to transfer any DTV channel video signal into its disk memory storage. It has the storage capacity and output logic for driving six hard copy printers.

Each of the six copy request units contains the controls for selecting and requesting a hard copy of any DTV channel. The logic for the copy request unit is contained in the system control unit.

Each of the six hard copy printers contains the logic and mechanics to convert the output from the display image buffer into a 6- × 8-in. image of the requested DTV display channel and print it out in about 11 sec.

Both the copy request units and the hard copy printers will be located in the various user areas of the SFOF. Hard copy requests will be initiated by the user at his copy request unit and the requested display will be outputted to the printer associated with that copy request unit.

When a copy request is made, it is routed to the computer subassembly by the system control unit where it activates the hard copy request sequence. The computer subassembly issues a DTV channel transfer and print command to the display image buffer (via the system control unit). The display image buffer then causes the video signal of the requested DTV channel to be transferred from the display generator into its video storage memory. This signal is then output in a format compatible to the hard copy printer.

8. Synchronization and Distribution of DTV Output

One video output signal from each DTV channel will be transmitted to the GCF TVA. The TVA, which is a part of the GCF wideband system, will supply the switching circuitry and amplifiers needed to distribute the DTV video signals to the TV monitors in the mission support areas.

In order for the DTV video signal to be correctly transmitted to the TVA, it must be synchronized with the TVA. Correct transmission will be accomplished by synchronizing the disk memory of each display generator to the TVA by means of a servo control system. The servo derives its timing control from the TVA synchronizing signals.

9. Capabilities of the DTV

The DTV will contain 60 channels of display, controllable from the IBM 360/75 computers. Each display will have an element resolution of 640 elements (horizontal) by 480 elements (vertical) and simultaneous graphic and alphanumeric display capability. The DTV will have a repertoire of 96 standard ASCII characters that it can generate and display in four sizes.

The alphanumeric content of any display can consist of either 40 or 80 characters per row and either 20 or 40 rows, depending on the character size used.

The DTV will have six hard copy printers, each with an associated copy request unit. The printout from these printers will cover a 6- by 8-in. area; the requested display will be produced element for element in 11 sec from start of printing. The print will be made on an 11-in.-wide roll of paper and the paper advance is designed to provide an 8½- × 11-in. sheet of paper with each printout.

Table 1. Capabilities of the DTV

Characteristic	Capability	Remarks
Display channels	60	Expandable to 120
Graphics generation	Yes	
Display resolution	640 × 480 elements	
Alphanumeric generation	Yes	
Character repertoire	96 ASCII characters	
Character sizes	4	
Characters per row	40 or 80	Dependent on character size
Character rows per display	20 or 40	Dependent on character size
TVA compatible	Yes	
Input source	Two IBM 360/75 computers	
Hard copy printers	6	Expandable to 19
Printed image size	6 × 8 in.	
Hard copy resolution	640 × 480 elements	
Hard copy print time	11 sec	
Number of copy request units	6	Expandable to 31

Because the DTV outputs into the television assembly, the DTV displays are capable of wide distribution throughout the SFOF on either 9- or 14-in. TV monitors. Furthermore, because every DTV display will be under the control of the IBM 360/75 computers, each will be fully capable of being updated in real time or of being completely replaced with a new display format.

The DTV has been designed for expansion capability to 120 display channels, 19 hard copy printers, and 31 copy request units. The capabilities of the DTV are shown in Table 1.

10. Conclusion

The DTV assembly is currently being implemented. The computer subassembly is an existing display computer system that is being modified to this requirement. The display subassembly is being manufactured by Data Disc, Inc., and is slated for delivery in September 1970.

B. Contrast Ratio Determination for the SFOF Video Image Display, *J. J. Volkoff*

1. Introduction

The SFOF Mark III central processing system will include a video processing facility that has the capability to display and print tonal images through video processing of digital image data. This display and print processing shall be performed within the video image display subsystem (VIDS) of the SFOF. The images will be displayed on high-resolution monitors that have the capability to display a large number of discernible grayshades requiring a minimum ratio for a given monitor configuration and ambient light condition.

This study is concerned with the effect of ambient light upon the contrast ratio, and upon the output light of the display. These results shall be used to establish a monitor configuration and ambient light condition to achieve a desired photometric quality and a discernibility requirement.

2. Definition

The contrast ratio of a cathode ray tube (CRT) display is defined as the quotient of the maximum luminance divided by the minimum luminance. Reduction of contrast ratio results from the scattering and the reflection of light. Light can be dispersed by reflections from the tube walls, scattering from CRT geometry, and scattered electrons. These dispersions can be significantly reduced through filtering, tube geometrical configuration, and tube design. Even though these dispersions reduce contrast ratio, their effect upon contrast ratio is small compared with the effect of the reflections from ambient light incident upon the faceplate of the CRT. Reflections from ambient light, however, can be reduced through the use of filter devices.

Consider the CRT configuration as shown in Fig. 1. An electron beam imparts energy to the phosphor screen producing light which is transmitted through the glass faceplate and filter screen of the CRT system. Ambient light passing through the filter screen is reflected from the glass faceplate and the phosphor screen. Combining the transmitted light generated by the electron beam with the reflected ambient light results in the total light emanating from the faceplate of the CRT.

The amount of light contributed by the total reflections of ambient light and emanating from the CRT system is

$$L_a(1-f)^2\,[t^2 r_s(1-r_g)+r_g]$$

where L_a is the total illumination of the ambient light incident upon the CRT system, t is the light transmission factor of the glass faceplate, r_s is the phosphor screen reflection coefficient, r_g is the glass faceplate reflection

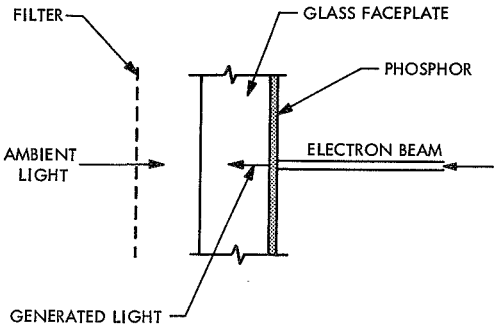


Fig. 1. Cathode ray tube configuration

coefficient, and f is the filtering factor of the filter screen. It is assumed that the reflection coefficient of the filter screen is negligibly low.

For a given luminance L_s generated at the phosphor screen by the electron beam, $(1-f)\,t\,L_s$ represents the amount of light transmitted from the CRT system. The total luminance L_t emanating from the CRT system, including the effects of ambient light, is then

$$L_t = (1-f)\{t\,L_s + L_a(1-f)[t^2 r_s(1-r_g) + r_g]\} \tag{1}$$

The highest contrast ratio which can be displayed by the CRT may be considered to be the luminance gain G. Gain is expressed as white-level luminance $L_{s\,(\max)}$ (maximum level), divided by the black-level luminance $L_{s\,(\min)}$ (minimum level), each produced at the phosphor screen such that $L_{s\,(\max)} = G\,L_{s\,(\min)}$. Using Eq. (1), the contrast ratio C_r of the displayed image, including ambient light incident upon the CRT system, can be expressed as

$$C_r = \frac{L_{t\,(\max)}}{L_{t\,(\min)}} = \frac{L_{s\,(\max)} + (1-f)\,k\,L_a}{L_{s\,(\max)}\,G^{-1} + (1-f)\,k\,L_a} \tag{2}$$

where

$$k = \frac{t^2 r_s(1-r_g) + r_g}{t}$$

3. Contrast Ratio Evaluation

For the evaluation of the contrast ratio of the VIDS high-resolution monitors as a function of ambient light, the following set of CRT configurations are established:

(1) The phosphor screen is comprised of Willemite (P4). The reflectivity coefficient r_s for a representative P4 screen was measured under actual florescent lighting in the SFOF and found to be 0.76.

(2) The faceplate consists of tinted glass that has an average light transmittance t of 0.63 integrated over the spectrum of light (Ref. 1). Reflectivity coefficient r_g was taken to be 0.04 (Ref. 2).

(3) The CRT gain G may be varied from a minimum value of 33 to a maximum value of 100.[1]

Using these values for the parameters in Eq. (1), total luminance L_t at white level as a function of ambient light for various neutral-density filter screens is shown in Fig. 2. This figure is based on a nominal output luminance of 50 ft-L emanating from the CRT at zero ambient light condition. By Eq. (2), contrast ratio also as a function of ambient light for various neutral-density filter screens and for a CRT gain of 33.0 is shown in Fig. 3.

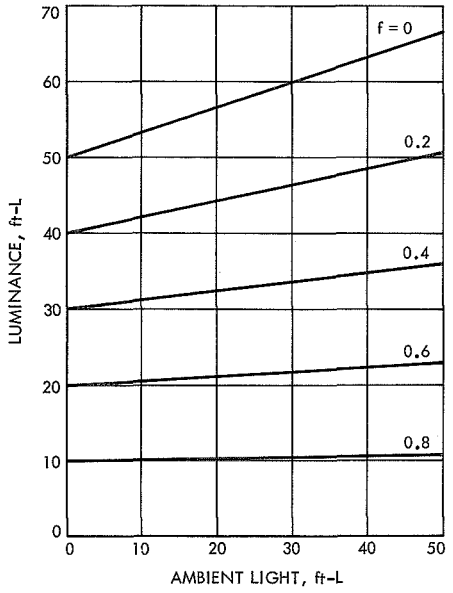


Fig. 2. White-level luminance from CRT system vs ambient light for various filtering coefficients f

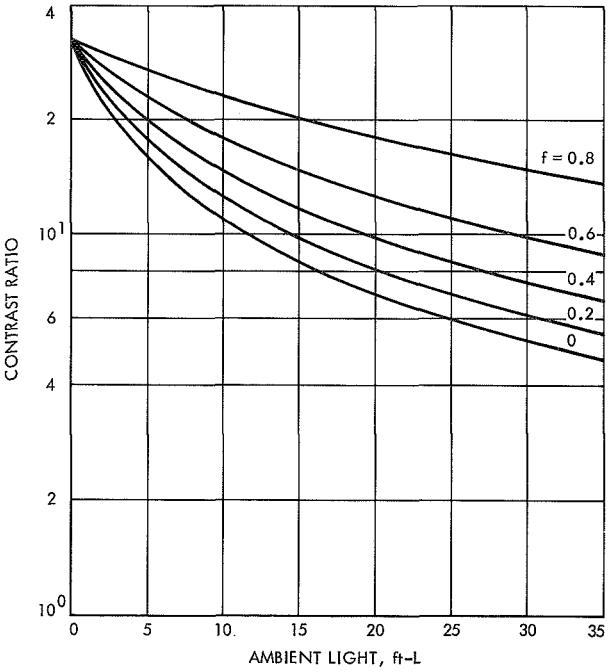


Fig. 3. Contrast ratio vs ambient light for various filtering coefficients f and a CRT gain of 33.0

[1]Diem, W., *Video Image Display Subsystem, Space Flight Operations Facility, Design Requirements for*, May 25, 1970 (JPL internal document).

4. The Effect of Gain

The effect of gain upon displayed contrast ratio of a monitor system, possessing a dc restorer which maintains the back porch despite changes in signal amplitude can be expressed by a factor ρ defined by

$$\rho = \frac{C_r(G)}{C_r(G_0)} \tag{3}$$

where $C_r(G_0)$ is the contrast ratio at a reference gain factor G_0 and $C_r(G)$ is the contrast ratio at some given factor G. From Eq. (2), ρ may be expressed as

$$\rho = \frac{L_{s\,(\max)}\, G_0^{-1} + (1-f)\, k\, L_a}{L_{s\,(\max)}\, G^{-1} + (1-f)\, k\, L_a} \tag{4}$$

Factor ρ as a function of ambient light for different filter factors is shown in Fig. 4, where G_0 is referenced as the minimum gain of the VIDS monitors. It can be seen from Fig. 4 that the effect of CRT gain to enhance the displayed contrast ratio grossly decreases with ambient light. However, CRT gain can be used to increase the degree of discernibility of the displayed image since contrast ratio increases with CRT gain. The maximum contrast ratio which can be developed by the VIDS monitor at a given ambient light and filter screen configuration can be found by combining the values from Figs. 3 and 4 into Eq. (3).

5. Circular Polarizer

A neutral-color circular polarizer screen filter can be used to enhance the contrast ratio of an image displayed on a CRT viewed under extreme ambient light conditions. A circular polarizer screen consists of a lamination of a linear polarizer layer and a retardation (quarter-wave) layer. The purpose of circular polarization is to attenuate the reflections of ambient light to negligible levels. This is evident from the typical transmission curves of a circular polarizer described in Fig. 5 (Ref. 3), which shows that about 35% of the light from the display may pass through a circular polarizer screen filter while attenuating the reflections of the integrated ambient light to a negligible level. An estimate of an application of circular polarization to the VIDS high-resolution monitor shows that the display should have a contrast ratio of about 31.8 under an ambient light condition of 100 ft-L. However, the maximum luminance of the display which can be observed is 17.5 ft-L.

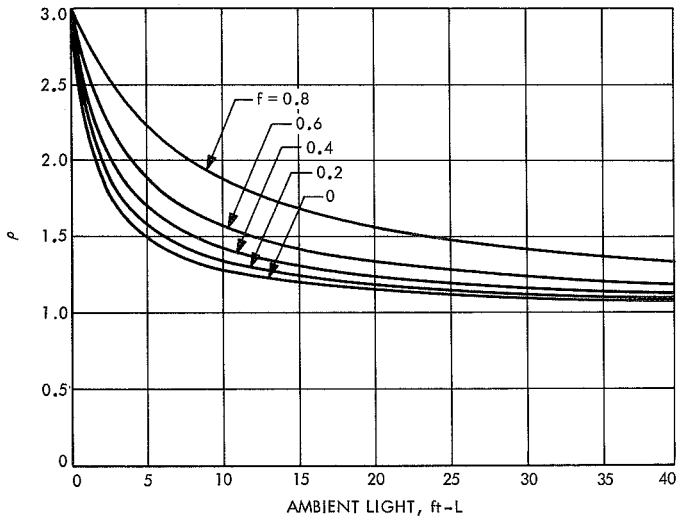


Fig. 4. Factor ρ vs ambient light for various filtering coefficients f

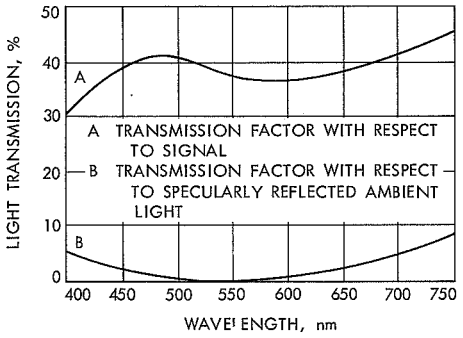


Fig. 5. Light transmission factor as function of wavelength for typical neutral-color circular polarizer screen filter

6. Conclusion

By knowing what contrast ratio is required to discern a desired number of grayshades displayed on the CRT as a function of white-level brightness, the maximum ambient light as a function of filter coefficient factor may be obtained from Fig. 3. Figure 3 represents a minimum gain condition for the VIDS monitor system. Also, the maximum total illuminance emanating from the monitor can be found from Fig. 2. From these evaluations, a trade-off can be made between monitor configuration, maximum ambient light, and the desired photometry of the displayed image.

References

1. *Optical Faceplates,* Ferranti-Microspot Cathode Ray Tubes Catalog, Vol. 1. Ferranti Electric, Inc., New York, Oct. 1967.

2. Born, M., and Wolf, E., *Principles of Optics.* Pergamon Press, London, England, 1959.

3. *Polarized Light,* FT 3374-A. Polaroid Corp., Cambridge, Mass., Feb. 1967.

C. DSN Mark IIIA Simulation Center Development,

R. G. Polansky

1. Introduction

The purpose of this article is to describe the current development activity that is taking place in the DSN Simulation Center (Simcen) in preparation for *Mariner* Mars 1971 and *Pioneer F* support.

2. Background

The current, or Mark II, configuration of the Simcen does not contain sufficient hardware to handle the requirements that have been levied upon it by the DSN and the *Mariner* Mars 1971 Project. Major hardware expansion is being carried out at this time to provide the necessary capabilities for support during the *Mariner* Mars 1971 era. The Simcen configuration being implemented now has been designated the Mark IIIA Simcen. Table 1 shows the characteristics of both the Mark II and Mark IIIA Simcens.

Table 1. DSN Simcen subsystem characteristics

Subsystem	Mark II	Mark IIIA
Telemetry	Fixed pattern Controllable pattern Locally controlled math model	Same as Mark II + dynamic/on-line responsive math models
	Long loop: 1 spacecraft to 2 DSSs simultaneously	Long loop: 2 spacecraft to 3 DSSs simultaneously
	Short loop: simultaneous 2-DSS outputs, 1 spacecraft, to SFOF	Short loop: simultaneous 3-DSS outputs, each of 2 spacecraft, to SFOF
Command	Operator entry to change telemetry	Same as Mark II + on-line inputs to telemetry model; simulation of DSS replies to SFOF commands
	Long loop: 1-DSS read-write-verify support	Long loop: processes command replies from 3 DSSs
	Short loop: 2-DSS read-write-verify support	Short loop: 3-DSS command system simulation
Tracking	Off-line generation of punched paper tape	On-line generation of tracking data to SFOF from Simcen
	8-tape playback, manual selection	Real-time selection of DSSs/formats; concurrent with telemetry simulation
Monitor	Fixed-pattern data blocks to SFOF	Dynamic/responsive DSS model to SFOF; simulates 3 DSSs simultaneously; on-line inputs to telemetry and tracking models
Operations control	SFOF clocks setable to simulated GMT	Same as Mark II
	Voice communication to DSN controllers simulating DSS operation	Same as Mark II

3. Hardware Description

Block diagrams for the Mark II and the Mark IIIA Simcens are shown in Figs. 1 and 2. A close look at these figures reveals that the following changes have been or are in the process of being made:

(1) The EMR 6050 input/output (I/O) computer's core memory has been expanded from 16k to 32k words.

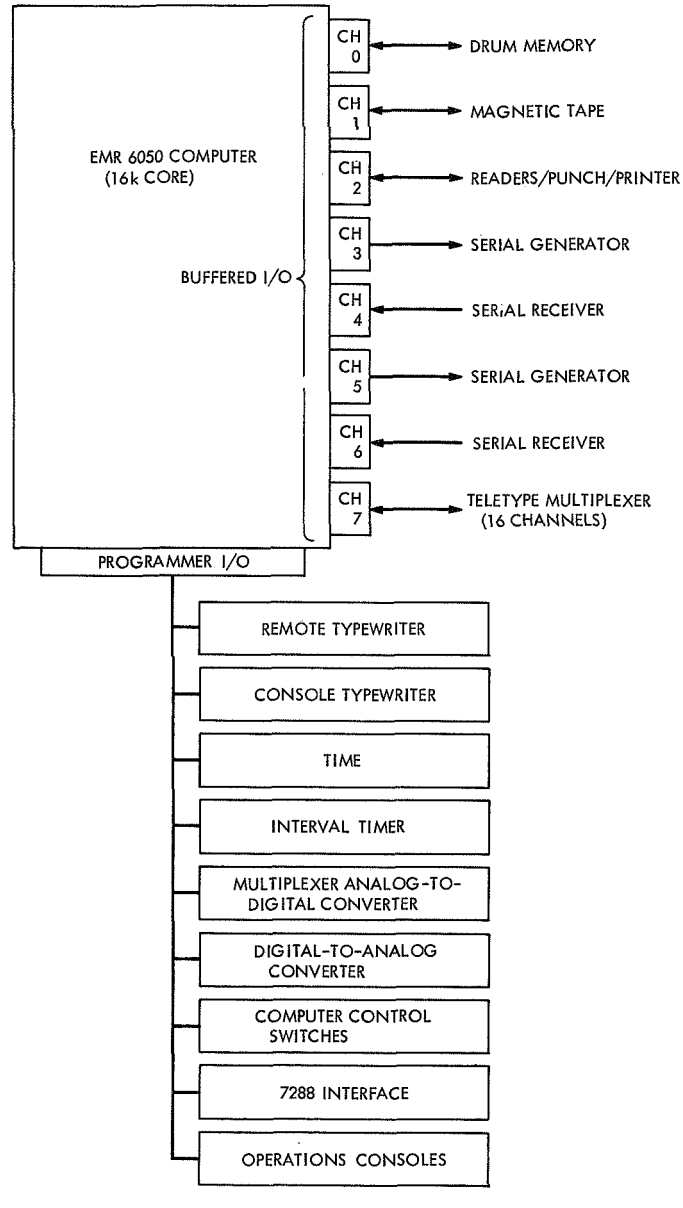


Fig. 1. Mark II DSN Simcen subsystem

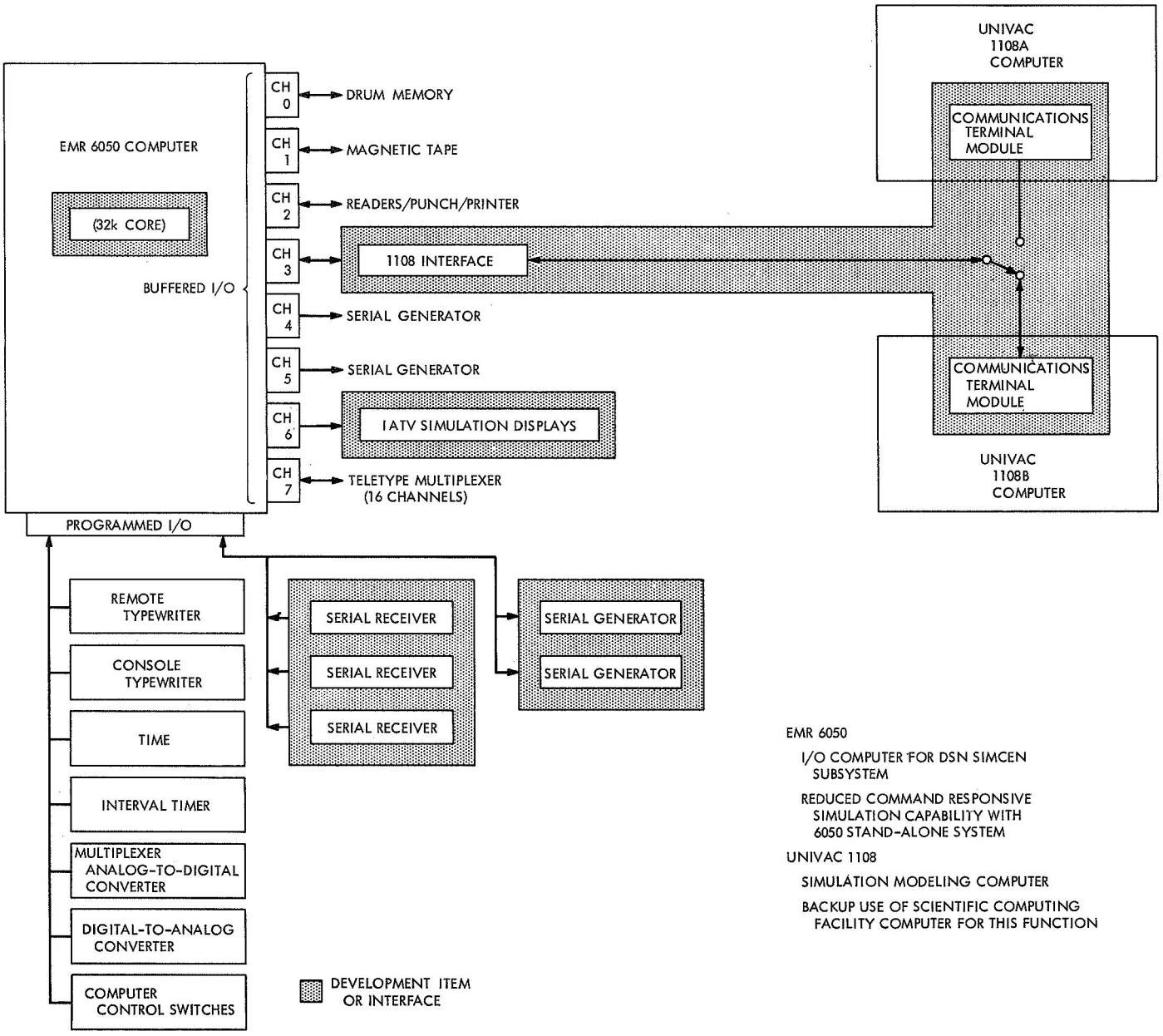


Fig. 2. Mark IIIA DSN Simcen subsystem

(2) The serial receivers which were interfaced with buffered I/O channels 4 and 6 of the EMR 6050 have been removed from the system.

(3) The serial generator that was interfaced with buffered I/O channel 3 of the EMR 6050 has been moved to channel 4.

(4) An EMR 6050/Univac 1108 interface device has been designed and will shortly be installed between buffered I/O channel 3 and the 1108 computer. (The 1108 will be used as the modeling computer for the Mark IIIA Simcen.)

(5) An interactive alphanumeric television (IATV) system will be interfaced with buffered I/O channel 6 of the EMR 6050.

(6) The 7288 interface and the operations consoles on the programmed I/O of the EMR 6050 are being removed.

(7) Three serial receivers and two serial generators were designed and built, and will shortly be interfaced with the programmed I/O channel of the EMR 6050.

A brief description of each of the new capabilities is presented next.

a. Core memory expansion. This item is self-explanatory. Two additional 8k-word memory modules were added to the 16k words of memory already in the machine. All supporting circuitry already existed in the computer.

b. 6050/1108 interface. This is a device that operates in a half duplex mode. In the 6050 to 1108 direction, it accepts parallel data words from buffered I/O channel 3, adds control and parity bits to each character, and transmits the data to the 1108 via a 50-kbits/s modem system. Provisions are made so that the modem can be patched to either 1108A or B.

In the 1108 to 6050 direction, sync, parity and end-of-text characters are stripped from the output of the 50-kbits/s modern system. The data itself is blocked into 24-bit words which are input in parallel form to buffered I/O channel 3.

c. Buffered I/O serial generators. The serial generators on buffered I/O channels 4 and 5 are each already capable of operating at whatever rate, 50-kbits/s or less, the external clock provides. In normal operation, one of the generators (either one) will be operated at the 50-kbits/s rate while the other will be operated at the 4.8-kbits/s rate. The data output from these serial generators will be routed via the GCF to either the SFOF or the DSSs.

d. IATV system. This hardware is to be used as the control and display portion of the Mark IIIA Simcen. Physically, it consists of eight interactive keyboard/cathode ray tube terminals, and two high-speed printers, each interfaced through a multiplexer to buffered I/O channel 6 of the EMR 6050 computer. Each keyboard is capable of entering requests into the computer and each cathode ray tube and printer is capable of displaying whatever data was requested by its associated keyboard, through the buffered I/O channel.

e. Programmed I/O serial receivers and generators. These generators and receivers are designed to operate at a variety of data rates, selectable by the resident program or manually by each device. The intent is that they will be used to interface with the 4.8-kbits/s GCF modems during the *Mariner* Mars 1971 era. The receivers will be used to "look at" command data going to the DSSs and at monitor data being received from the DSS. The generators will, as described previously, output data via GCF hardware to either the SFOF or the DSSs.

D. Diagnostics for the SFOF Mark IIIA Central Processing System: Standalone Acceptance and Maintenance Routines, *R. A. Wells*

1. Philosophy

To verify proper performance of all elements in the SFOF central processing system equipment configuration and its interfaces with other facilities, it has become standard practice in previous developmental periods to produce and deliver a complete set of diagnostic software to permit the isolation and correction of hardware faults which might complicate or jeopardize the accomplishment of mission objectives. These routines are particularly essential where JPL-developed devices and linkages are involved and for which no vendor-furnished diagnostics are available. In these instances, it is required that comprehensive standalone diagnostics[1] be originated by JPL for the early stages of hardware checkout. Vendor diagnostics, often called customer engineering or CE diagnostics, are generally provided for testing of computer mainframes (arithmetic functions, memory, etc.) and on-line peripherals of their own manufacture. While these are also of a standalone nature, they are not treated herein since they are developed, maintained, and generally operated by vendor or contract maintenance personnel and are not compatible with SFOF software. Also of no concern here are wired-in equipment test features which require no special programming.

2. Objectives

In the process of equipment development, it becomes necessary to have some means to:

(1) Exercise the equipment under development.

(2) Act as a data generator/sensor for related equipment.

(3) Gather performance (i.e., error) statistics.

(4) Enable equipment acceptance and buyoff from the developer or vendor.

(5) Demonstrate equipment capabilities to users.

(6) Perform maintenance and exercise selected or suspected functions.

(7) Verify proper device or channel operation.

[1]Standalone diagnostics have been arbitrarily defined in the Mark III era as independently loadable routines not *primarily* under the control of the resident 360/75 diagnostic monitor.

Since it is usually impractical or uneconomic to tie up major portions of the facility for these purposes, independent standalones are the most suitable solution.

3. Background

In the Mark II era (7044/7094 computer strings), two families of diagnostics were produced:

(1) SYSCKO-INERT,[2] which required total dedication of the 7044 (and sometimes the 7094) for operation of its dependent routines.

(2) SYSCKO-LIVE,[2] which resided and executed compatibly with other jobs (such as mission data processing) under the 7044 redesign operating system, and required dedication of only those system elements under test.

In addition, many independent standalones were produced for equipment acceptance and other specialized purposes.

4. Mark IIIA Evolution

In the Mark IIIA environment, major emphasis has been placed on live diagnostics, the inert monitor concept has been superseded, and standalones have been confined to the following major areas:

(1) Mission display board (MDB).

(2) Digital TV (DTV).

(3) Simulation center (Simcen).

(4) High-speed data (HSD).

(5) Wideband data (WBD).

Figure 1 shows the equipment which can be tested, along with the hierarchal relationships of each standalone diagnostic. Table 1 gives the definitions of nomenclature used in Fig. 1.

5. Mission Display Board

Because of the limited memory of the MDB processor, its standalone diagnostics must be loaded for each occasion and displace the live operational software. Electromechanical responses of the VIGICON scribing projectors can be checked or, as a separate test, the integrity of the 50-kbits/s half-duplex ASCII serial channels to the 360/75s. (CE diagnostics test all standard DDP 416 devices and computer functions.)

[2]SYSCKO = system checkout.

Table 1. Nomenclature

BMXR	block multiplexer
CCTV	closed-circuit television
CHAN	channel
DIAG	diagnostic
DIS	data instrumentation subsystem
I/O	input/output
LDA	line driving amplifier
MUX	multiplexer
MX	multiplex
PCM	pulse-code modulated
PROJ	projector
RCVR	receiver
RT	real time
SCF	Scientific Computing Facility
TCP	telemetry and command processor
TTY	teletype
TVA	TV assembly
TWR	typewriter

6. Digital TV

These diagnostics are loaded into either CDC 3100 computer and exercise the DTV display generators, the display buffers for hard copy output, and the ability of the DTV copy request panels to select the desired output. If the 360-to-3100 data channel is to be tested, the 360 will be required to pass known data patterns across this one-way link in 8-bit parallel bytes. (CE diagnostics test all CDC 3100 functions and on-line CDC peripherals.)

7. Simulation Center

This family of diagnostics checks all JPL-unique attachments to the EMR 6050 (Fig. 1). The 50-kbits/s half-duplex serial channel from the EMR 6050 computer is tested by transmitting known data blocks to the Univac 1108 computer, which returns them unaltered to the Simcen for checking. Pulse-code-modulated serial generators and receivers can be tested within the Simcen by loop patching. The GCF diagnostic utilizes the support of a comparable 360 program to verify proper functioning of the HSD input and output equipment and lines.

8. High-Speed Data

This standalone, while requiring some services of the live 360/75 diagnostic monitor, provides more extensive capabilities than are to be found in the on-line HSD diagnostics. Specifically, in addition to passing 1200-bit data blocks at 4800 bits/s across the various HSD links, it can also gather error statistics for long-range analytical studies. It also acts as the generator/sensor element to support testing of individual GCF components and lines.

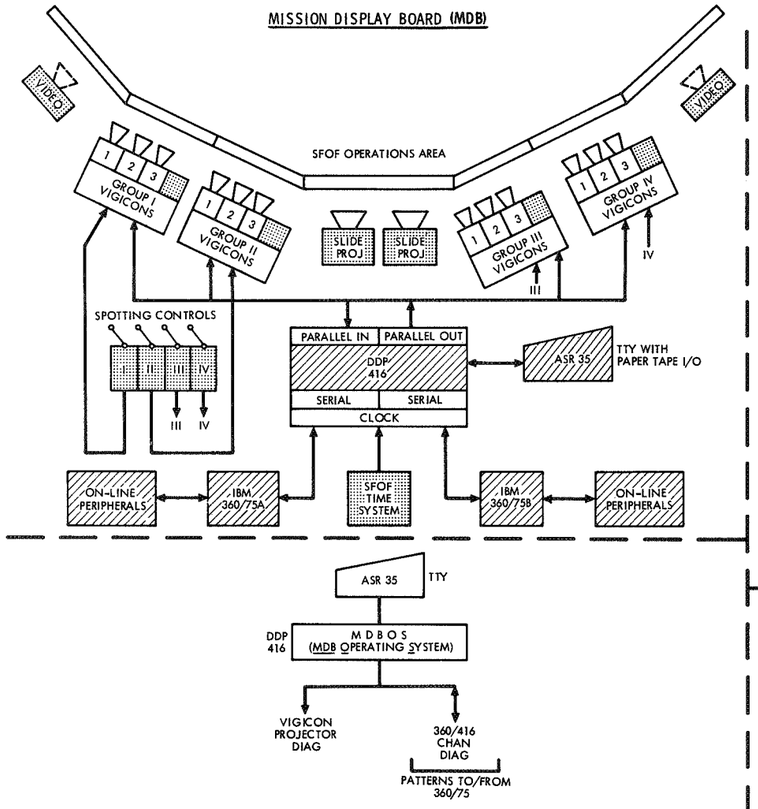
9. Wideband Data

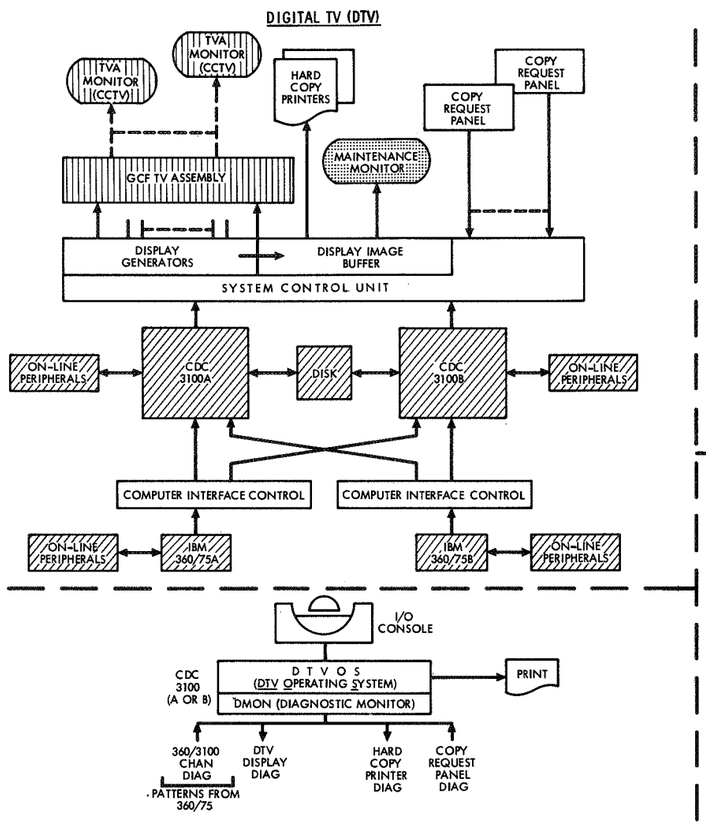
This test routine is essentially a 50-kbits/s version of the HSD standalone routine. Because of hardware limitations, only the WBD sensor program is employed to check incoming wideband data from the DSS 14 or CTA 21 telemetry and command processors. As with the HSD sensor diagnostic, error statistics may be obtained with the aid of this program.

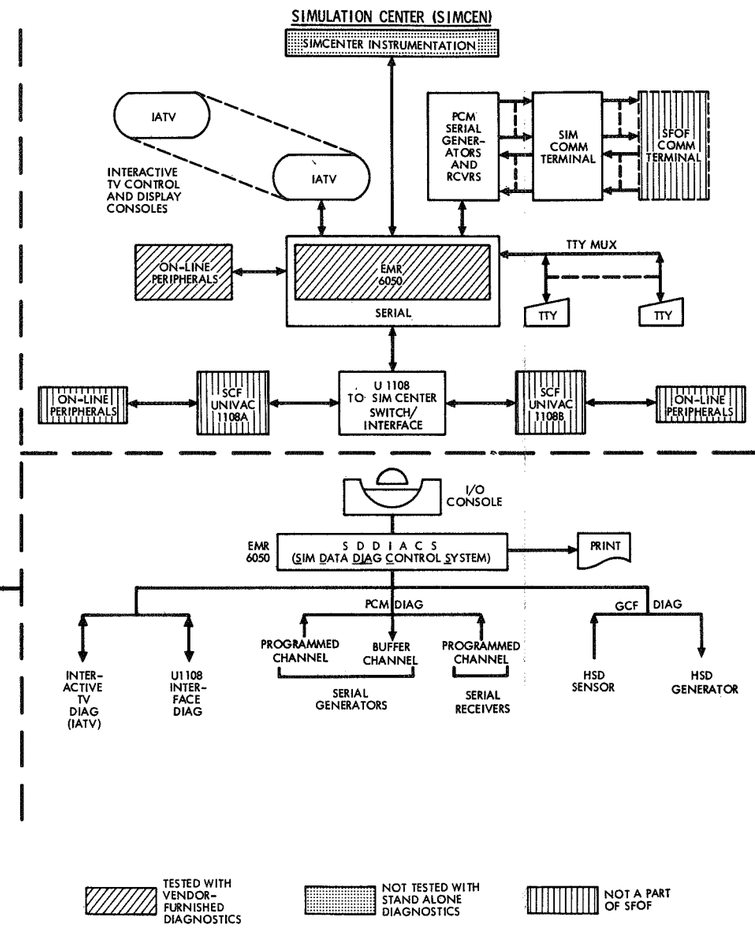
Since it is the intent to keep these standalone diagnostics as useful and comprehensive as possible, the cognizant hardware subsystem engineers seek to improve the programs as seems advantageous. For example, as performance idiosyncrasies of particular devices or channels are uncovered, the appropriate diagnostic is upgraded to exercise these potentially troublesome functions.

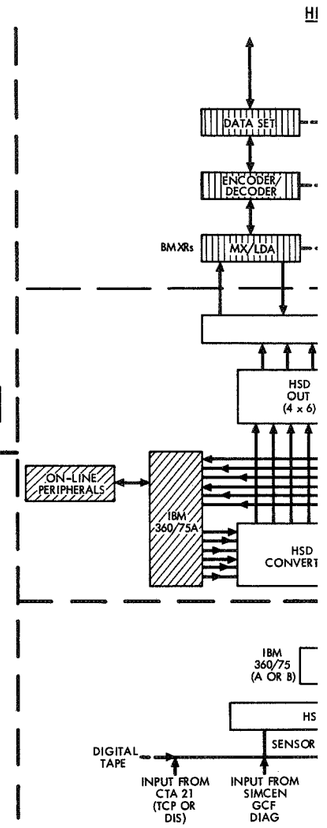
10. Conclusion

All of the standalone routines described are in use for SFOF subsystem and assembly checkout, except for the MDB diagnostics which remain to be completed. Utilization of these methods has enabled installation and acceptance to proceed with a minimum of competition for SFOF resources.









TESTED WITH VENDOR-FURNISHED DIAGNOSTICS

NOT TESTED WITH STAND ALONE DIAGNOSTICS

NOT A PART OF SFOF

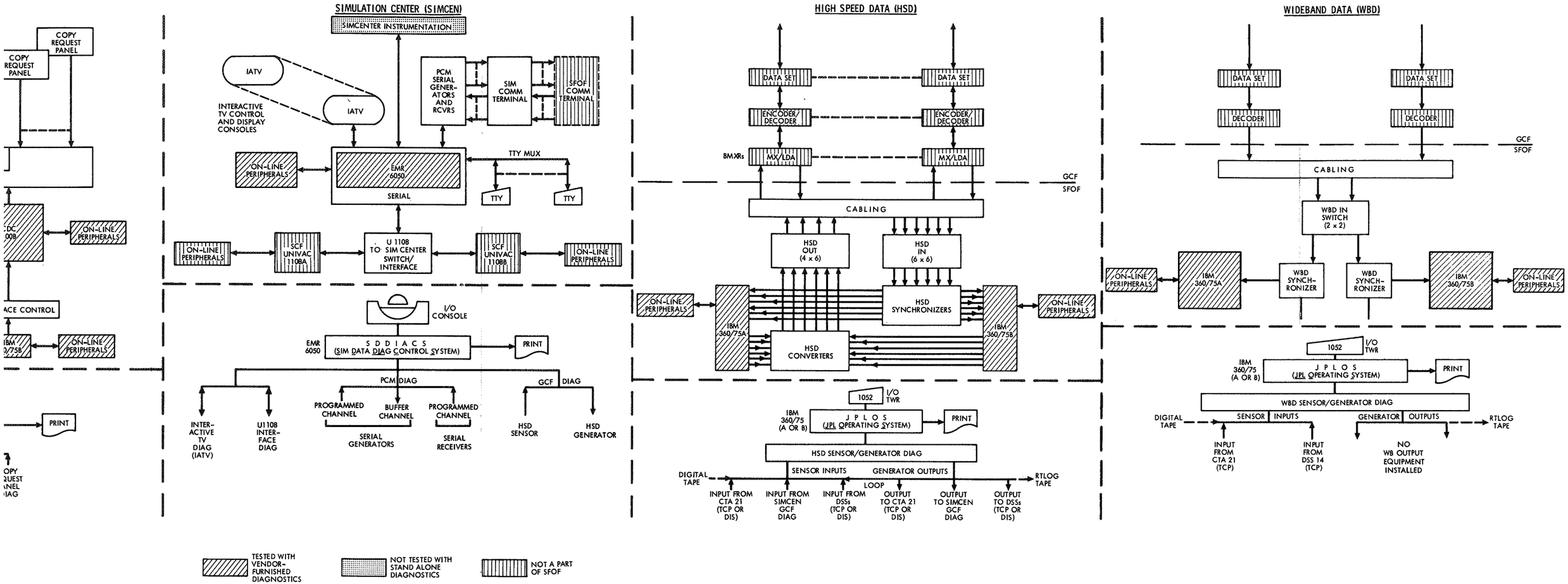


Fig. 1. SFOF and Simcen standalone diagnostic configurations

2

3

E. High-Speed Data, SFOF Outbound Communication, *P. G. Mullen*

As part of the SFOF Mark III configuration to support *Mariner* Mars 1971 and *Pioneer* missions, there will be high-speed data (HSD) transmission capability for both inbound and outbound communication. The outbound HSD data communication will be used to transmit command blocks, predicts, standards and limits, sequence of events data, and scheduling information. The HSD line rate is 4800 bits/s with data transfers, using the standard DSN 1200-bit blocks. The SFOF HSD block format consists of a header, the data body, and error detection sections.

The outbound data flow in the SFOF is from the central processing unit (CPU) to an IBM 2909 input/output channel to the HSD converters through a switcher to the GCF. In the GCF, the data blocks enter the block multiplexer, are sent to an encoder and then to the data set for transmission on the HSD lines. The CPU consists of two IBM 360/75 computers designated 75A and 75B, one acting as the primary computer and the other as an alternate or back-up computer. The data block header and data body will be originated here and transferred out of the computer through an IBM 2909 asynchronous data channel, general-purpose output (GPO) subchannel. The HSD converter accepts the 2909 GPO subchannel 32-bit parallel word and converts it to the serial format for transmission. The HSD converter must communicate and coordinate with both the 2909 and the block multiplexer to perform this task. The HSD outbound data switcher allows any one of the converters to interface with any one of the HSD lines. Thus, any 2909 GPO subchannel can output to any data line. The data block is then sent to the block multiplexer (BMXR) in the GCF. The BMXR interleaves the data blocks with filler blocks and thus provides a continuous outbound data stream. The encoder accepts the first 1164 bits of each block from the BMXR, either data or filler blocks, and adds the remaining data to complete the 1200-bit block. The complete data and filler blocks are then sent to the data set for transmission to the desired location. Among the stations that will receive outbound data are the DSIF stations, Boulder, Colorado, Ames Research Center, and CTA-21.

Present plans call for six HSD lines used for both inbound and outbound communication. Because the data sets used do not allow for splitting of the transmit and receive data lines, the six lines will provide a capability of six total simultaneous transmissions. The 2909 asynchronous data channels for 75A and 75B have a present capability of four HSD output subchannels each. The HSD converters will be designed to interface with either 2909 GPO subchannel and through the switcher can communicate with each of the HSD lines. Thus, six of these units will be built to handle the present requirements. The converters will be allocated as follows:

Converter	IBM 2909 GPO subchannel	IBM 360/75 computer
1	300	75A
	300	75B
2	310	75A
	310	75B
3	320	75A
4	330	75A
5	320	75B
6	330	75B

Either computer can transmit from four subchannels with the alternate computer, using the remaining two HSD lines if needed.

All units in the SFOF outbound communication link are commercially available items except the HSD output switcher and the HSD converters, which will be developed by JPL. The switcher will be a single unit capable of handling eight converter inputs and sending to eight HSD lines. A modular concept will be used in construction of the converters to allow for future expansion by addition of new units.

IX. GCF Development

DEVELOPMENT AND IMPLEMENTATION

A. GCF Wideband Digital Data System, *J. P. McClure*

In the fall of 1970, the GCF will install a 50-kbits/s wideband digital data capability between the SFOF, DSS 14, and CTA 21. These capabilities will first be used in support of the *Mariner* Mars 1971 mission.

Figure 1 delineates the major functional components used in this new GCF system. CTA 21 and DSS 14 will each be connected to the SFOF by a full duplex synchronous transmission capability which will permit 50-kbits/s data to be interchanged simultaneously in both directions between the SFOF and either field location.

In the inbound (station-to-SFOF) direction, data from two or more station telemetry and command processor computers will be merged, on a data block basis, by a block multiplexer. The merged data, plus any filler blocks necessary for synchronous operation, will be output by the block multiplexer at a 50-kbits/s rate to the error detection encoder. This encoder will append an error detection code to the end of each block before it is transmitted by the data set. The data set translates the digital signal to an audio form appropriate for transmission over 48-kHz microwave or cable facilities. In the Goldstone case, a regenerator at the DSCC area communications terminal will retime and clean up the signal prior to forwarding it toward the SFOF. In the SFOF, the signal will be received and demodulated to digital form by the data set. The error detection decoder will detect the presence of any transmission errors and will output the data to line drive amplifiers for the SFOF central processing system and the *Mariner* Mars 1971 mission test computer.

In the opposite direction, the SFOF wideband transmit terminal will accept blocked simulation data from the Simulation Center 6050 computer, process it to audio form, and route it to a patching capability in the basement of the SFOF. This signal can then be routed either to CTA 21 or DSS 14. The outbound path (SFOF-to-station) includes a demultiplexer at each field location. These devices will remove the filler blocks inserted by the block multiplexer.

Transmission between DSS 14 and the Goldstone DSCC Communications Center will be made via the existing Goldstone area microwave link. The signals will be relayed between the DSCC and the SFOF over modified channels of the Goldstone–SFOF microwave link provided by Western Union. Communications between CTA 21 and the SFOF will use existing shielded cables capable of 50-kbits/s operation.

The system is expected to yield an error rate of 3×10^{-5} or better on a 24-h basis. In the inbound direction, erroneous blocks sensed by the decoder will be tallied by an

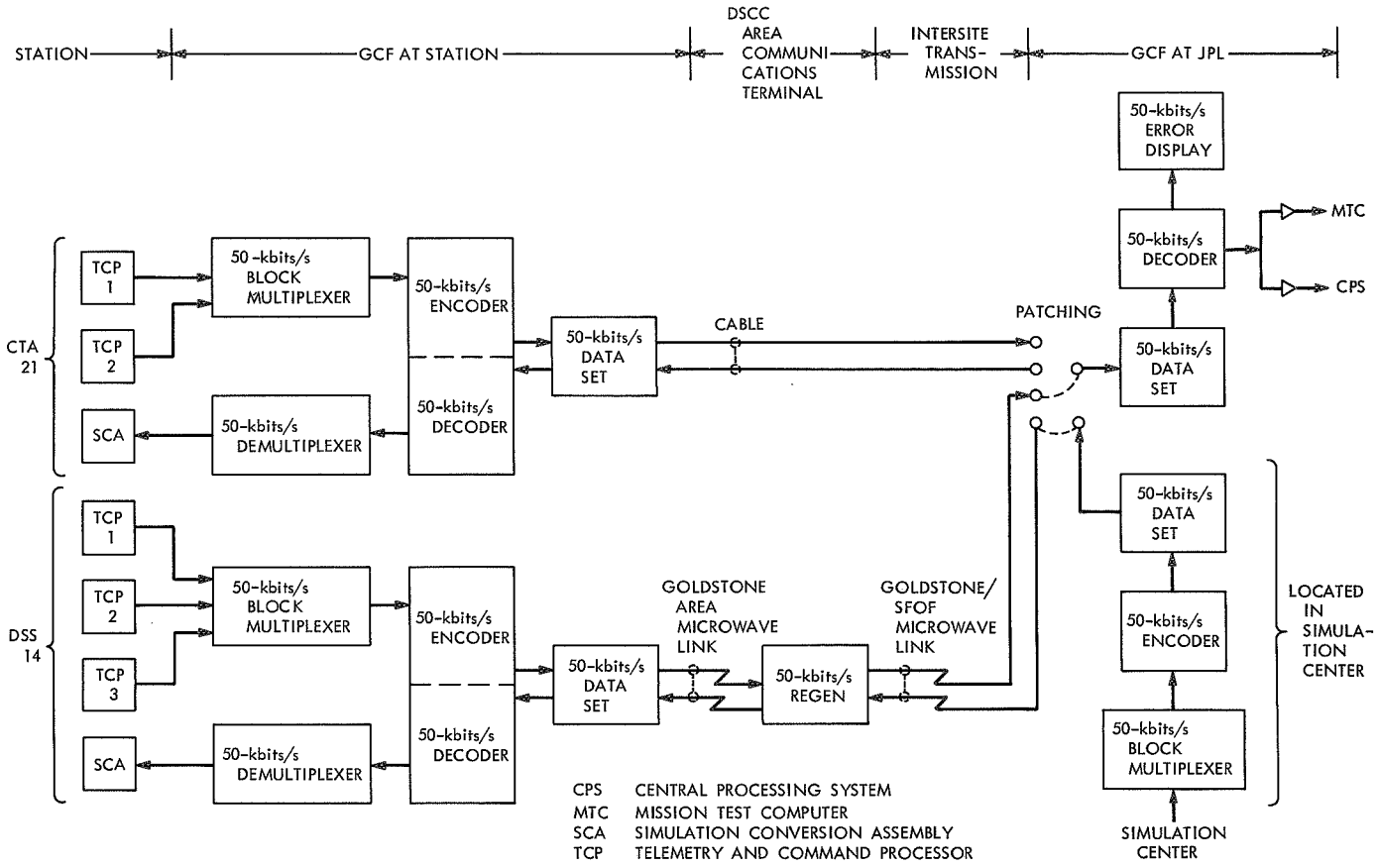


Fig. 1. Ground Communications Facility 1971–1972 wideband system

error display, thus permitting real-time quality evaluation. In the operationally critical DSS 14-to-SFOF link, dual terminal equipments and transmission links will be used to provide high service availability.

Inbound data (to the SFOF) will initially consist of one or two *Mariner* Mars 1971 digital video streams of 16 kbits/s or less each. Each stream will be output from an active telemetry and command processor. In the opposite direction, the Simulation Center will send simulated video data outbound to the station simulation conversion assembly where the blocked data will be transformed into spacecraft rate information and used as simulated data inputs at several locations within the station.

X. DSIF Development

DEVELOPMENT AND IMPLEMENTATION

A. CTA 21 to TRW Passive Microwave Link,

M. E. Wyatt

1. Introduction

To establish compatibility between the DSN and the various spacecraft prior to launch, Compatibility Test Area 21 (CTA 21) was implemented at JPL to represent a standard DSIF station. Via microwave link, CTA 21 can establish compatibility with a spacecraft in the Spacecraft Assembly Facility, the Environmental Test Laboratory, or the Space Simulator. However, the *Pioneer* Project has determined it impractical to bring the *Pioneer F* and *G* spacecraft to JPL for the necessary compatibility tests. Therefore, it has been proposed that an S-band microwave link be established between CTA 21 and TRW in Redondo Beach so that the compatibility tests may be integrated with the presently scheduled tests at TRW.

2. Plan

The plan at present is to establish a test link between JPL and TRW to determine if an operational S-band link is feasible over this path at the available signal levels. If the test link proves feasible, a permanent installation is planned.

When clearance to transmit S-band is received, it is planned to first set up an EMA910 field intensity receiver (borrowed from the JPL Electromagnetic Interference Group) on TRW Building M2 (2 story) with a tripod-mounted antenna to determine the signal and noise levels at the DSIF frequencies and to determine the amplitude modulation of a carrier transmitted from an HP8614 frequency generator with synthesizer and tripod-mounted antenna located at Henninger Flats in Angeles National Forest.

If preliminary tests indicate that the link direct to Building M2 is very unstable, the feasibility link will be set up with another relay on Building E2 (7 story), if TRW consents.

The feasibility link will consist of 4- or 6-ft antennas as follows: a 6-ft antenna at CTA 21 (existing), two back-to-back antennas at Henninger Flats, one antenna on M2 at TRW, and, if required, two antennas back-to-back on Building E2 at TRW. The exciter and receiver at CTA 21 will be used in conjunction with a test transponder to establish the phase and amplitude characteristics of the link. If determined feasible, the final link will consist of 6-ft dishes throughout.

3. Status

The task was assigned on June 23, 1970 and a clearance to transmit 2295 and 2110 MHz was requested along with a "use permit" to establish a relay site in the Angeles National Forest. A temporary use permit (effective until September 1, 1970) to set up a temporary relay site in the Angeles National Forest has been received. On July 30, the RF clearance was received but no testing has been accomplished.

Studies of topographic maps indicated several sites in the San Rafael Hills (south of JPL) and in the Angeles National Forest (east of JPL) could be used as relay points for one-bounce between JPL and TRW. However, the San Rafael Hills are private or city property and investigations have indicated that negotiations with the U. S. Forestry Service would be the most feasible and expeditious because JPL has secured "use permits" from the U. S. Forestry Service before.

Further investigations using helicopter and topographic maps indicate that the Henninger Flats helipad area would provide the optimum location because (1) it can be seen from JPL and TRW, (2) it is accessible by helicopter or car, (3) it extends the largest angle between the JPL mesa and CTA 21, and (4) it extends the largest angle between JPL and TRW.

B. Waveguide Switch Protector, *R. B. Kolbly*

The waveguide switches used by the Deep Space Instrumentation Facility are driven by a gear-reduced dc motor, with no power supplied to the motor armature with the switch in its operating position. If, for any reason, the motor operates with the switch rotor "hung up," the motor will destroy itself. A fuse is provided, but it is required to protect several switches, and does not reliably protect an individual switch motor.

In general, waveguide switch failures can be divided into two types:

(1) The switch rotor "hangs up" between positions, preventing the rotor from reaching its intended position. The motor continues to receive power, eventually burning it up.

(2) An internal relay or limit switch fails, allowing the motor to continue to operate, even though the switch rotor has reached a mechanical stop.

The common feature of these two types of failure is that the motor continues to draw power for an excessive period of time, until it destroys itself. The key to protection then is to sense when the motor is operating for an excessive period of time.

The circuit of this waveguide switch protector consists of a timing pulse generator, a counter, a current detector, and relay and alarm circuits. Figure 1 is a block diagram of the basic protection unit.

In this device, 60-Hz line voltage is converted to timing pulses by means of an *exclusive or* circuit utilizing a μL914 dual two-input gate. These pulses are applied to a $\div 256$ circuit (FF1–FF8). During the period when the waveguide switch is not being operated, the reset line of FF1–FF8 is high (approximately +3.6 V), preventing the flip-flops from counting. When the waveguide switch is activated, a sharp increase in voltage occurs at the output of the current-sensing transformer, which in turn is rectified and inverted by I1, causing the flip-flop reset lines to go low, allowing the counter chain to count the 60-Hz pulses. If the waveguide switch completes its cycles and the motor stops before FF8 is "set" (approximately two seconds), the output of the inverter goes high, forcing all flip-flops to set and stop the count, effectively returning the circuit to its starting point.

If the waveguide switch does not stop drawing current before FF8 is set, there is trouble within the switch. In this case, when FF8 is set, relay K1 is closed, giving an external alarm to the operator and removing power from the waveguide switch by means of slave relay K2. One set of contacts on K1 is used to hold K1 closed until the circuit is manually reset. This prevents the circuit from automatically recycling, which would set up an oscillation, causing eventual destruction of the switch motor.

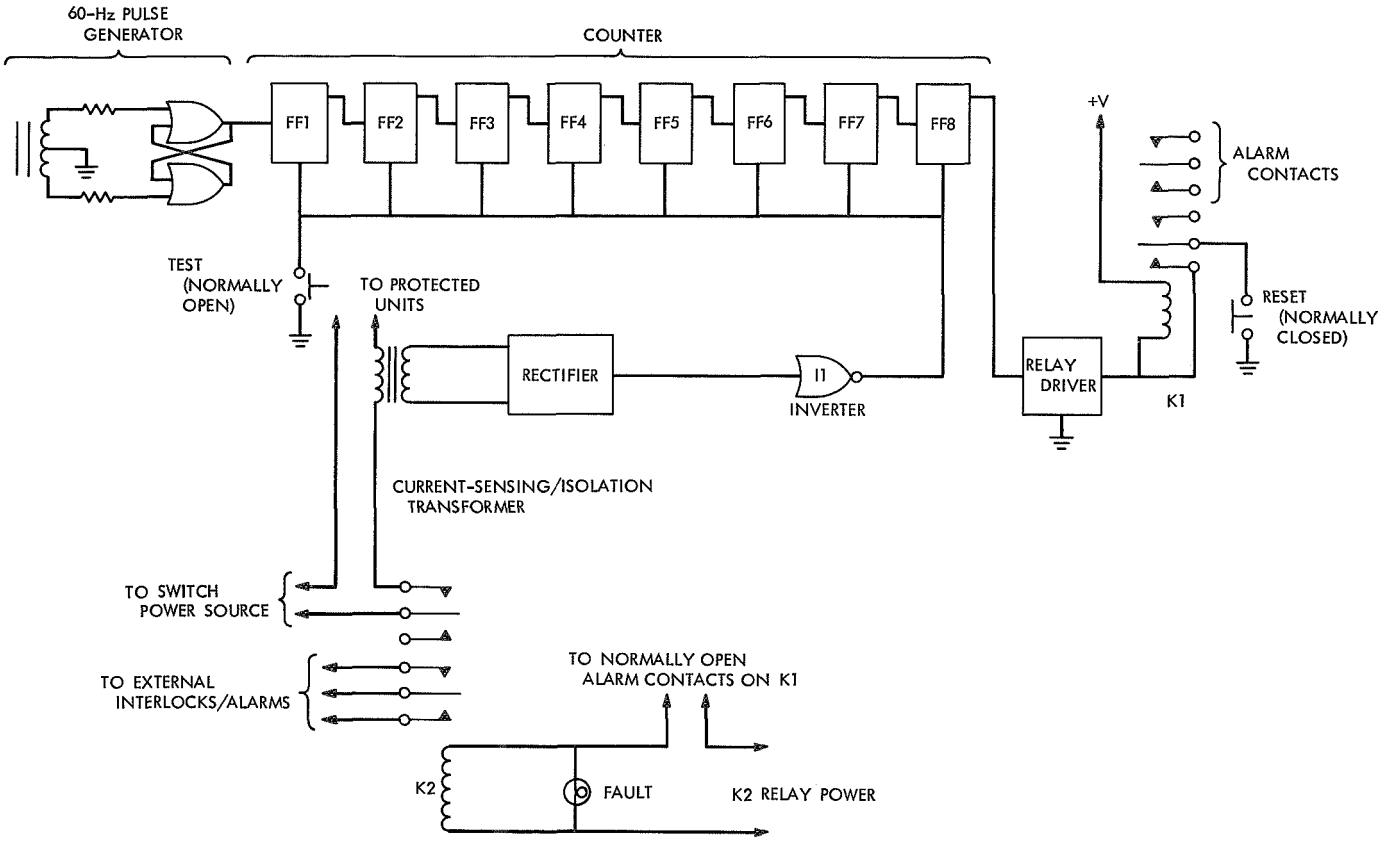


Fig. 1. Waveguide switch protector

A *test* function is provided to enable the operator to determine that the protective circuit is operating properly. A normally open push button is held closed for the two seconds or so required to set FF8 and activate the alarm. Momentary operation of the RESET push button returns the circuit to normal operation.

The time required to activate relay K1 can be changed by modifying the counter chain, or the repetition rate of the pulse generator, or both. By proper selection of the current-sensing transformer and inverter circuit constants, a number of waveguide switches can be protected by this circuit.

In addition to using this circuit to protect waveguide switches, it could be used to protect other types of intermittent-operation devices, such as machine-tool positioners, spacecraft antenna positioners, etc. Because it is digital in operation, this unit is not sensitive to variations in temperature and voltage. Also, it can be made very small and light, as well as low in cost.

C. Block IV Receiver Automatic Carrier Acquisition, R. C. Bunce

1. Introduction

A new method of automatically acquiring down-link (spacecraft) carriers for phase-lock receiver tracking loops has been designed in conjunction with the Block IV DSN Receiver Development Program. The method has been verified as operationally feasible during the prototype receiver checkout at DSS 13. At this station, automatic acquisition of the *Pioneer* and *Mariner* series spacecraft carriers has been achieved, with signal levels in the vicinity of -170 dBmW ± 5dB.

As a system function (see Fig. 1), the method combines the frequency sweep capabilities of the programmed local oscillator (PLO) assembly with an automatic acquisition detector. Acquisition search is conducted with the receiver in an open-loop mode, the input to the voltage-controlled oscillator (VCO) being shorted to ground to eliminate false locks. During search the PLO frequency sweeps the receiver local oscillator, tuning across the carrier uncertainty range until zero beat between the carrier and the local receiver reference frequency occurs. At this instant the acquisition detector generates an acquisition-trigger-at-zero-beat (ATZ) signal. The ATZ signal is fed to the PLO, causing it to discontinue sweeping and closes the receiver automatic phase control (APC) loop by removing the VCO ground. The loop then completes acquisition by automatically pulling in to phase lock.

The method and design have been optimized to give a maximum zero-beat signal-to-noise ratio in the acquisition detector channel.

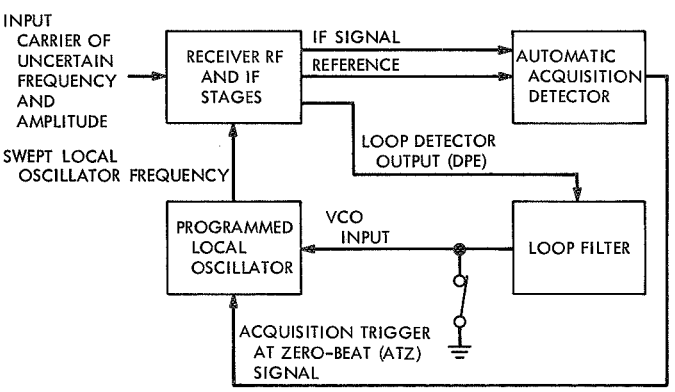
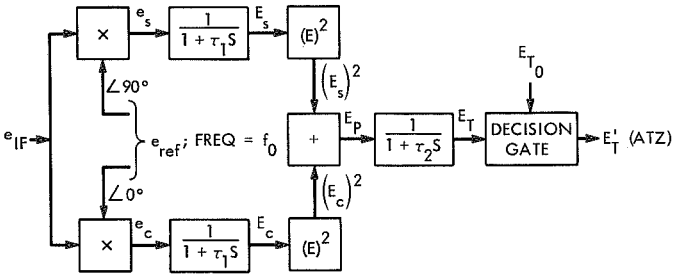


Fig. 1. Block IV automatic acquisition simplified block diagram

2. The Automatic Acquisition Detector

a. General description. The automatic acquisition detector is functionally diagrammed in Fig. 2. The IF signal is fed to the inputs of two identical phase detectors which are referenced in quadrature. The detectors act on the swept signal as frequency-differencing mixers, producing an output of linearly varying frequency that passes through zero beat as the swept signal frequency passes the receiver reference frequency. At zero beat, the multiplier outputs are voltages proportional to the arbitrary phase angle of the signal at that instant, which cannot be predicted. However, since the detectors are in quadrature, the phase difference between these two outputs (e_s and e_c) is 90 deg. (See Table 1 for definitions of symbols used in this article.) This quadrature is necessary to hold the trigger-pulse level constant. Errors in the 90-deg quadrature reference angle lead to amplitude variations in the trigger-pulse height; such errors must be minimized for stable and repeatable operation. Typical signal waveforms e_s and e_c for a linear



SIGNAL FREQUENCIES (HIGH-ORDER TERMS NEGLECTED):

FREQ $(e_{IF}) = 2\pi f_0 + 2\pi \dot{f}\tau$ (rad/sec); $-T \le \tau \le t$, $t = 0$ AT ZERO BEAT

FREQ $(e_{ref}) = 2\pi f_0$ (rad/sec)

FREQ $(e_s, e_c) = 2\pi \dot{f}\tau = \dot{\phi} = d\phi/d\tau$, $\dot{f}$ = SWEEP RATE, Hz/sec

PHASE (LINEAR SWEEP):

$$\text{PHASE } (e_s) = \int_{-T}^{t} \dot{\phi}(\tau)\, d\tau = \pi \dot{f} t^2 + \phi_0, \quad \left[\phi_0 = -\pi \dot{f} T^2 + \phi(-T)\right]$$

$$\text{PHASE } (e_c) = \int_{-T}^{t} \dot{\phi}(\tau)\, d\tau + \frac{\pi}{2} = \pi \dot{f} t^2 + \phi_0 + \frac{\pi}{2}$$

WAVEFORMS:

$$e_s = A \sin\left(\pi \dot{f} t^2 + \phi_0\right) \qquad e_c = A \cos\left(\pi \dot{f} t^2 + \phi_0\right)$$

$$E_{s,c} = \frac{1}{\tau_1} \int_{-T}^{t} \exp\left[\frac{-(t-\tau)}{\tau_1}\right] e_{s,c}(\tau)\, d\tau \qquad E_P = E_s^2 + E_c^2 \text{ [INDEPENDENT OF } \phi_0]$$

$$E_T = \frac{1}{\tau_2} \int_{-T}^{t} \exp\left[\frac{-(t-\tau)}{\tau_2}\right] E_P(\tau)\, d\tau \qquad E'_T = \begin{cases} 0, & E_T < E_{T_0} \\ 1, & E_T > E_{T_0} \end{cases}$$

Fig. 2. Automatic acquisition detector functional diagram

Table 1. Nomenclature

A	sine-wave signal peak voltage, V
A_M	signal level at minimum operating condition, V
A_0	design-point signal peak voltage (minimum), V
e_c	multiplier (0 deg or cosine ϕ) output, V
e_{IF}	IF signal (out of limiter), V rms
e_{ref}	multiplier reference signal, V rms
e_s	multiplier (90 deg or sine ϕ) output, V
E_s, E_c	parallel band-pass filter outputs (sine and cosine channels), V
E_P	summing-junction voltage, V (E_{P_S} = signal component when noise is considered; E_{P_N} = noise component)
E_T	post-filter output trigger pulse, V (E_{T_S} = signal component; E_{T_N} = noise component)
E_{T_0}	trigger level for acquisition decision, V
E'_T	decision gate output (binary state), V
f_0	multiplier reference (zero beat) frequency, Hz
$\dot{f}$	sweep rate of signal (around f_0), Hz/sec
$\dot{f}_0$	design value (usually minimum) of $\dot{f}$, Hz/sec
f, f'	frequency, Hz
ΔF_N	frequency error due to noise of trigger point, Hz
K	probability of acquisition, sigma units (normal curve of error integral)
N_0	noise spectral density (referred to receiver output), W/Hz
P	probability density function
P_A	probability of acquisition (zero to unity) per zero-beat pass
P_{TT}	probability of transient (false) triggering; fraction of time noise voltage dwells above E_{T_0}
R	ratio of τ_2/τ_1
t	time, sec
T	time of sweep preceding zero beat, sec
T_{TT}	time between transient triggering
X	independent variable, as defined in text
σ_N^2	variance of noise at summing junction (voltage due to noise power), V
$\sigma_{N_T}^2$	variance of noise at post-filter output, V
τ	normalized time t/τ_1 (dummy variable for t), sec/sec
τ_1	parallel filters time constant, sec
τ_2	post-filter time constant, sec
ϕ	phase angle, rad or deg (ϕ_0 = phase angle at zero beat)

frequency sweep are shown at the top of Fig. 3 for an arbitrary phase difference (ϕ_0) of 60 deg. Note the frequency variation around zero beat and the symmetry with respect to this point.

The effect of a filter on the waveforms e_s and e_c is not easily modeled mathematically; only a rather cumbersome series solution can be obtained in the general case. The function resembles the Fresnel integral, and its value at zero beat can be analytically determined, as shown later. The waveforms shown were machine-integrated, using an approximation to the convolution integral form, and the results were automatically plotted. The integral is noted in Fig. 2 and typical results (E_s and E_c) are shown on Fig. 3. Recorded waveforms during sweep are shown at the top of Fig. 4 for comparison.

When these waveforms are of the same amplitude and are squared and summed, theoretical independence of ϕ_0 is easily established. Upon squaring and summing, not only is the ϕ_0 effect eliminated but, as zero beat is approached, the voltage at the summing junction increases smoothly. This was first observed on a machine plot and later verified on the feasibility model (E_P, Figs. 3 and 4). The smoothness of E_P in the negative regions of t increases as the integrator gain decreases. The approach is far from monotonic when the integrator gain is infinite. Analytical determination of the effect of the integrator is seemingly impossible, since it involves approaching a finite time with negative and positive values from an infinite negative limit; even the series solution was not of much help. However, as a practical result, the method of pre-filtering, squaring, and summing of the signals produces a clean pulse with a smooth rise and rapid decay; the pulse amplitude at zero beat is some fraction of its peak height, and small variations in amplitude follow the peak.

Noise present on E_s and E_c is also detected as part of E_P, where it appears as an average dc value with residual noise variations around it. The zero beat signal level-to-dc noise component ratio was investigated (for fixed sweep rate $\dot{f}$) as a function of the filter time constant τ_1, and found to be optimum for an approximate value of $\tau_1 = [2\pi\dot{f}]^{-1/2}$. The noise variance is theoretically the same as the mean dc noise value. To reduce the probability of noise voltage peaks exceeding the trigger level of the decision gate, a post filter (with time constant τ_2) was added. This filter also has the effect of

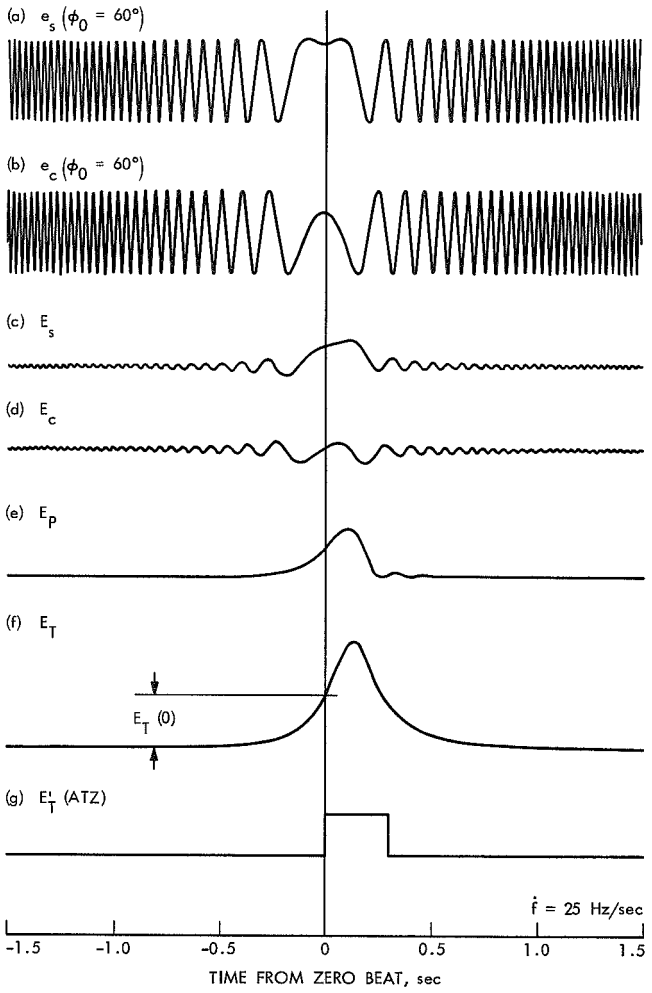


Fig. 3. Theoretical waveforms of acquisition detector

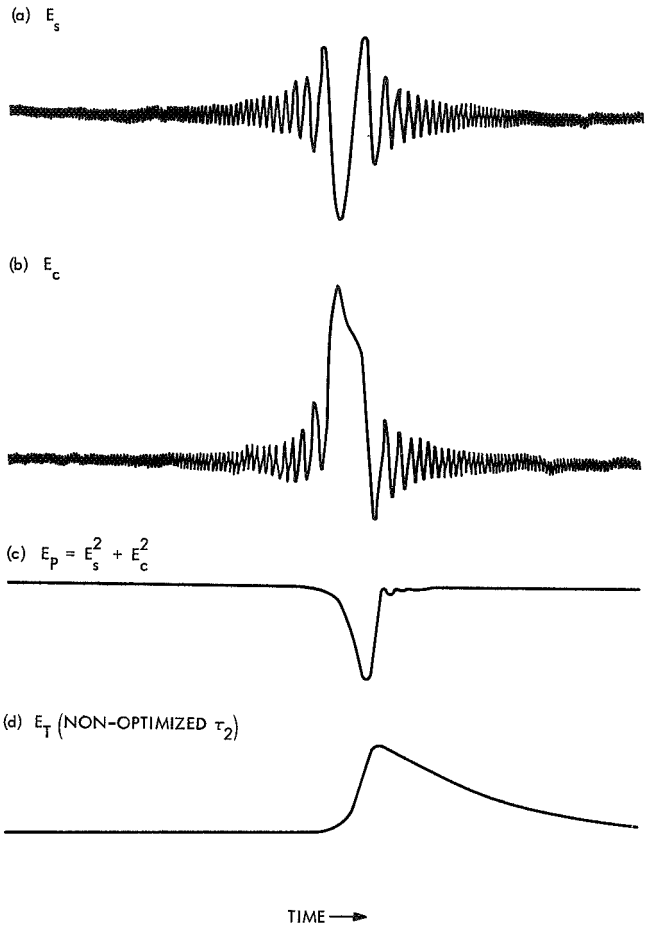


Fig. 4. Actual waveforms of acquisition detector

suppressing the amplitude of the E_P pulse, and its peak value-to-noise variance ratio was also investigated and found to optimize for a value of $\tau_2 \cong 1.68\ \tau_1$. The optimized form of E_T, the signal output of the post filter, is shown in Fig. 3 (machine plot).

The trigger level E_{T_0} is adjusted as a system procedure to equal the zero-beat amplitude of the E_T pulse at a minimum operating signal level (MOC), using a minimum sweep rate for this condition (covered later). At higher signal levels, the sweep rate is increased to suppress the zero-crossing level. In all cases, when the trigger level is exceeded, the decision gate changes state, producing a binary ATZ signal. When the detector is used as part of the acquisition system, the ATZ signal disables the local sweep and closes the receiver APC loop by removing the ground at the VCO input. The receiver then automatically phase locks to the incoming signal. During lock, the ATZ signal is normally a constant dc level, since a constant zero-beat signal is present. However, when the in-lock carrier is noisy, the detected pulse occasionally drops below the trigger level, causing momentary loss of the ATZ signal. A two-second delay is used within the PLO to prevent this type of ATZ loss from falsely initiating a search during in-lock operation; this is covered more fully later.

b. Analysis. The basic mathematical model for the detector channel is given in Fig. 2. A linear frequency sweep and sine-wave signal are assumed in this analysis; higher-order sweep rate terms and carrier frequency jitter that may be present in practice are neglected. The integrals, in general, have no analytic solution and machine integration was used to obtain numerical results and automatically plot the waveforms which led to the optimization of the detector. Most waveforms were functionally verified by comparing with those obtained from operating equipment. The notations on Fig. 2 refer to signal only; when noise is considered in the following expressions, suitable subscripts are used.

The optimizing design criteria and parameter limits are given in expressions (8), (14), (16), (17), and on Figure 11. A list of nomenclature terms is given in Table 1. All gain factors in the following expressions have been normalized to unity.

(1) *Optimization of* τ_1. The summing junction signal voltage is

$$E_{P_S}(t) = \left[\frac{A}{\tau_1}\int_{-T}^{t} e^{-(t-\tau)/\tau_1}\sin(\pi\dot{f}\tau^2+\phi_0)\,d\tau\right]^2 + \left[\frac{A}{\tau_1}\int_{-T}^{t} e^{-(t-\tau)/\tau_1}\cos(\pi\dot{f}\tau^2+\phi_0)\,d\tau\right]^2 \quad (1)$$

and the summing junction noise voltage is

$$E_{P_N}(t) = 2N_0\int_0^\infty \frac{1}{1+(2\pi f\tau_1)^2}\,df + V_n(t,N_0)$$
$$= 2N_0\left|\frac{1}{2\pi\tau_1}\tan^{-1}(2\pi f\tau_1)\right|_0^\infty + V_n(t,N_0)$$
$$= \frac{N_0}{2\tau_1} + V_n(t,N_0) \quad (2)$$

The noise voltage consists of a dc component and time-variable part (N_0 assumed white and gaussian). For large T, at zero beat, expression (1) has the particular solution:

$$\lim_{T\to\infty} E_{P_S}(0) = \pi A^2\left[\frac{1}{2\tau_1\sqrt{\pi\dot{f}}}\right]^2\left[1-2(S_a+C_a)+2(S_a^2+C_a^2)\right]$$
$$S_a = \sqrt{\frac{2}{\pi}}\int_0^{1/2\tau_1\sqrt{\pi\dot{f}}}\sin(t^2)\,dt \qquad C_a = \sqrt{\frac{2}{\pi}}\int_0^{1/2\tau_1\sqrt{\pi\dot{f}}}\cos(t^2)\,dt \quad (3)$$

The summing junction zero-beat signal-to-noise dc component is

$$|S/N|_{dc\Sigma} = \frac{\lim_{T\to\infty} E_{P_S}(0)}{\frac{N_0}{2\tau_1}} \quad (4)$$

For fixed $\dot{f}$, this varies with τ_1 according to

$$|S/N|_{dc\Sigma} = \frac{A^2}{N_0}\sqrt{\frac{\pi}{\dot{f}}}(X)\left[1-2(S_a+C_a)+2(S_a^2+C_a^2)\right]$$
$$X = \frac{1}{2\tau_1\sqrt{\pi\dot{f}}}$$
$$S_a = \sqrt{\frac{2}{\pi}}\int_0^X \sin(t^2)\,dt \qquad C_a = \sqrt{\frac{2}{\pi}}\int_0^X \cos(t^2)\,dt \quad (5)$$

By numerical integration, expression (5) was found to have a maximum value at

$$X(|S/N|_{dc\Sigma}\,\max) \approx \frac{1}{\sqrt{2}} \text{ (see Fig. 5)} \quad (6)$$

yielding optimum conditions at summing junction:

$$|S/N|_{dc\Sigma}\,\max = \left(\frac{A^2}{N_0}\right)\left(\sqrt{\frac{\pi}{\dot{f}_0}}\right)(0.238)$$
$$= 0.422\frac{A^2}{N_0\sqrt{\dot{f}_0}} \quad (7)$$

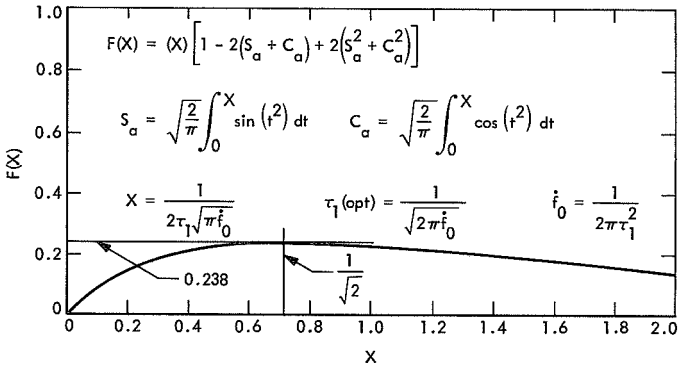


Fig. 5. Optimization of τ_1 for constant $\dot{f}$

and

$$\tau_1\,(\text{optimum}) = \frac{1}{\sqrt{2\pi\dot{f}_0}} \tag{8}$$

where $\dot{f}_0$ is the selected sweep rate for which τ_1 is optimized. The detected pulse shape associated with this optimization is shown on Fig. 6.

(2) Optimization of τ_2. The time-variable component of noise at the summing junction has a spectral density obtained by convoluting the filter output spectral density with itself:

$$S_{V_N}(f) = 4N_0 \int_0^{+\infty} \frac{1}{1+[2\pi\tau_1 f']^2} \times \frac{1}{1+[2\pi\tau_1(f-f')]^2}\,df'$$

$$= \frac{N_0}{2\tau_1} \times \frac{1}{1+(\pi\tau_1 f)^2} \tag{9}$$

The variance of the variable noise component at the post-filter output is therefore:

$$\sigma^2_{N_T} = \frac{N_0}{2\tau_1}\int_0^{\infty} \frac{1}{1+(\pi\tau_1 f)^2} \times \frac{1}{1+(2\pi\tau_2 f)^2}\,df$$

$$= \frac{N_0}{(2\tau_1)^2}\left[\frac{1}{1+\dfrac{2\tau_2}{\tau_1}}\right] \tag{10}$$

The signal at the post-filter output, as the filter time constant varies, can be expressed as (at $\dot{f} = \dot{f}_0$):

$$\left.\begin{aligned} E_{T_S}(\tau) &= \frac{A^2}{R}\int_{-\infty}^{\tau} e^{-(\tau-T)/R}\,E_P(T)\,dT \\ \tau &= t/\tau_1 = \sqrt{2\pi f_0}\,t \qquad R = \tau_2/\tau_1 \\ E_P(T) &= \left[\int_{-\infty}^{T} e^{-(T-\tau)} \sin\left(\frac{\tau^2}{2}\right)d\tau\right]^2 + \left[\int_{-\infty}^{T} e^{-(T-\tau)} \cos\left(\frac{\tau^2}{2}\right)d\tau\right]^2 \end{aligned}\right\} \tag{11}$$

The expression (11) was numerically approximated by machine for various ratios τ_2/τ_1; typical results are shown on Fig. 7. The peak results were divided by expression (10) to yield:

$$\left|S/N\left(\frac{\tau_2}{\tau_1}\right)\right| E_T(\text{PK}) = \frac{[E_{T_S}(\text{PK})]^2}{N_0} \times (2\tau_1)^2 \times \left[1+\frac{2\tau_2}{\tau_1}\right] \tag{12}$$

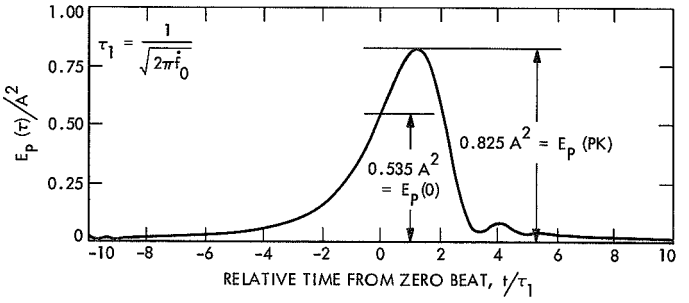


Fig. 6. Optimum detected pulse E_{P_S}

The results, in dB enhancement by the post filter, are shown on Fig. 8. This gave the optimum ratio:

$$R_{\text{opt}} = \frac{\tau_2}{\tau_1}\,(\text{opt}) \approx 1.68 \tag{13}$$

or

$$\tau_2\,(\text{optimum}) \approx 1.68\,\tau_1 \tag{14}$$

The pulse shape is plotted on Fig. 9.

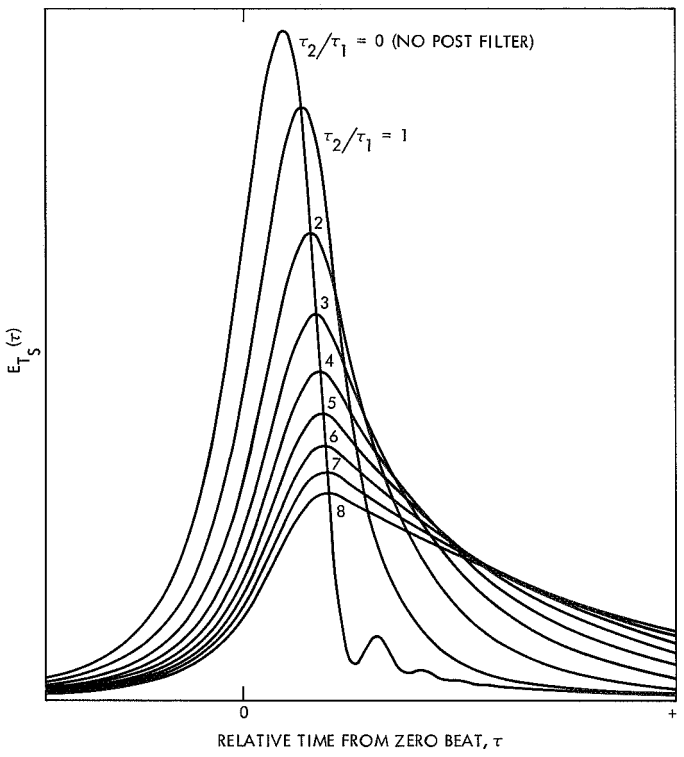


Fig. 7. Effect of post-filter τ_2 on trigger pulse shape

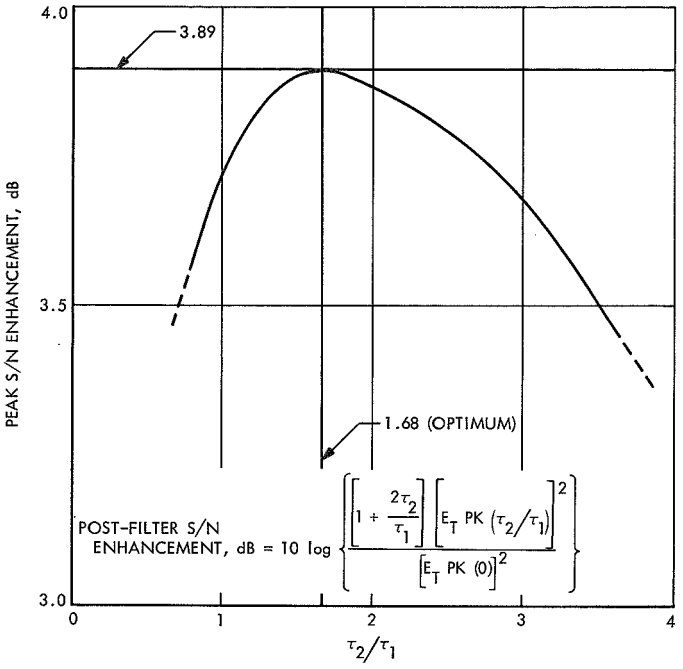


Fig. 8. S/N enhancement as a function of τ_2/τ_1

3. Choice of Minimum Sweep Rate and Trigger level

Assuming the acquisition system is associated with a receiver reference loop of design-point bandwidth $2B_{L_0}$, then the mean-peak-trigger-pulse-signal-voltage-to-noise-dc-voltage ratio at the receiver loop design point will be (from expression 2 and Fig. 9), gain normalized:

$$E_{T_N}(\text{dc}) = \text{noise dc} = \frac{N_0}{2\tau_1}$$

$$\text{signal (design pt)} = N_0(2B_{L_0}) = \frac{A_0^2}{2}$$

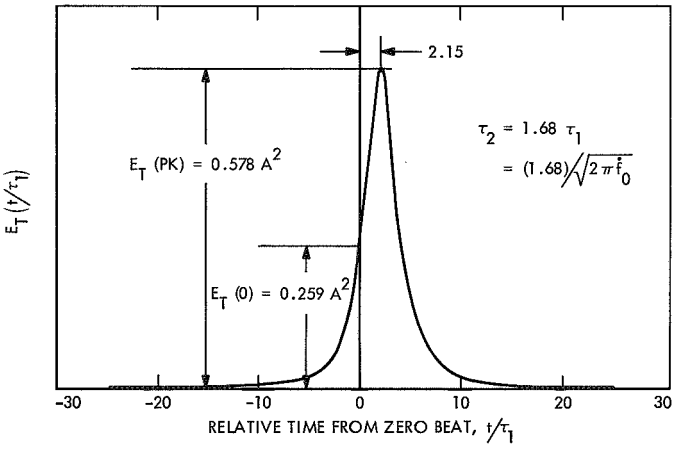


Fig. 9. Optimum trigger pulse E_{T_S}

$$\frac{\text{mean peak signal}}{\text{noise dc}} = \frac{0.578\,A^2}{\frac{N_0}{2\tau_1}} = \frac{1.156\,N_0(2B_{L_0})}{N_0} \times (2\tau_1)$$

$$= 1.156(2B_{L_0}) \times \frac{2}{\sqrt{2\pi \dot{f}_0}} = 0.922\left[\frac{2B_{L_0}}{\sqrt{\dot{f}_0}}\right] \qquad (15)$$

Choosing this ratio as *unity* (compatible with main loop design point criteria) yields

$$\dot{f}_0 \approx 0.850\,(2B_{L_0})^2 \text{ Hz/sec} \tag{16}$$

The optimum setting for the trigger level is a function of the signal level and sweep rate; it is obvious that the least acquisition error will occur if the trigger-pulse amplitude (nominal) crosses the trigger level at zero beat. If the sweep rate of expression (16) is used at an arbitrary low signal level, the correct setting is

$$\frac{E_{T_0}}{E_{T_N}} = 1 + 0.448\left(\frac{A_M}{A_0}\right)^2 \tag{17}$$

where E_{T_N} is the dc noise component at the post-filter output in the absence of signal, and A_M is an arbitrary minimum-operating-condition (MOC) signal level discussed more fully later. At the MOC, acquisition trigger-point frequency error is a function of the signal level uncertainty and the amplitude displacement due to noise. At signal levels above the MOC, the sweep rate is increased to suppress the zero-crossing level. Machine integration yielded the suppression shown on Fig. 10, which is expressed in dB on Fig. 11. In the Block IV equipment, the MOC is set at $A_M = \sqrt{10}\ A_0$ (10-dB margin). Above this level the sweep rate is increased to compensate for the signal level increase corresponding to the suppression factor given on Fig. 11. Limiter action restricts the maximum sweep rate to about:

$$\dot{f} \text{ max (strong signal)} \approx 20\,\dot{f}_0 \tag{18}$$

corresponding to a total signal level range of 23 dB at the limited output.

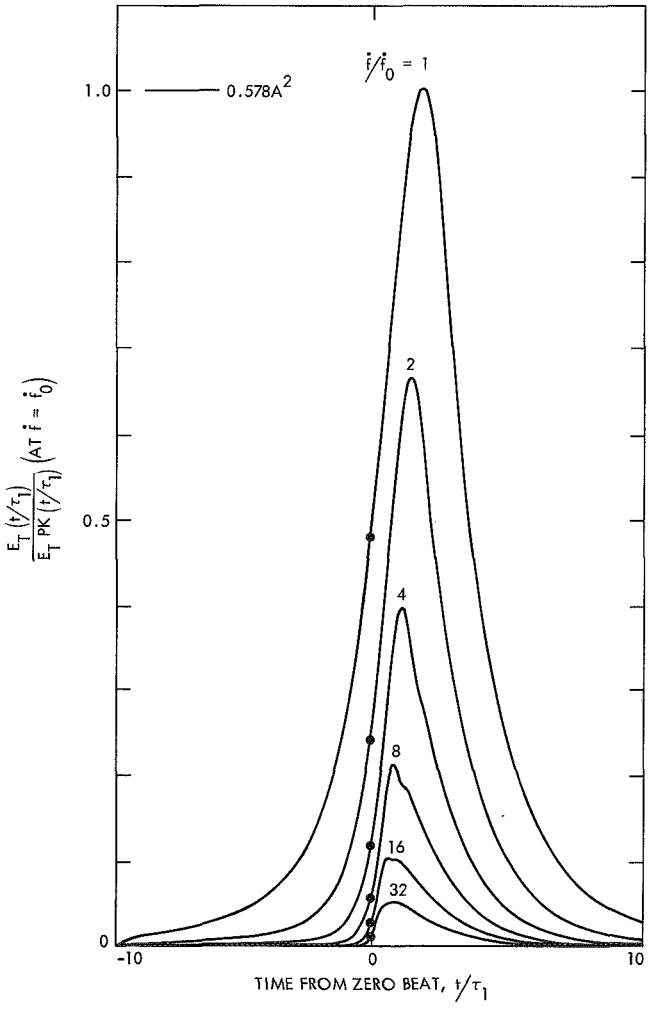


Fig. 10. Trigger pulse suppression as a function of sweep rate

4. Acquisition Probabilities and Trigger Error

The probability of acquisition is directly related to the probability of the pulse (Fig. 9) exceeding the trigger level E_{T_0} as the signal passes through zero beat. This does not imply that this event will occur at $t = 0$; the acquisition decision will, in general, occur with a frequency error associated with the precise time at which the pulse amplitude, perturbed by noise and

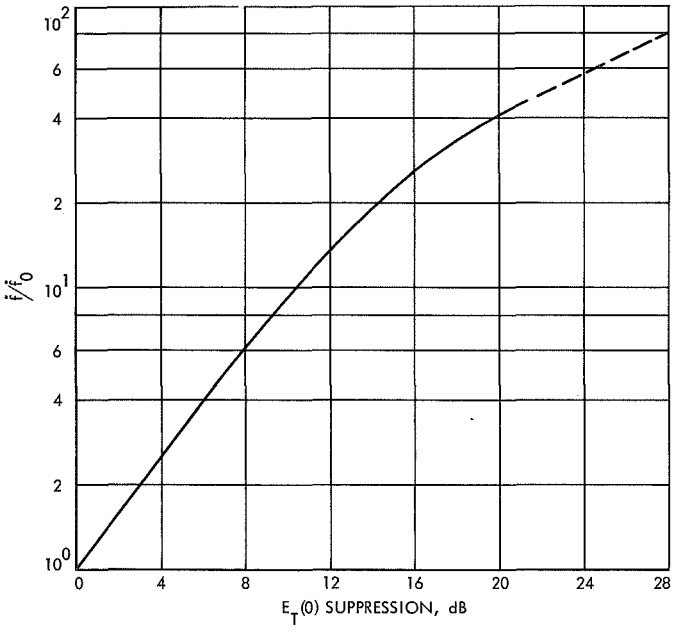


Fig. 11. Zero-crossing amplitude suppression as a function of sweep rate

signal level uncertainty, crosses the trigger level. However, as long as this error does not exceed the acquisition bandwidth of the main receiver loop, acquisition will occur.

A full statistical study of the subject amplitude has not been attempted; however, a lower bound on this probability can be estimated for the signal level region in which the trigger level is set based on the minimum expected signal level:

$$\text{Probability of acquisition (min)} = P_A = \frac{1}{2} + \text{erf}\,[K_1]$$

$$K_1 \approx \left[1 - \sqrt{\frac{0.259}{0.578}}\right]\frac{A}{A_0} \tag{19}$$

This probability is the major bound on the arbitrary MOC selection. At $A = A_0$, this is about 0.71/pass; at $A = \sqrt{10}\ A_0$, the selected Block IV MOC, this is about 0.96/pass. These figures assume that the pulse amplitude variance is gaussian and equal to that of the noise at $A = A_0$. The actual probability undoubtedly exceeds this, since the pulse variance occurs in the entire region of zero beat, not simply at the instant of peak amplitude, as expression (19) implies. Under stronger signal conditions, even if sweep rate is increased to suppress the pulse amplitude, P_A will continue to increase as the signal suppression factor increases, decreasing the relative signal variance.

The probability of a transient ATZ signal occuring is the same as the probalility of a noise spike exceeding the trigger level and will cause a short false decision period in the absence of signal. The moments of the probability density function (PDF) at the post-filter output are very cumbersome to calculate, and this was not attempted. The PDF at the summing junction, however, can be shown to be exponential (variable noise component only):

$$\left.\begin{aligned} P(E_{P_N}) &= \frac{1}{(\sigma_N)^2}\exp\left[-\frac{E_{P_N}}{(\sigma_N)^2}\right] \\ \sigma_N^2 &= \frac{N_0}{(2\tau_1)^2} \end{aligned}\right\} \tag{20}$$

This was roughly confirmed by sample data taken on the feasibility model. The optimized ratio of expression (12) results in a variance reduction of

$$(\sigma_{N_T})^2 = \left[1 + \frac{2\tau_2}{\tau_1}\right]^{-1}(\sigma_N)^2 = 0.229\,(\sigma_N)^2 \tag{21}$$

The PDF of this function tends away from the exponential toward the Rayleigh (Ref. 1); however, solutions available in the literature are only approximate, and most treat only the signal-plus-noise case. A rough approximation can be obtained, however, by simply substituting σ_{N_T} of expression (21) into expression (20) and integrating to obtain the approximate cumulative function:

$$P\,(E_{T_N} > E_{T_0}) \approx 1 - \int_0^{E_{T_0}} \frac{1}{0.229(\sigma_N)^2}\exp\left[-\frac{X}{0.229(\sigma_N)^2}\right]dX$$

$$\approx \exp\left[-\frac{E_{T_0}}{0.229(\sigma_N)^2}\right] \approx P_{TT} \tag{22}$$

The probability of transient (false) triggering, P_{TT}, must be interpreted carefully; it simply represents an estimate of the fraction of time that the noise voltage exceeds E_{T_0}, and is not in itself a "mean time between transients." This can be estimated by imagining a sample rate at twice the post-filter bandwidth; that is,

$$\overline{T_{TT}} \approx 4\exp\left[\frac{E_{T_0}}{0.229(\sigma_N)^2}\right]\tau_2, \qquad (\sigma_N)^2 = E_{T_N}$$

or, using expressions (17), (16), (8), and (14),

$$\overline{T_{TT}} \approx \frac{2.91}{(2B_{L_0})}\exp\left[4.37 + 1.96\left(\frac{A}{A_0}\right)\right]^2 \text{ seconds}$$

$$= \text{estimated mean time between transient triggering} \tag{23}$$

At design point ($A = A_0$), expression (23) is

$$\overline{T_{TT}}\,(\text{min}) \approx \frac{1633}{2B_{L_0}} \text{ seconds} \tag{24}$$

In the Block IV system, this is a minimum of 16 seconds for the $2B_{L_0}$ max (design value) of 100 Hz. Expression (23) clearly becomes large enough to be disregarded at $(A/A_0)^2 = 10$, the "MOC." At this point,

$$\overline{T_{TT}} \approx \left[\frac{7.5 \times 10^{10}}{2B_{L_0}}\right] \text{seconds} \tag{25}$$

Table 2. Block IV R&D receiver automatic acquisition design parameters

Predetection Bandwidth B_G, Hz	Closed-loop Bandwidth $2B_{L_0}$, Hz	Minimum sweep rate $\dot{f}_0$ at MOC, Hz/sec	Maximum sweep rate $\dot{f}$ at strong signal, Hz/sec	τ_1, sec	τ_2, sec
200	1.0 (3.0)	0.85	16.9	0.433	0.727
2000	10.0 (30.0)	85	1690	0.0433	0.0727
20000	100.0 (300.0)	8500	169000	0.00433	0.00727

Frequency errors at the trigger instant occur from two sources: (1) displacement of $E_T(0)$ by noise perturbations, and (2) by errors in the estimate of the signal amplitude A (see Fig. 9). The order of magnitude of the first is that resulting from the peak displacement in time; that is, at the MOC:

$$\Delta F_N\,(\text{order of magnitude}) = 0.923\,(2B_{L_0}) \qquad (26)$$

This is at least an order of magnitude less than the acquisition bandwidth of loops in the receiver under design; additional displacements due to errors in signal amplitude estimates would thus have to be very large to prevent acquisition after triggering. Statistical measurements on the feasibility model showed that the error stated above is apparently normally distributed, and that expression (26) represents about 2 standard deviations of the error distribution.

5. Design Values and Performance of the Block IV R&D Receiver

The Block IV R&D receiver has six closed-loop design point bandwidths with three associated predetection filters, one for each two closed-loop bandwidths. To simplify the design of the automatic acquisition detector, three sets of time constants and associated gain changes were used, one set for each predetection filter. The receiver bandwidths used for the design are the narrower of the two possible in each case to assure acquisition in both modes. Final figures are summarized in Table 2. All spacecraft carrier acquisition to date has been done in the 3-Hz mode, using the figures in the first row of the table. All conditions shown on the table are, however, selectable in the hardware.

Actual recommended sweep rates at the PLO (1/24 of S-band rate) in use at DSS 13 in the narrow-band mode and for the stated MOC, using a value of system temperature of 20°K, are given in Table 3 as a function of carrier level. The rates were derived generally from Fig. 11 and are added to the signal doppler rate to achieve minimum trigger frequency error. The signal level range covers the expected signal levels of the *Mariner* and *Pioneer* spacecraft.

Table 3. Acquisition sweep rate as a function of expected signal level (Block IV receiver, DSS 13)

Signal level, dBmW	Sweep rate at PLO, Hz/sec	Sweep rate at S-band, Hz/sec
−175	0.04	0.85
−174	0.04	0.9772
−173	0.05	1.2410
−172	0.07	1.6358
−171	0.09	2.1544
−170	0.12	2.7877
−169	0.15	3.5243
−168	0.18	4.3508
−167	0.22	5.2510
−166	0.26	6.2068
−165	0.30	7.1979

The Block IV receiver system does not achieve the low probability of expression (25) for an MOC of 10 dB above the 1-Hz design point. An occasional false ATZ signal occurs for reasons not fully understood. The period implied by expression (25) was, however, achieved on the feasibility model evaluated in the laboratory. The causes of the periodic false triggering in the Block IV receiver are presently under investigation. The problem is not serious since a transient triggering, not followed by receiver lock, merely causes a two-second delay in acquisition.

An additional problem arose while acquiring the *Pioneer IX* spacecraft. During this acquisition, ATZ instability led to the need for a manual over-ride procedure to maintain phase lock after acquisition. The instability was not unexpected; the spacecraft signal level was near the MOC and probably had a somewhat broad spectrum. However, to assure stable operation under such conditions, the modification that added a two-second delay within the PLO was implemented. This modification delays the initiation of PLO sweep and loop-opening for two seconds after loss of the ATZ signal. If the ATZ signal reoccurs during the two-second interval, lock is not interrupted. The interval represents the approximate maximum pull-in time for a 1-Hz loop when the frequency error is within the pass-band of the acquisition detector. The delay is, in effect, additional filtering required in the vicinity of zero beat to assure stable acquisition for weak signals. The modification also causes a momentary delay during search in the event that a noise pulse exceeds the ATZ trigger level. The only effect in this case is a longer acquisition period, which is probably increased by not more than 10%.

In summary, the Block IV open-loop acquisition method appears to be a distinct improvement over the Block III design since it avoids inherent false-lock conditions and gives promise of performing automatically over the entire signal level dynamic range. It is not theoretically as rapid as a fully computerized process (such as automatic spectral analysis by fast-fourier transform) might prove to be; however, it is considerably less complex and expensive than the computer approach.

Reference

1. Davenport, W. B., and Root, W. L., *Random Signals and Noise*, p. 199. McGraw-Hill Book Co., Inc., New York, 1958.

D. The Recording of SDA Outputs for *Mariner* Mars 1971, *G. Hamilton*

In the event of failure in the data path after the subcarrier demodulator assembly (SDA), in such a manner that the telemetry and command processor (TCP) tapes do not furnish all required data, it is desirable that the output of the SDA be recorded for possible playback at a later time and insertion into the symbol synchronizer assembly (SSA) for recovery of data at the station. It is a requirement of *Mariner* Mars 1971 that, for this purpose, data be recorded prior to integration. It is also necessary to make provisions for recording two channels (engineering and science) simultaneously for each of two spacecraft. In addition, science may use either a low or high data rate channel. Because of other requirements on the recording subsystem, only two channels are available for recording SDA outputs.

Signal characteristics for recording purposes are as follows:

SDA	Data	Characteristics[a]
A	Engineering	33 bits/s, rectangular waveform, 5-V signal + noise
B	Science	50 bits/s to 10.8 kilosymbols, rectangular waveform, 5-V signal + noise
C	Engineering	33 bits/s, rectangular waveform, 5-V signal + noise
D	Science	50 bits/s to 10.8 kilosymbols, rectangular waveform, 5-V signal + noise

[a]At DSS 14, science data rate is 50 bits/s to 86 kilosymbols.

Since the required dynamic range is so large, i.e., noise plus signal is large compared to the signal, a recording technique which provides good accuracy (1%) and is reasonably conservative of bandwidth and which extends to zero frequency is indicated. The technique which best satisfied these criteria is wideband FM (±40% deviation) multiplexing on a direct record track.

Configuration of the multiplex equipment for the *Mariner* Mars 1971 85-ft stations will be SDAs A and B, representing spacecraft 1, multiplexed together using voltage-controlled oscillators (VCOs) with 13.5- and 108-kHz center frequencies, respectively (Fig. 1). The mixed data will be fed to a mixer amplifier and recorded on tape using direct recording techniques. The bandwidth required is 183.6 kHz (108 kHz × 1.40 + 3 × 10.8). This indicates a recording speed of 15 in./s during encounter. SDAs C and D are processed identically.

At the 210-ft site (DSS 14), VCOs with 13.5 and 525-kHz center frequencies will be used. The bandwidth required is 994.2 kHz (525 kHz × 1.40 + 3 × 86.4). This dictates a recording speed during encounter of 120 in./s.

Suitable discriminators are also being furnished for verification of recordings and VCO setup and for possible playback of a selected recorded SDA output into the SSA, if required.

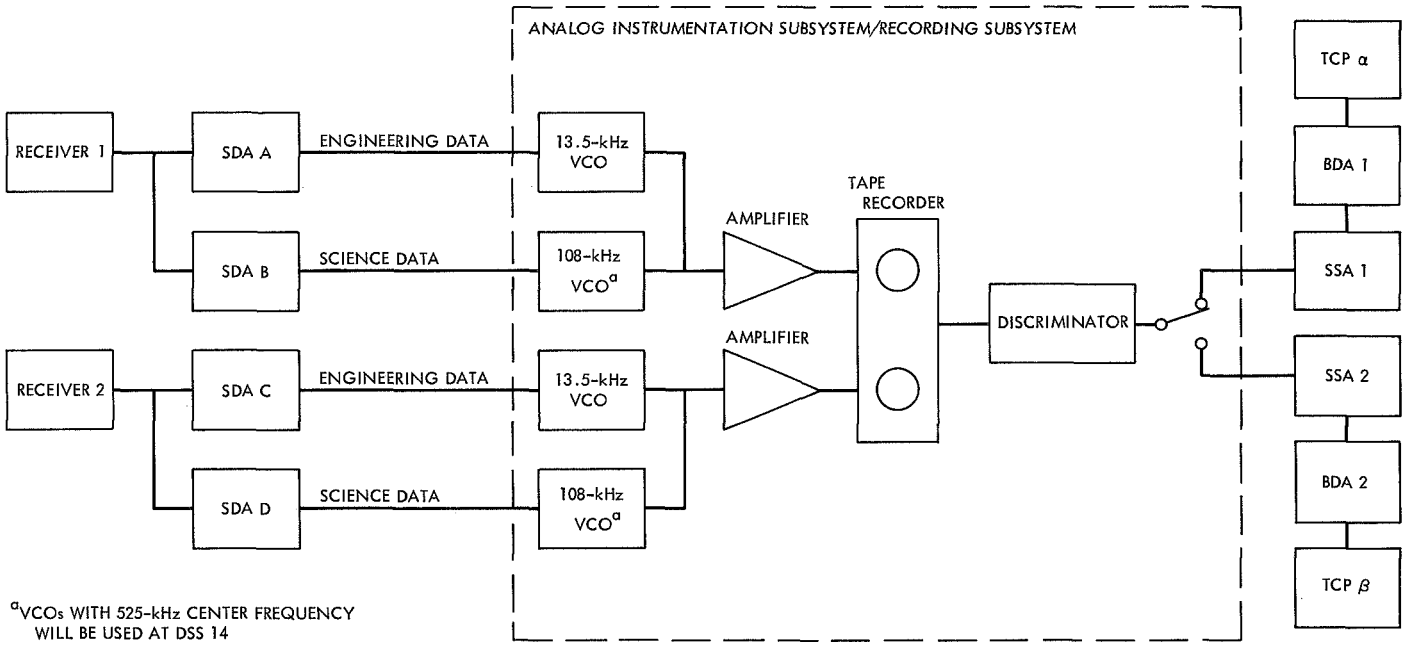


Fig. 1. *Mariner* Mars 1971 SDA output recording and playback data path

XI. DSN Operations

OPERATIONS AND FACILITIES

A. Real-Time Selection and Validation of Telemetry Data in the SFOF, *W. Kinder and W. Kizner*

1. Background

The DSN has the responsibility of delivering to the Flight Project a master data record (MDR) within 24 h after a station track. In the SFOF, the DSN Telemetry Analysis Group determines when gaps or data quality require recall of fill-in recorded data, verifies the quality of the MDR, and coordinates certification of the record with the Flight Project and the DSN Operations Control System. To accomplish this task in real-time, it was decided that DSN partial status, i.e., equipment indicators based on the presence of a signal, would accompany the telemetry data into the SFOF, where the IBM 360/75 computer would utilize these indicators for the generation of an MDR in real-time. However, a problem arises in that some of the partial status indicators from the DSIF are not indicative of the present status of the data, due to their built-in response time delay.

The DSN Telemetry System Design Team conducted a study with the intention of recommending an algorithm for data selection and validation based on these partial status observables traveling with the telemetry data. This article presents an interim algorithm for incorporation into the 1971 era Telemetry System; a presentation of the possible application of discriminant analysis to data validation; and finally, an outline of possible future work with discriminant analysis for the evaluation of these partial status observables.

2. Recommended SFOF Data Selection and Validation Algorithm

The SFOF telemetry processor should select and validate data on the following partial status indicators:

(1) Observe and record data $= F \cdot I \cdot N \cdot \bar{E}$.

F = frame sync acquired, leading and trailing edge.

I = interval timing, SFOF checks on DSIF ground received time.

N = noise with data, telemetry and command processor calculated signal-to-noise ratio estimate exceeding some minimal value.

$\bar{E}$ = no error on GCF service, highspeed and wideband data blocks.

where

Observe = drive display devices in the SFOF analysis and support areas.

Record = on-line data validation recorded on telemetry system data record.

Recall = post-track replay of facility analog, digital records.

(2) Observe and recall data = any one or all of the above observables in default.

On the selection of a data stream, the SFOF telemetry processor should function automatically. Rankings of signal-to-noise ratio estimates, transmission errors, and errors in the frame synchronization word are available for selection decision. If two or more identical data streams from a common spacecraft are equal in all observables, then the selection automatically favors the rising deep space station. The selection functions can be manually overriden and forced to favor a particular deep space station. All of the incoming data streams from a particular spacecraft, whether selected to drive display devices or not, are logged in the SFOF. Once a data stream is selected, the other streams are inhibited from driving the display devices.

The values, or favored states of these partial status indicators, are supplied by the DSN/Project and initialized manually in the SFOF before a track by action on the part of the DSN Telemetry Analysis Group. During a track, real-time changes to the partial status standards, as when going from a coded to an uncoded mode, are handled by this same Analysis Group as real-time Telemetry System standards and limits changes coordinated with the DSN Operations Control System. As shown in Fig. 1, the validation routine should be identical for either real-time telemetry or recall data.

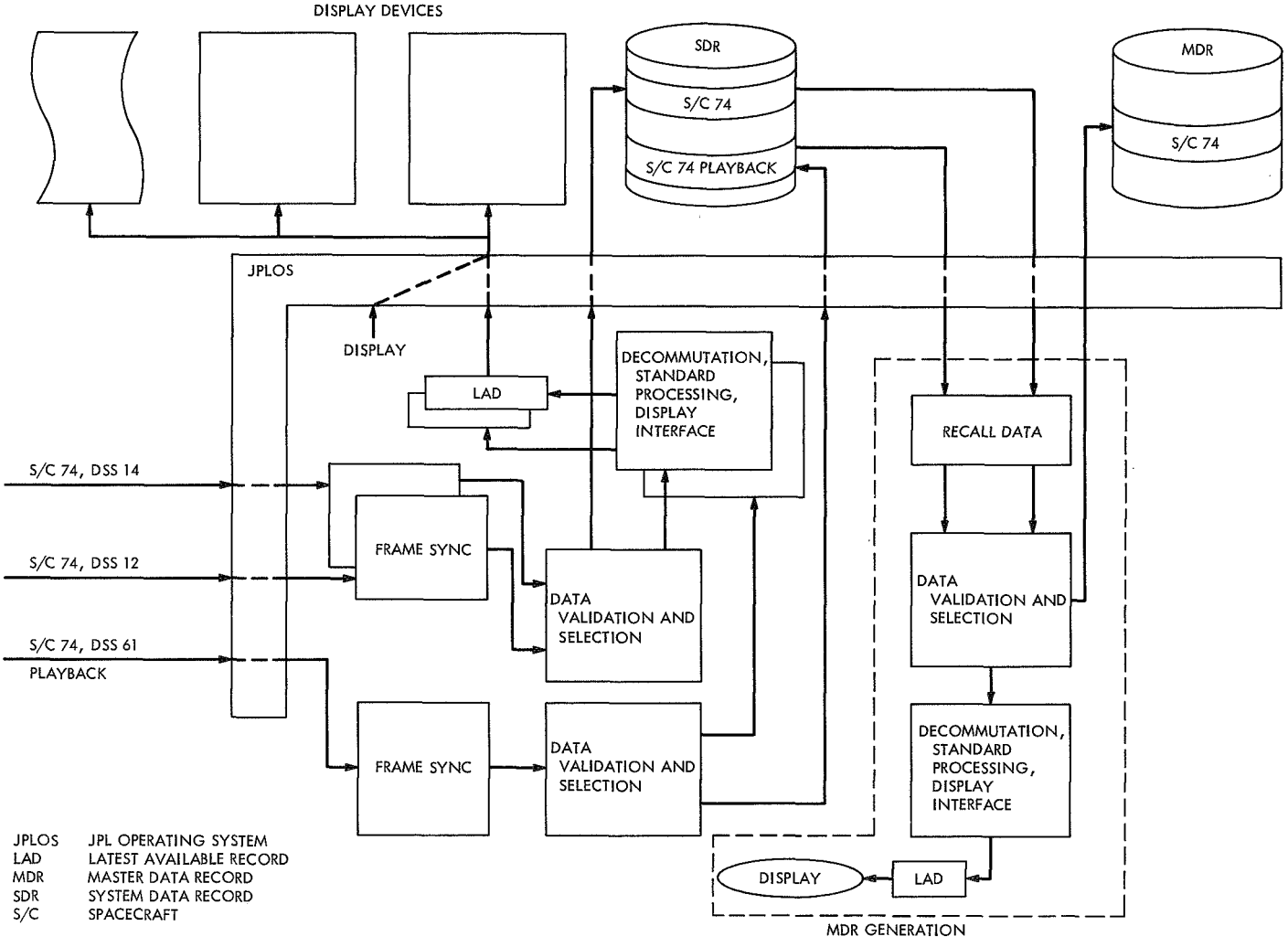


Fig. 1. SFOF telemetry data validation and routing

When a default condition is experienced, the data stream should not be inhibited automatically; however, the SFOF telemetry processor should alarm the Analysis Group with the following message:

(1) Time of outage, full sync frame prior and after bad data.

(2) Display of latest available applicable partial status information, including all lock indicators in the telemetry high-speed data blocks.

(3) Data identification, type, station, spacecraft, bit rate, recall/real-time.

This alarm routine should be the basis for a future, fully automatic telemetry recall capability.

Whenever the aforementioned validation function is overridden, the processor should tag all accumulated data during this condition as "observe and recall." Supplementary analog records are available for the Project for non-real-time processing. In the SFOF processor's periodic (hourly) and post-track telemetry summaries for the Analysis Group, the data outages should be tabulated. Based on SFOF outage time, partial status indicators, and delay characteristics inherent in their design, the Analysis Group should specify standardized replay times for recalling data from the various digital and analog facility logs.

Data selection and validation should be based on data time progression and data word repetition in the telemetry stream. This process should be automated in keeping with the DSN requirement of delivering an MDR within 24 h after a track. Manual intervention should be limited to system initialization and possible post-track recall of facility logs. Alarms to the Analysis Group should be a minimum in keeping with the Group's response capability.

3. Application of Discriminant Analysis to Data Validation

A method of combining measurements to differentiate between two (or more) alternatives is called discriminant analysis. Here the two alternatives are whether the data is good or bad. Let k be the number of measurements on which a decision is based.

In order to make use of these measurements, some knowledge of the kinds of measurements which occur for each alternative is needed. If the probability distribution of these measurements for each alternative were known, a standard procedure would be to employ the Neyman–Pearson likelihood ratio criterion (Ref. 1) for testing between the alternatives. When these distributions are not known, an applicable technique is the use of discriminant analysis. Under certain assumptions, such as that the distribution for the measurements for each alternative is a k-dimensional normal distribution having the same covariance matrix, the method of discriminant analysis is equivalent to the use of the likelihood ratio.

It is convenient to define scatter, which is related to the variance of the distribution. Let $(X_1, X_2, \cdots, X_n)$ denote a sample from a one-dimensional distribution with a cumulative distribution function $F(X)$. The scatter $S_{X_0, n}$ of the sample about the pivotal point X is defined by

$$S_{X_0,n} = \sum_{i=1}^{n} (X_i - X_0)^2 = S_{\bar{X},n} + n(\bar{X} - X_0)^2 \qquad (1)$$

where $\bar{X}$ is the sample mean. Thus, the scatter of a sample is a minimum when taken about the sample mean.

It is assumed that a sample of the measurements for each category is available; the sample represents previous cases when the measurements were made. It is required to find some single linear combination of these variables Z such that the scatter of Z between the means for each sample is as large as possible relative to the sum of the internal scatter of Z for each sample. This is roughly equivalent to finding a linear combination of variables such that the square of the difference of its expectation in the two populations is maximized, subject to the condition that the sum of the variance for each sample is a constant.

An exact statement and solution of this problem will now be given. Suppose

$$(X^{(\gamma)}_{1\xi_\gamma}, \cdots, X^{(\gamma)}_{k\xi_\gamma}, \xi_\gamma = 1, \cdots, n_\gamma), \qquad \gamma = 1, 2$$

where $n_1 > k, n_2 > k$ are two samples with k variables. Let $(\bar{X}^\gamma_1, \cdots, \bar{X}^\gamma_k), \gamma = 1, 2$, be the vectors of means. Let the matrix $\|\mu^\gamma_{ij}\|, \gamma = 1, 2$, be the internal scatter matrices defined by

$$\mu^{(\gamma)}_{ij} = \sum_{\xi=1}^{n_\gamma} (X^{(\gamma)}_{i\xi} - \bar{X}^{(\gamma)}_i)(X^{(\gamma)}_{j\xi} - \bar{X}^\gamma_j) \qquad (2)$$

Let $(\bar{X}_1, \cdots, \bar{X}_k)$ be the vector of sample means, and $\|\mu_{kj}\|$ the internal scatter matrix of the ground sample composed of the two samples pooled together. Let

$$Z^{(\gamma)}_{\xi_\gamma} = \sum_{i=1}^{k} C_i X^{(\gamma)}_{i\xi_\gamma}, \xi_\gamma = 1, \cdots, n_\gamma, \qquad \gamma = 1, 2 \qquad (3)$$

Then $(Z_1^{(1)}, \cdots, Z_{n_1}^{(1)})$ and $(Z_1^{(2)}, \cdots, Z_{n_2}^{(2)})$, except for scaling, are one-dimensional samples. Let $\bar{Z}^{(1)}$ and $\bar{Z}^{(2)}$ be the means of the two samples of Z's and $\bar{Z}$ the mean of the pooled samples. Let

$$S_W = \sum_{\gamma=1}^{2} \sum_{\xi_\gamma=1}^{n_\gamma} (Z_{\xi_\gamma}^{(\gamma)} - \bar{Z}^{(\gamma)})^2 \tag{4}$$

$$S_B = \sum_{\gamma=1}^{2} n_\gamma (\bar{Z}^{\gamma} - \bar{Z})^2 \tag{5}$$

It will be noted that, if S is the scatter of the grand sample obtained by pooling the two samples of Z's, then $S = S_W + S_B$. S_W is the within sample component and S_B is the between sample component of S. The basic problem is to determine $(C_1, \cdots, C_k)$ so as to maximize S_B for a fixed value of S_W. The solution is given by the following theorem, whose proof can be found in Wilks (Ref. 1).

THEOREM. *Let $\|\mu_{ij}^W\|$ and $\|\mu_{ij}^B\|$ be within sample and between sample scatters of two samples from a k-dimensional distribution where the sample sizes both exceed k and where $\|\mu_{ij}^W\|$ is positive definite. Let S_B and S_W be defined as in Eqs. (4) and (5). The value of $(C_1, \cdots, C_k)$, say $(C_1^1, \cdots, C_k^1)$, which minimizes the ratio*

$$Q = \frac{S_W}{S_W + W_B} \tag{6}$$

so that S_W has a fixed value $C \neq 0$, is the solution of

$$\sum_{j=1}^{k} (\mu_{ij}^B - l_1 \mu_{ij}^W)\, C_j = 0, \qquad i = 1, \cdots, k \tag{7}$$

where l_1 is the non-zero root of the characteristic equation

$$|\mu_{ij}^B - l\mu_{ij}^W| = 0 \tag{8}$$

given by

$$l_1 = \frac{n_1 n_2}{n_1 + n_2} \sum_{i,j=1}^{k} \mu_W^{ij} (\bar{X}_i^{(1)} - \bar{X}_i^{(2)})(\bar{X}_j^{(1)} - \bar{X}_j^{(2)}) \tag{9}$$

where

$$\|\mu_W^{ij}\| = \|\mu_{ij}^W\|^{-1} \tag{10}$$

Thus, the vector $(C_1, \cdots, C_k)$ can be determined and used to differentiate or discriminate between two alternatives. Furthermore, the minimum value of Q is $1/(1 + l_1)$.

These equations have been programmed and checked. One check on the accuracy of the computation is to calculate Q by means of Eq. (6) and $1/(1 + l_1)$, and to compare the results. The two values have agreed to about seven places in all the cases that have been run.

One problem which may occur is the case where $\|\mu_{ij}^W\|$ is very nearly singular and cannot be inverted. This condition can be determined by the use of the JPL library routines which calculate the singular value decomposition of matrices. It can readily be seen that $\|\mu_{ij}^W\|$ will be singular if one or more variables are given as a linear combination of the others. Conversely, if $\|\mu_{ij}^W\|$ is singular or very nearly singular, then the data is consistent with the hypothesis that one or more variables are given as linear combinations of the others. In this case, nothing is lost by eliminating those variables which are linear combinations of the others. To do this, the singular value decomposition is used to determine which variable should be eliminated so that the resulting matrix $\|\mu_{ij}^W\|$ with one less variable should have the best condition number.

For these matrices, the condition number is equal to the ratio of the largest to the smallest eigenvalue. This procedure may be repeated any number of times until a well-conditioned within-scatter matrix results.

To implement this method, one needs past histories of indicators or measurements for good and bad telemetry data, from which the means and $\|\mu_{ij}\|$, $\|\mu_{ij}^W\|$, and $\|\mu_{ij}^B\|$ can be calculated. Then the theorem is used to calculate the linear combination determined by $(C_1, \cdots, C_k)$. If necessary, some variables may be eliminated if they appear to be linear combinations of others.

4. Future Work

Some of the partial status observables are redundant for this application. Some, because of their time response, are meaningless. We still need data on how these indicators, when displaying outages, propagate through the Telemetry System. We also need data on how the fixed observables, such as measured signal-to-noise ratio and errors in the fixed frame sync word, behave as threshold is approached. Tests are planned to acquire field data on the above unknowns and to evaluate the indicators using the discriminant analysis technique.

Reference

1. Wilks, S. S., *Mathematical Statistics*, Sec. 18.7. John Wiley & Sons, Inc., New York, 1962.

B. DSN Tracking System Operations,

J. Heller and R. B. Miller

1. Introduction

The purpose of this article is to describe the operations of the DSN Tracking Operations Analysis Group in the Mark IIIA era. Most of the functions described are being performed at the present time by manual procedures and with the use of the software available on the IBM 7044/7094 computer system. The design of the IBM 360/75 software is geared towards minimizing manual operations in the accountability and validation of tracking data and towards using computer displays to monitor data flow and system performance.

2. Functions

The DSN Tracking Operations Analysis Group performs the following functions:

(1) Monitors the operations of the SFOF portion of the Tracking System editing and validation functions and the DSN Tracking System in general.

(2) Coordinates the tracking editing function and the presentation and display of tracking data.

(3) Monitors the partial status indicators inserted in the tracking data stream and provides regular reports and alarms to DSN operations control.

(4) Generates and provides tracking standards and limits in response to operational and status information received from the spacecraft project and the DSN Operations Control System. These will include such items as predict tolerances, doppler tolerances, and frequency precision tolerances.

(5) Generates tracking prediction information as required, and certifies its availability for transmission. The DSN is capable of generating its own predicts without project inputs, but the project is expected to provide validated trajectories to ensure that the most up-to-date information is used to generate predicts.

(6) Coordinates the tracking calibration function.

(7) Coordinates any orbit determination function which may be necessary for the DSN generation of initial conditions for prediction purposes.

3. Software

The DSN Tracking Operations Analysis Group is the cognizant user engineer for the SFOF Tracking System software. This assignment coincided with the DSN decision to implement configuration C for the *Mariner* Mars 1971 mission. This configuration calls for the DSN Tracking System software to reside in the IBM 360/75 computer while the orbit determination and the trajectory functions are accomplished in double precision on the Univac 1108 computer. An integral SFOF tracking subsystem software design was evolved which operates in real time in the IBM 360/75 under the JPL operating system. Figure 1, a functional diagram of this design, shows the main software elements and data files. A brief description of these components will serve to clarify the diagram.

a. Teletype tracking data input processor (TYDIP). TYDIP is a multimission real-time program responsible for the decommutation and reformatting of tracking data received on teletype from various stations and forwarded by the Univac 490 communications processor. These data will be examined by TYDIP for acceptable format; unacceptable data are stored in a rejected data file; data condition codes and data values are checked to determine that they are within allowable ranges; the accepted data are then converted to DSIF Tracking Subsystem (DTS) format and a subroutine, teletype tracking data outage subroutine, is called to compute sample rates and detect sample rate changes and data outages; a pass identification code is attached to the data stream, which is then routed to the Pseudo-Residual Program.

b. DTS tracking data input processor (DTDIP). DTDIP is a multimission real-time program responsible for the decommutation and reformatting of tracking data and messages received over the high-speed data lines (HSDL). Messages are routed to the DTDOP message processor described below. Tracking data in TTY format are routed to TYDIP for processing, as described above. Tracking data from the DTS, which already contains sample rates and pass identification codes, is simply routed to the Pseudo-Residual Program. Both TYDIP and DTDIP have the same output format.

c. DTS tracking data output processor (DTDOP). DTDOP interprets control messages received from the DTS and composes and transmits over HSDLs appropriate response messages. These include data transmission requests, predict transmissions, predict instructions, etc. In particular, DTDOP detects outages and initiates data recall automatically.

d. Pseudo-Residual Program. This is a real-time multimission program which operates on all incoming tracking data and differences the data from predicts. It interfaces with the Predicts Program to obtain prediction data which

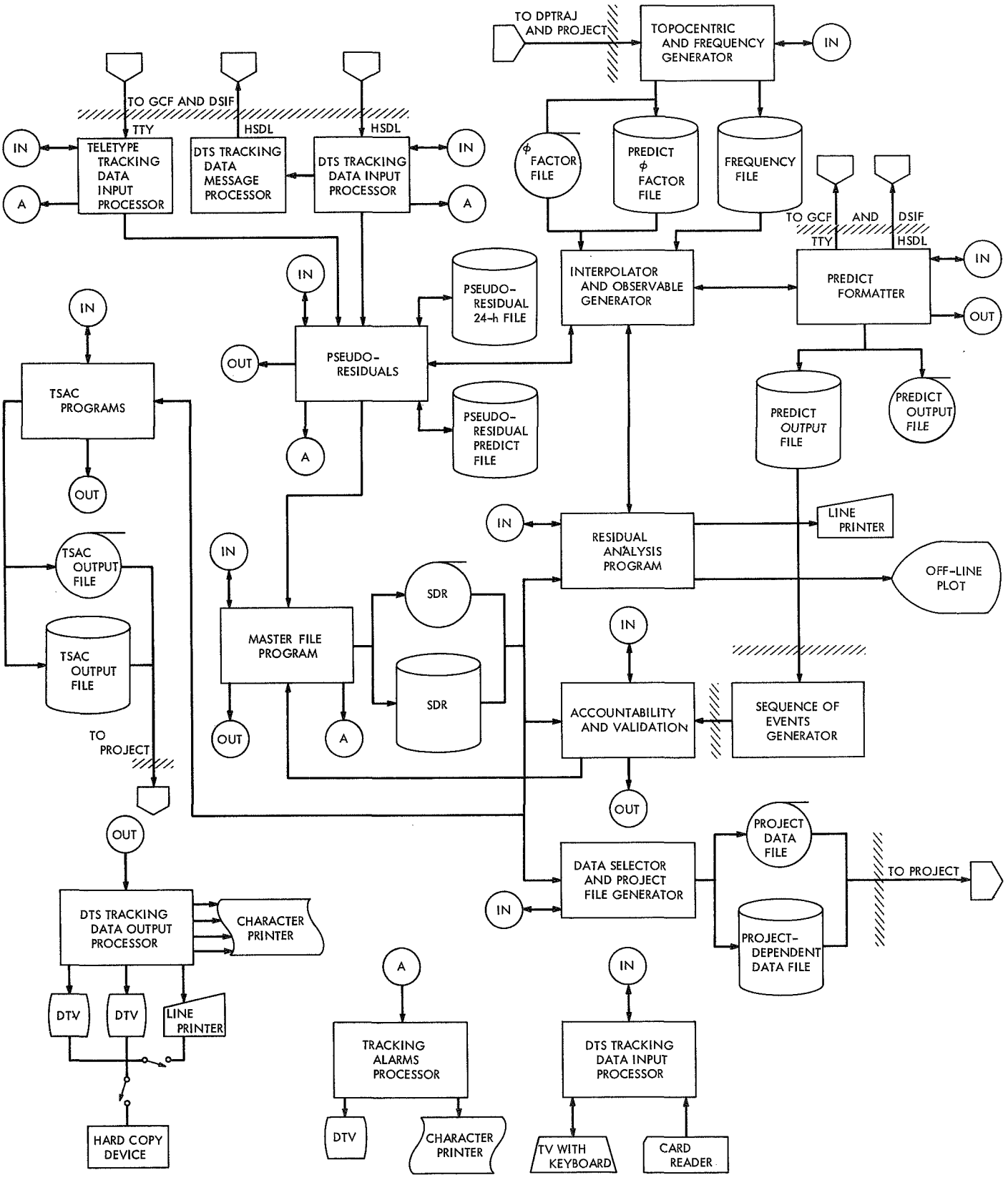


Fig. 1. SFOF tracking subsystem software design

are stored on disk, as described below. It also computes data noise for doppler and ranging data and outputs all these computations as printed listings and for digital TV (DTV) plots. In addition, the program generates a data quality indicator which is placed alongside the non-DTS tracking data (the DTS data has a quality indicator already). All tracking data are then routed to the Master File Program. The last 24 h of pseudo-residual data are stored on disk for analysis.

e. Master file program (MFP). The primary responsibility of the MFP is the creation and management of the system data record (SDR) file. This file contains up to 2 weeks of tracking data associated partial status code and accountability information and resides on disk. The MFP accomplished four main functions:

(1) Processes tracking data received from TYDIP or DTDIP and writes it on the SDR file. Recall data can be merged in time sequence.

(2) Allows the DSN Tracking Operations Analysis Group to add partial status code (such as transmitter on times and frequencies, range bids, etc.) and to perform data editing functions, i.e., delete, suspect, ignore data.

(3) Computes in real time an accountability summary for each pass listing data received, good data received, and percentages for each data type.

(4) Outputs the SDR or portions of it onto tape, transfers it to the Univac 1108 via the electrical interfaces, prints it out on the line printer or displays it on DTV.

f. Accountability Program (AP). The function of the AP is to perform non-real-time accountability on the SDR file. It will compute over any time interval the number of data points, number of good data points, and percentage of good data for each data type, and output summaries on DTV or printer.

g. Tracking alarms processor (TAP). The TAP receives alarms from TYDIP, DTDIP, pseudo-residuals, and MFP. In addition, it receives some status and accountability information from these programs and formats all of this information into a continuous log of alarms and outages. This log is output on a printer sequentially and is displayed on DTV in a special format which is updated automatically to show the last alarm and data status for each data stream.

h. Residual Analysis Program (RAP). The RAP is a non-real-time data analysis program. For an input time interval, it computes data residuals for each data type based either on a set of predicts or on a least squares fit of the data itself. Once the residual set is obtained, the RAP computes upon request the mean and standard deviation of the set, the cumulative residual and the local standard deviation at each point, and the autocorrelation function and power spectrum for the set. The output is available in either printed or plotted form.

i. Tracking system analytic calibration software (TSAC). The TSAC is under the cognizance of the Mission Analysis Division. TSAC consists of two non-real-time IBM 360/75 routines: (1) the platform observables subassembly, which provides calibrations for DSS locations, polar motion, and time, and (2) the transmission media subassembly, which provides calibrations for the troposphere, ionosphere, and space plasma effects, using differenced range versus integrated doppler. This software will be operated by the DSN Tracking Operations Analysis Group and the resulting polynomials will be furnished to the project.

j. Predicts Program (PREDIX). The function of PREDIX is to compute predicted observables for use both in the acquisition of the spacecraft by the DSS and in the validation of the data by the Tracking System. The input to PREDIX is a probe ephemeris tape generated by the Double-Precision Trajectory Program (DPTRAJ) on the Univac 1108. This input is transferred via tape or via the Univac 1108/IBM 360/75 electrical interface. PREDIX then generates for each station a file of angles, range, and range rate (ϕ factors) computations which are stored on disk in the IBM 360/75. The ϕ factor file contains up to 2 weeks of prediction data for each active spacecraft. At the time of transmission, an interpolator and observable generator routine uses a stored frequency file to convert the factors to doppler frequencies and range units interpolated for the desired sample interval and makes these available for transmission by teletype (TTY) or HSDL. The interpolator routine is also accessed by pseudo-residuals and RAP when these programs need predicts. In addition, the program computes and stores a set of view periods for extended periods of time. Predict output is available on tape and DTV. Further interfaces of the program include the tracking simulation program and the sequence of events generator. View periods interface with the scheduling software.

4. Hardware

The DSN Tracking Operations Analysis Group will perform its real-time functions in the SFOF from an area within the Network Analysis Team area. Figure 2 shows

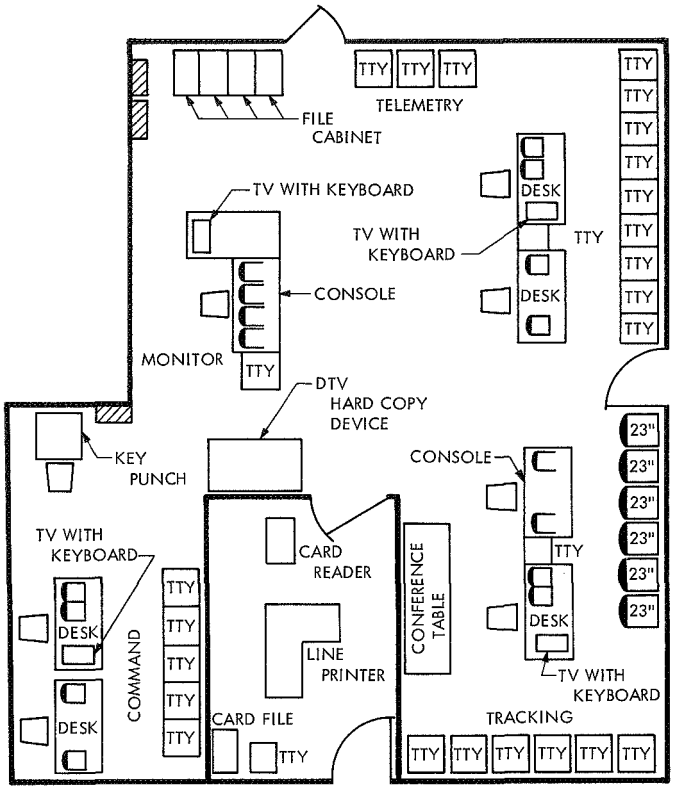


Fig. 2. Tracking Operations Analysis Group area in SFOF

the layout of the area and the hardware complement allocated to support the DSN Tracking Operations Analysis Group. In addition to the desks and chairs, the hardware includes an appropriate operational voice communications system capability and the following input/output devices:

(1) An IBM 2260 with keyboard. This device is an input message processor consisting of a keyboard and a cathode-ray tube display. It will be used for composing most of the real-time input and control messages for the Tracking System software.

(2) A shared IBM 2501 card reader for bulk inputs via punched cards.

(3) Seven TTY printers for display of tracking data, pseudo-residuals, and tracking alarms.

(4) A shared IBM 1443 line printer for bulk listings of tracking data and predicts.

(5) Four closed-circuit TV monitors, which include 3 DTV output channels dedicated to the Tracking System software. These DTV channels will display pseudo-residual plots, tracking alarms, and data summaries. A fourth DTV channel, driven by the Monitor System software, will be dedicated to displaying the tracking system parameters gathered by the Monitor System.

(6) A shared DTV hard copy device which will deliver on request a print of a selected DTV display.

(7) An Olivetti programmed 101 desk top computer to perform spacecraft frequency determinations, blunder point calculations, etc.

5. Operations

The DSN Tracking Operations Analysis Group will maintain its operational position on a round-the-clock basis and perform most of its functions in real time. The operational functions fall primarily in the following areas.

(1) *Tracking data functions.* The group will operate the Tracking System software and monitor the tracking data flow. It will initiate data recall, merge the partial status code into the data, validate the quality and quantity of data on the SDR written 24 h after the end of a tracking pass, transfer the data to the user via the Univac 1108/IBM 360/75 electrical interface, and write the data onto an SDR tape for safekeeping and delivery to the users.

(2) *Tracking system functions.* The group will monitor the Tracking System performance in real time both by the content of the tracking data via pseudo-residuals and by the Monitor System display of tracking parameters gathered throughout the network. The group will alert the Operations Control Team of any malfunction or anomaly and will attempt to isolate the facility involved.

(3) *Predict functions.* The group will be responsible for obtaining the necessary trajectory and frequency information, operating PREDIX, and generating the appropriate prediction data in a timely manner. It will also validate the output before transmission to the station or delivery to the user.

(4) *Tracking system analytic calibrations.* The group will be responsible for gathering the raw data and operating the software for the TSAC function. The resulting calibration polynomials will be validated by cognizant Mission Analysis Division personnel before delivery to the user.

These real-time functions will be backed up with a small amount of non-real-time support dedicated to the in-depth analysis of problems and data, the generation of operating procedures, and the writing of reports.

XII. SFOF Operations

OPERATIONS AND FACILITIES

A. Man–Machine Interaction in the Post-1971 SFOF, *L. R. Huesmann*

1. Introduction

The basic design for the *Mariner* Mars 1971 SFOF system was derived from JPL's IBM 7094/7044 system and the Manned Spacecraft Center's IBM 360/75 real-time operating system (RTOS). The *Mariner* Mars 1971 mission is the first of a type of mission with some fundamental characteristics quite different from previous unmanned missions and also from the type of mission the Manned Spacecraft Center's RTOS was designed to support. It seems likely that missions in the post-1971 time period will exacerbate the problems created by these fundamental changes to the point that major alterations may need to be made in the SFOF system. The purpose of this research was to outline some of the general problems that the SFOF must face in the post-1971 time period and to suggest how advances in the field of man–machine interaction can be applied to these problems.

2. Characteristics of the Post-1971 Missions

It seems probable that a large number of missions in the post-1971 time period will possess a set of characteristics that has not been typical of earlier missions. First, missions will involve multiple spacecraft active during the same time span, and several missions will be concurrently active. Many spacecraft may require equal attention from the SFOF during a given period. Second, the active period for spacecraft may increase beyond 90 days. Third, the round-trip light time to the spacecraft will usually be long enough to make any sort of real-time control from earth impossible. Fourth, the delay between final testing (launch) and operation of the spacecraft will increase greatly. Fifth, the number of scientific experiments executed during any mission can be expected to increase. Each spacecraft in a multiple-craft mission may carry much more complex scientific equipment than previous flights and engage in a much larger number of experiments. Finally, as a result of these factors, one can expect an exponential increase in the command and control functions that the SFOF system must handle.

Let us now consider some of the major specific problems posed by these changes in mission characteristics.

3. Reliability Problems

The increased complexity, and particularly the increased duration of missions in the post-1971 time period, will

raise serious reliability problems both in the hardware and in the human management of the missions. The probability of a failure simultaneously in both the prime and back-up computing facility at the SFOF cannot yet be estimated but must be kept low for a 3- to 6-month active mission. In addition, the on-board equipment, and particularly the on-board computer, must have a reliability comparable to the earth-based computing facilities. Hence, hardware failure rates low enough for missions in the 1960's will probably be too high for post-1971 missions. In addition, the operator error rates are likely to increase as the missions lengthen. The word "operator" in this context means any human involved in the control of the mission. In particular, the decision makers on the various science and engineering teams can be expected to commit errors on long duration missions requiring constant attention to detail. Jenkins (Ref. 1), Yntema (Ref. 2), and others have found that operator error rates which are quite low for limited periods increase exponentially as the operator is faced with the same task for longer periods and at irregular intervals. A decision maker's performance is likely to be worst on routine tasks requiring concentrated attention at irregular intervals. The important point to remember is that a decision maker's reliability will decrease beyond the extrapolation of daily reliability rates when the above contingencies are encountered.

4. Time and Resource Problems

The second major group of problems to be expected in the post-1971 time period arises from the increased use of fixed resources and the increased activities in any time period.

Presumably, it is desirable on any mission to maintain as short a regular planning cycle as possible during the operation of the mission. Yet, one can expect the regular planning cycles to become very cramped in the post-1971 time period unless they are lengthened beyond the range proposed for the *Mariner* Mars 1971 mission. One can divide a planning cycle into five stages: data acquisition, data analysis, decision making, event scheduling, command generation, and command transmission. It is obvious that with multiple spacecraft, increased scientific experimentation, and more-complex flight plans, the analyses, decision making, scheduling, and command generation times will expand. In addition, since light time will be great enough to prohibit any sort of real-time control, the planning process will be complicated by the need for basing plans on predicted sequences of events and the need for the development of contingency commands that may or may not be executed by the spacecraft. One should also not forget the small but significant strain put on the cycle by the loss of planning time directly due to light time.

The time from acquisition of the last data used in planning to transmission of the last command will be at least one round-trip light time less than the spacecraft activity cycle. For example, for a craft in the vicinity of Jupiter, round-trip light time is about 90 min. Hence, to achieve a 24-h lag between the occurrence of data and the spacecraft's response would require a 22.5-h earth planning cycle. Finally, on post-1971 missions the planning cycle will also be more susceptible to unpredictable overloads. With multiple-craft missions and multiple missions many events will be occurring in parallel. Hence, unusual combinations of unexpected events are much more likely to swamp the system.

Regardless of the planning cycle, at some point the decision makers are faced with the problem of scheduling the events that have been decided upon. As mentioned above, many operations may be executed at the same time. What makes the scheduling particularly difficult is the likelihood of conflicts between events. A number of events may require the same resource or a resource that is only available at specified times, e.g., DSS 14. This is as true of activities within the DSN and SFOF as on-board activities. The events must be scheduled to be completed in as short a time as possible without conflicts. Scheduling tasks in this way is by no means an easy job, and it is one that is beyond the capability of the present *Mariner* Mars 1971 scheduling routines.

5. Real-Time Control Problems

As described above, the impossibility of any real-time control of spacecraft activities increases the complexity of the planning cycle; but this phenomenon also prevents some activities entirely and increases the cost of others. The utility of real-time control is quite mission dependent. On orbiter missions repeated observations can be made at low cost until one is sure the craft obtained the desired information. However, on fly-by and landing missions real-time control can be extremely useful. For example, one would like to be able to employ complex real-time instrument-adjustment algorithms, redirect the scan platform on the basis of received data, and (on lander and exploration missions) avoid obstacles or sample ground material intelligently.

6. Principles for an Advanced SFOF System

Four related approaches for ameliorating the difficulties discussed above have survived scrutiny.

(1) Upgrade the on-board computer.

(2) Incorporate additional redundancy into the human system.

(3) Increase the speed of information transmission and carefully control the dispersal of information to operators.

(4) Automate a much greater portion of the monitoring, analysis, and decision making.

Each of these alterations would contribute to improvement in the three major problem areas expected to develop in the post-1971 time period.

a. Upgrading the on-board computer. Upgrading the on-board computer would have the obvious effect of permitting some real-time control of activities. In addition, and perhaps more importantly, greater reliability of on-board hardware can be achieved with a sophisticated on-board computer. Exactly what sort of increment can be achieved in these areas would depend on how much computing power were provided. However, even a very small on-board computer with an arithmetic central processing unit and priority interrupts could accomplish a great deal.

The ability of such a computer to improve the reliability of the on-board hardware (including itself) is not only within the state of the art, but a prototype of such a computer has been built at JPL. The self-testing and repairing (STAR) computer is capable of diagnosing and correcting both transient and permanent hardware faults in itself (Ref. 3). The only major limitation on extending the computer's diagnosing and repairing functions to other on-board hardware would be the cost of the additional computing power required and the replacement hardware. However, it is known that a specified reliability can be achieved with less hardware redundancy when fault detection, hardware testing, and repair by replacement are instituted. It is appropriate, then, to balance the cost of the additional computing power against the savings in redundant hardware. Fault monitoring could be carried out in real-time by the on-board computer with problems that exceed the on-board computing power relayed to earth. Conceivably, faults fatal to missions without this real-time capability could easily be handled by a small on-board computer. Avizienis (Ref. 3) presents a more comprehensive discussion of the advantages of the replacement method over other means of achieving reliability.

In addition to permitting a repair-oriented approach to reliability, upgrading the on-board computer allows some sophisticated forms of operational control that would otherwise be impossible. Two particular applications will be briefly discussed here: video field control on fly-by missions, and obstacle avoidance during robot exploration of other planets. Many workers in pattern recognition—most recently, Guzman (Ref. 4)—have demonstrated that, in situations where human operators appear to be using very complicated rules to select patterns, a very simple set of rules actually suffices.

The state of the art is such that a reasonably simple algorithm could be developed to direct the scan platform to acquire video targets of exceptional interest in high resolution. One should not underestimate the problems in on-board video processing; but neither should one reject the approach under the false assumption that an appropriate decision algorithm could not be programmed for the on-board computer. To construct the algorithm one would need to study how the human planning teams make the same decisions, extract the principles they use, and construct the algorithm about those principles. Such an approach has been used with success in many problem-solving fields, e.g., Greenblatt (Ref. 5) and Tonge (Ref. 6). Obstacle avoidance and pathfinding by robots has already been studied extensively by artificial intelligence workers, e.g., Nilsson (Ref. 7), Kirk and Lim (Ref. 8). Again, the conclusion to be drawn is that algorithms could be written for an on-board computer so that it could effectively guide an explorer robot. While it is true that the performance of an on-board computer on these two tasks would probably be inferior to its performance if it were controlled by a human operator, one must remember that light time prevents any sort of human control.

Perhaps the major advantage accruing from a sophisticated on-board computer would be the reduction in activities required of the SFOF. Contingency programs and monitoring routines in the on-board computer would reduce the strain on the SFOF.

b. Human redundancy. The major nemesis of human reliability over a long mission is boredom and uncertainty. Errors in routine tasks will become much more probable than a simple extrapolation of daily reliability

rates would indicate. Performance in highly stressful decision-making situations that occur repetitively at irregular intervals is especially affected. The difficulty is that problem solvers must be in a highly attentive, motivationally aroused state to perform at their best. This aroused state simply cannot be maintained over a long period of time for tasks with low probabilities of requiring nonroutine actions at any time (Ref. 9).

The situation can be improved with several modifications to the human information-processing system of the SFOF. First, continuous monitoring of parameters by human operators should not be attempted. The ground-based computing facilities in conjunction with the on-board computer should not only monitor relevant raw data, but should also execute sophisticated analyses of the data in near real-time. The computing system should be capable of determining whether parameters are within acceptable bounds, predicting probabilities that they will continue to be acceptable, and composing commands to adjust parameters, if necessary. Human monitoring should be viewed as a redundant monitoring system to be used if the other system fails. If a critical problem occurs, the system should be able to raise an alarm, display processed data, suggest a remedy, and begin a remedy if time does not permit human intervention. During normal operation, at prescribed intervals during the day, the system would present reports for the operators to study. This technique of involving the human decision maker only at established regular intervals maximizes the reliability of human performance over long periods (Ref. 9).

A second modification would be to introduce operational redundancy into the human decision-making process. Crucial decisions, command construction, and data analyses performed by one man should be verified by at least one other person who operates independently of the first. This is already done informally but should be systematized. The cost of such redundancy is not small and should be reserved for areas where fatal errors are not improbable. For example, a team's input to the command generation program should be verified by an independent decision maker, and a team's confirmation that the command sequence is correct (on the basis of the simulated execution of the commands) should be verified by an independent decision maker.

c. Improving information transmission. The SFOF computing system is a time-sharing system. As with any such system, the response time to a request varies greatly with the load on the system. As a result, one can predict that decision makers frequently will spend significant amounts of time waiting for output that may arrive in 1 s or 10 min. As Nickerson (Ref. 10) has pointed out, even the most sophisticated decision makers find it impossible to operate efficiently under such response uncertainties. Of course, the problem is serious only if users are engaged in several simultaneous tasks and are pressed for time. But this is exactly the situation one expects to occur at the SFOF.

The only solution that has been proposed for this problem is quite difficult for systems managers to accept, since it involves a deliberate delay of information. Yet it undoubtedly increases user efficiency (Ref. 11). The monitor computes a conservative estimate of when the system can respond to the user's request, informs the user of the response time, and responds at that time even if the information is available sooner. Of course, one would want to be able to override this facility in certain emergencies. However, one should note that such a procedure in no way increases the computer's idle time or overhead. Output is simply stored until the appropriate delivery time.

Information transmission between human operators must also be facilitated. The SFOF management system for *Mariner* Mars 1971 divides responsibilities primarily by function. Communication among decision makers is vital for the success of this approach. However, past experience (Ref. 9) provides some evidence that concurrent missions involving many spacecraft will require more division of responsibility by spacecraft for reliable operation. In this situation the ease of communication among operators performing parallel tasks becomes even more crucial. More emphasis will need to be placed on locating hardware and personnel to facilitate communication, selecting operators, and training the human system by simulation.

The SFOF operating system can also ensure improved human reliability and efficiency by carefully regulating the normal distribution of information to the planning teams. One might think that more information is always desirable. Yet it has been found that it is relatively easy to overwhelm decision makers with irrelevant information to the point that they cannot extract the relevant information (Ref. 9). Data distribution to planning teams should be designed to communicate on request exactly what is desired and no more. Though it may seem counter-intuitive, continuous presentation of data is more likely to lead to human error and inefficiency than regular periodic presentations of summary data.

Some strain is also placed on the planning cycle by the SFOF's hardware and software limitations on information transmission speed. For example, the current command transmission rate of 1 bit/s is an order of magnitude too small. Advances in the state of cathode ray tube terminals can be expected to permit more rapid presentation of data to decision makers and the elimination of most mechanical output devices. Advances in the use of engineering psychology in designing terminals should lower error rates as well. Significant advances in the speed of input devices are less likely. However, the input requirements of programs can be designed in ways that minimize the likelihood of user error and minimize transmission times. For example, mnemonics are generally preferable to numeric commands, and an immediate acknowledgment for any input is desirable.

d. Automating human decision making. The fourth conclusion of this report was that greater use needs to be made of the computing facilities in the decision-making process. Actually, many of the above suggested modifications to the SFOF system have involved such a change. It has already been suggested that the on-board computer could make several types of operational real-time decisions and that the SFOF system could assume a large portion of the fault monitoring, parameter adjusting, and repair switching functions human managers have performed in the past. Besides incrementing reliability, this approach would reduce the load on the human planners and improve their performance on other decisions. However, to keep the planning cycle reasonably short, it is proposed that, in addition, the SFOF computing facility must assume a greater burden of the planning functions.

An axiom about decision making to which artificial intelligence researchers have learned to subscribe is that most human decisions are systematic whether the operator knows it or not (Ref. 2). If one can discover the algorithm the decision maker is using, one can program a computer to make the same decision. The computer will usually make the decision more rapidly and consistently than the human operator, although it may suffer defects in its adaptability to new situations (Ref. 2). The typical approach to discovering the algorithm has been to query the decision maker and to study his behavior in controlled situations. This approach has been successfully employed on tasks ranging from scheduling jobs in a factory to playing chess. It is a conclusion of this study that there are excellent chances for the success of such an approach in discovering many of the decision algorithms the planning teams employ. The resulting program would certainly be somewhat inferior to the human planner in handling new situations; so obviously one would employ the human decision makers to verify the decisions. Again, the proposed change is one that makes the human operator a back-up system, reducing the strain on him, and improving his performance on crucial tasks. One specific planning task for which such an approach is virtually a necessity is the previously described problem of scheduling parallel events requiring limited resources. There are no known algorithms for solving this type of problem. Huesmann (Ref. 12) has constructed a program that models human behavior on these problems and does much better than human operators in any reasonable amount of time.

Many will question the ability of the computing facilities to perform the decision-making functions previously executed by men. This skepticism is justified in the light of past experience. However, recently the field of artificial intelligence has advanced to a stage where the implementation of previously experimental techniques is producing large increments in system performance. Several examples have been cited above (Refs. 4, 5, 6, 12). In direct competition with man we have medical diagnoses programs that are more accurate than a single excellent doctor (Ref. 13). Edwards (Ref. 14) has shown that, while sophisticated decision makers are good at forming subjective estimates of the probability of a single event, they are quite bad at combining incoming data and *a priori* probabilities to judge the likelihood of an event. No one is suggesting that all decisions can or should be automated. Rather it is proposed that a selective automation of some decisions would improve the reliability and quality of many decisions.

7. Conclusions

A study was undertaken of the problems the SFOF can be expected to encounter in the post-1971 time period. It was concluded that many of the difficulties could be relieved by four major changes in the man–machine interface. It is proposed: (1) that the on-board computer should be upgraded to assume a portion of the functions previously executed by ground-based machines and decision makers, (2) that crucial human decision-making be executed in parallel, (3) that the ground-based system be designed to disperse information in a way that improves human reliability and efficiency, and (4) that the computing systems assume a greater portion of the decision-making and planning functions. Such alterations should improve the reliability of performance,

reduce the required time for a planning cycle, and permit some desirable activities that cannot be accomplished without real-time control.

References

1. Jenkins, H. M., "The Effect of Signal Rate on Performance in Visual Monitoring," *Am. J. Psychol.*, Vol. 71, pp. 647–661, 1958.
2. Yntema, D. B., and Torgerson, W. S., "Man-Computer Cooperation in Decisions Requiring Common Sense," *IEEE Trans. Human Fac. in Elec.* HFE-2, pp. 20–26, 1961.
3. Avizienis, A., *Design Methods for Fault-Tolerant Navigation Computers,* TR. 32-1409, Jet Propulsion Laboratory, Pasadena, Calif., Oct. 15, 1969.
4. Guzman, A., "Decomposition of a Visual Scene Into Three-Dimensional Bodies," *AFIPS Proceed.*, pp. 291–304, Fall 1968.
5. Greenblatt, R., Eastlake, D., and Crocker, S., "The Greenblatt Chess Program," *AFIPS Proceed.*, pp. 801–810, Fall 1967.
6. Tonge, F., "An Heuristic Program for Assembly Line Balancing." In *Computers and Thought.* Edited by E. Feigenbaum and J. Feldman. McGraw-Hill Book Co., New York, 1963.
7. Nilsson, N., "A Mobile Automaton," in *Proceed. Internat. Joint Conf. Art. Intell.*, Washington, 1969.
8. Kirk, D. E., and Lim, L. Y., "A Dual-Mode Routing Algorithm for an Autonomous Roving Vehicle," in *IEEE Trans. Aero. Elec. Sys.*, AES-6(3), pp. 290–294, May 1970.
9. Green, B. F., "Man-Computer Systems," in *Digital Computers in Research,* McGraw-Hill Book Co., New York, 1963.
10. Nickerson, R. A., "Man-Computer Interaction: A Challenge for Human Factor's Research," *IEEE Trans. Man–Mach. Sys.*, pp. 164–180, Dec. 1962.
11. Simon, H. A., "Reflections on Time-Sharing From a User's Point of View," *Comput. Sci. Res. Rev.*, pp. 43–51. Carnegie Institute of Technology, Pittsburgh, Pa., 1966.
12. Huesmann, L. R., "A Study of Heuristic Learning Methods for Optimization Tasks Requiring a Sequence of Decisions," *AFIPS Proceed.*, Spring 1970.
13. Kleinmuntz, B., "The Processing of Clinical Information by Man and Machine," in *Formal Representation of Human Judgment.* Edited by B. Kleinmuntz, John Wiley & Sons, New York, 1968.
14. Edwards, W., "Conservatism in Human Information Processing," in *Formal Representation of Human Judgment.* Edited by B. Kleinmuntz, John Wiley & Sons, New York, 1968.

XIII. DSIF Operations

OPERATIONS AND FACILITIES

A. Radio Science Support, *T. Sato, L. Skjerve, and D. Spitzmesser*

1. RAES Panel Activities

The Radio Astronomy Experiment Selection (RAES) panel has completed review of three proposals, and each proposal received a favorable evaluation. Table 1 contains a general summary of these experiments. At the present time, the panel is reviewing three new proposals.

2. Radio Science Operations

Since May 1, 1970, five radio science experiments have been supported at DSN facilities. Outlines of these experiments are given below:

a. SCO-XR1 observations. As reported in SPS 37-63, Vol. II, pp. 98–100, Dr. C. S. Boyer, University of California at Berkeley, submitted a proposal to the RAES panel to conduct extensive observations of the 13-cm radio emissions from SCO-XR1, an X-Ray source in the constellation Scorpio. The objective of the experiment is to determine if the radio emissions from this object are variable and correlated with previously observed optical variability. The observations were conducted in May and June of 1970 at DSS 14.

Since the radio emissions from this object are very weak, a stabilized radiometer (*Subsection VI-C*) developed by the Telecommunications Division was employed, which considerably enhanced the probability of a successful experiment.

b. Trans-Pacific very long baseline interferometer experiment. On June 13, 1970, Drs. A. Moffet (associated with Owens Valley Radio Observatory) and D. Robertson (principal officer of Space Research Group, Weapons Research Establishment, Adelaide, Australia) conducted another 24-h trans-Pacific very long baseline interferometer (VLBI) experiment between the Goldstone DSCC in California and the Canberra DSCC in Australia (SPS 37-62, Vol. II, pp. 125–127). This was the first observation under the auspices of the RAES panel (SPS 37-63, Vol. II, pp. 99–100). During the experiment, which was supported by DSSs 12, 14, and 42, data was taken on 29 radio sources. The data tapes are now being evaluated by the experimenters.

Table 1. Approved experiments

Experimenter	Affiliation	Objective	Comments
Dr. A. Moffet	Owens Valley Radio Observatory	To detect pulsar emissions at 7840 MHz	Experiment completed
Dr. A. Moffet	Owens Valley Radio Observatory	To study polarization characteristics of pulsar emissions at 2295 MHz	Scheduled for Fall 1970
Dr. B. Burke	Massachusetts Institute of Technology	To conduct a precision test of general relativity by use of very long baseline interferometry	Scheduled for October 1–15, 1970
Dr. T. Clark	Goddard Space Flight Center		
Dr. R. Goldstein	Jet Propulsion Laboratory		
Dr. A. Rogers	Haystack Radio Observatory		
Dr. I. Shapiro	Massachusetts Institute of Technology		

c. X-Band pulsar observations. An experimental low-noise X-band feed cone (SPS 37-62, Vol. II, p. 80) was part of the tricone feed assembly installed at DSS 14 early this year. An experiment to study X-band emissions of pulsars using this new experimental capability, was conducted on July 4, 1970, and data was obtained. The principal investigator was Dr. A. Moffet.

d. X-Band planetary radar. Two X-band planetary radar experiments were conducted jointly with MIT's Lincoln Laboratory in May, June, and July 1970. The first experiment attempted to obtain a radar return from Venus and from Callisto, a satellite of Jupiter. Lincoln Laboratory's Haystack Facility radiated 300 kW at the target and DSS 14 attempted to receive the return. The received data were transmitted via microwave to DSS 13 for post-reception processing. Runs were made on six evenings in May and June. The data has not been fully evaluated at this time.

The second experiment was an attempt to radar-range the planet Mercury near superior conjunction. This experiment was designed as a test of general relativity. The observations, utilizing the same facilities as the Callisto experiment, were conducted over a 3-wk period in June and July 1970. The data is still being evaluated.

This combination of facilities represented the most powerful planetary radar system employed to date in the world.

B. Computerized Receiver and Telemetry SNR Predictions Program, *W. Porché*

1. Introduction

A computerized program has been developed to provide in one document, predicted receiver signal strength, receiver margin, predicted telemetry ST_b/N_0, and parity error rate or bit error rate. The program Receiver and Telemetry SNR Predictions is used to generate predicts for *Mariner* and *Pioneer* spacecraft.

The receiver and telemetry SNR predictions program is presently being used by the real and non-real time analysis groups, operations engineering, and system data analysis personnel. These predictions are sent to the various tracking stations and are used by the stations to ascertain normal operation of the telemetry system.

2. Operation

Figure 1 is a listing of the program output. This listing is obtained through the following data inputs and program functions:

a. Identification input. The ID card provides the following inputs:

Modulation index
System noise temperature

Receiver bandwidth
Antenna polarizer in or out
85- or 210-ft network
Spacecraft number
Month to start predicts
Day of month to start predicts
Year to start predicts
Pass number to start predicts
Number of days to run predicts
Mission (*Pioneer* or *Mariner*)
Telemetry mode (single or dual station)

b. Data input. Data input is by cards providing day of year and spacecraft range in kilometers. The range data is obtained from the spacecraft trajectory printout.

c. Temperature input. The temperature card is used when dual-station telemetry is to be output. The system noise temperature is provided for stations operating in the dual-telemetry mode.

d. Range interpolation. The program uses a four-point La Grange interpolation routine. The day of year and spacecraft range is input to the program and interpolated to provide range for daily computations. The program will take up to 99 inputs and can compute predicts for 999 days.

e. Prediction computations. The equation for free-space loss (Ref. 1) is

$$L_{fs} = 92.45 + 20 \log (f) + 20 \log (d)$$

PIONEER 8 RECEIVER AND TELEMETRY SNR PREDICTIONS. 85 FOOT NETWORK

DUAL STATION-COMBINED TELEMETRY MODE

RECEIVER BANDWIDTH = 3.0 HERTZ
SYSTEM NOISE TEMPERATURE:
DSS 11 = 40.00 DEGREES K
DSS 12 = 17.00 DEGREES K
DSS 61 = 42.00 DEGREES K
DSS 62 = 30.00 DEGREES K
POLARIZER IN OR OUT = IN
BAD PER = ******
SNR=ST/NO
MODULATION INDEX = 51.60 DEGREES

DATE	DAY	PASS NO.	RECEIVER SIGNAL STRENGTH	DEEP SPACE STATIONS	PARITY ERROR RATE(PER) 8BPS	16BPS	64BPS	TELEMETRY SNR 8BPS	16BPS	64BPS
6- 1-70	152	902	-164.10	DSS 11/12	.00000	.00000	******	13.36	10.78	5.03
				DSS 61/62	.00000	.00568	******	11.69	9.12	3.37
6- 2-70	153	903	-164.11	DSS 11/12	.00000	.00000	******	13.34	10.77	5.01
				DSS 61/62	.00000	.00585	******	11.67	9.10	3.35
6- 3-70	154	904	-164.13	DSS 11/12	.00000	.00000	******	13.32	10.75	5.00
				DSS 61/62	.00000	.00602	******	11.66	9.09	3.33
6- 4-70	155	905	-164.15	DSS 11/12	.00000	.00000	******	13.31	10.73	4.98
				DSS 61/62	.00000	.00619	******	11.64	9.07	3.32
6- 5-70	156	906	-164.16	DSS 11/12	.00000	.00000	******	13.29	10.72	4.96
				DSS 61/62	.00000	.00636	******	11.62	9.05	3.30
6- 6-70	157	907	-164.18	DSS 11/12	.00000	.00000	******	13.27	10.70	4.95
				DSS 61/62	.00000	.00653	******	11.61	9.03	3.28
6- 7-70	158	908	-164.20	DSS 11/12	.00000	.00000	******	13.25	10.68	4.93
				DSS 61/62	.00000	.00671	******	11.59	9.02	3.27
6- 8-70	159	909	-164.22	DSS 11/12	.00000	.00000	******	13.24	10.66	4.91
				DSS 61/62	.00000	.00688	******	11.57	9.00	3.25
6- 9-70	160	910	-164.23	DSS 11/12	.00000	.00000	******	13.22	10.65	4.90
				DSS 61/62	.00000	.00706	******	11.56	8.98	3.23
6-10-70	161	911	-164.25	DSS 11/12	.00000	.00000	******	13.20	10.63	4.88
				DSS 61/62	.00000	.00724	******	11.54	8.97	3.22
6-11-70	162	912	-164.27	DSS 11/12	.00000	.00000	******	13.19	10.61	4.86
				DSS 61/62	.00000	.00741	******	11.52	8.95	3.20

Fig. 1. Two-station dual telemetry predictions

where

L_{fs} = free space loss, dB

f = received frequency, GHz

d = spacecraft range, km

The predicted received signal strength is computed as follows:

SS = (spacecraft gain) + (ground antenna gain) − (modulation loss) − L_{fs} − (antenna pointing loss = 0.5 dB)

where the spacecraft gain

PN-6 = +48.9 dB

PN-7 = +48.2 dB

PN-8 = +49.1 dB

PN-9 = +49.1 dB

Ground antenna gain for

85-ft antenna = +53.3 dB

210-ft antenna = +61.4 dB

Modulation loss is 10 log (modulation index); the modulation index is 51.6 deg for PN-6 through 9. The modulation index for the *Mariner* spacecraft is variable. The receiver margin M, (Ref. 2) is computed by

$$M = SS - \left[10 \log\left(\frac{KT_s\,(BW)}{10^3}\right)\right]$$

The ST_b/N_0 of the telemetry (Ref. 2) is computed by

$$DP = SS + 20 \log \quad \tan\,(\text{modulation index})$$

$$SNR = \frac{DP}{KT_s(BR)} - \text{loss}$$

where

M = receiver margin, dB

DP = telemetry data power, dB

SS = predicted receiver signal strength, dBmW

SNR = telemetry $\dfrac{ST_b}{N_0}$

K = Boltzmann constant

T_b = 1/bit rate

T_s = system noise temperature, °K

BW = receiver bandwidth, Hz

$KT_s = N_0$ = noise

BR = telemetry bit rate

Loss = fixed system loss depending on bit rate

Output is available for single-station telemetry or the dual-station combined-telemetry mode of operation. Dual-station combined-telemetry tracking by DSSs 11 and 12, and DSSs 61 and 62 have been experimentally tried and determined to have a definite improvement in telemetry performance. The theoretical maximum improvement for combined telemetry of two stations of equal gain is 3 dB (SPS 37-63, Vol. III, pp. 116–120). The computer program computes the dual-station combined-telemetry SNR as follows:

$$SNT = SNR_l + 10 \log\left(1 + \frac{T_l}{T_h}\right)$$

where

SNT = total SNR for combined telemetry

SNR_l = SNR of lowest system noise temperature station

T_l = lowest system noise temperature, °K

T_h = highest system noise temperature, °K

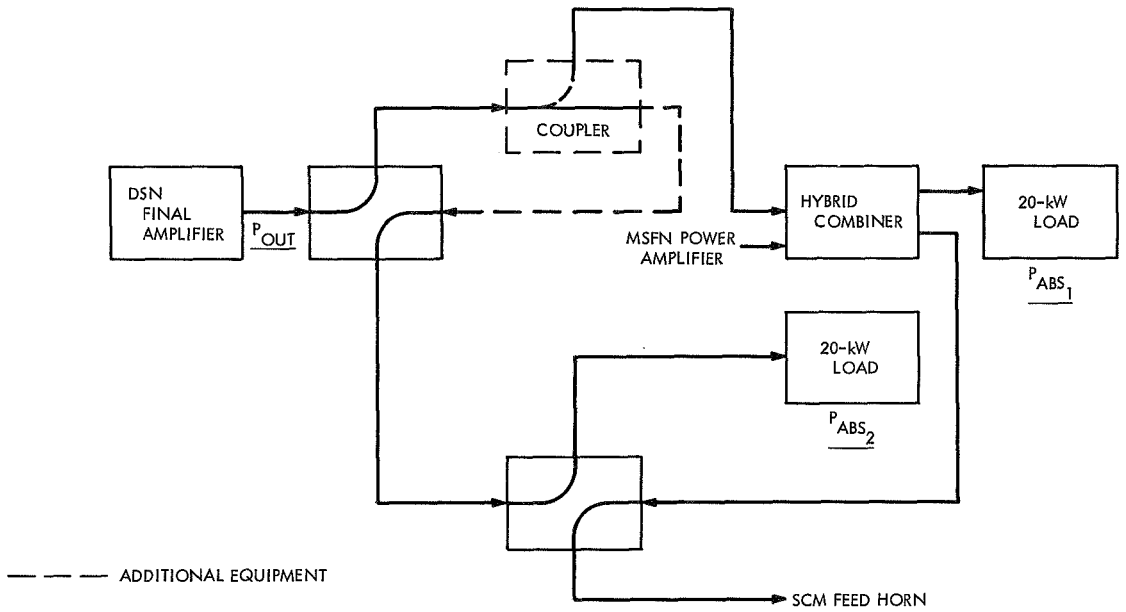
$$T_l \leq T_h \text{ and } \left(1 + \frac{T_l}{T_h}\right) \leq 2$$

Figure 1 provides an illustration of the output listing for dual-station combined telemetry. The predictions are printed for each day at 00:00 GMT. Currently, these predictions are being used by all stations in the Deep Space Network.

References

1. *Reference Data for Radio Engineers,* pp. 34-3, McGraw-Hill Book Co. Inc., New York, 1949.

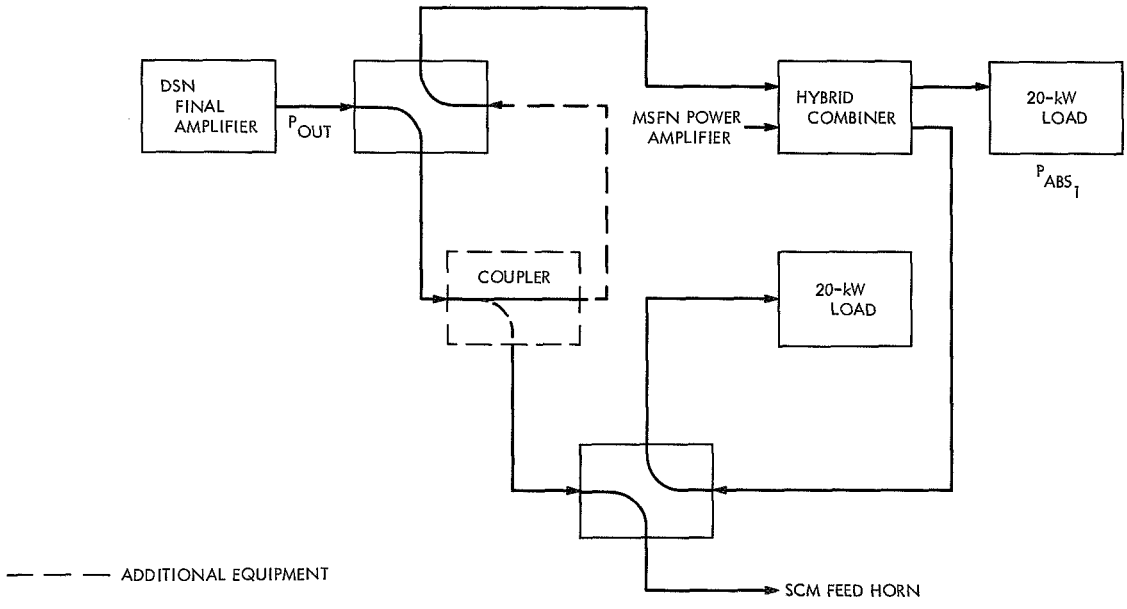
2. *Deep Space Network/Flight Project Interface Design Handbook,* DSN Standard Practice 810-5, pp. 2-32 to 2-78, Jet Propulsion Laboratory, Pasadena, Calif., March 1, 1970.



COUPLER FACTOR, dB	P_{ABS_1}, W	P_{ABS_2}, W	P_{ANT}, W
30	10	19,980	10
20	100	19,800	100
10	1000	18,000	1000

P_{ANT} = RADIATED POWER
MAXIMUM P_{ANT} = 20 kW
P_{OUT} = 20 kW (CONSTANT)

Fig. 1. Directional coupler in an 85-ft-diam antenna mutual station using combiner in attenuated RF path



COUPLER FACTOR, dB	P_{ABS_1}, W	P_{ABS_2}, W	P_{ANT}, W
30	9990	9990	20
20	9900	9900	200
10	9000	9000	2000

P_{ANT} = RADIATED POWER
MAXIMUM P_{ANT} = 10 kW
P_{OUT} = 20 kW (CONSTANT)

Fig. 2. Directional coupler in an 85-ft-diam antenna mutual station

C. Low Transmitted Power Operation,

R. L. Riggs

1. Introduction

This report describes a method by which the radiated output power level from a DSIF antenna may be adjusted while maintaining the transmitter at constant power and bandwidth levels. The constant power output from the transmitter is split into appropriate levels between a dissipative load and the antenna feed horn.

Two devices accomplish the splitting. One device is basically a fixed high-power directional coupler (or combination of couplers), and the other device is a high-power microwave power divider that has a capability of being continuously variable.

Current instrumentation techniques for monitoring the output power level are not changed with either device, and the overall accuracy of the system is degraded only by the tolerances in the calibration of the added device(s).

2. Description

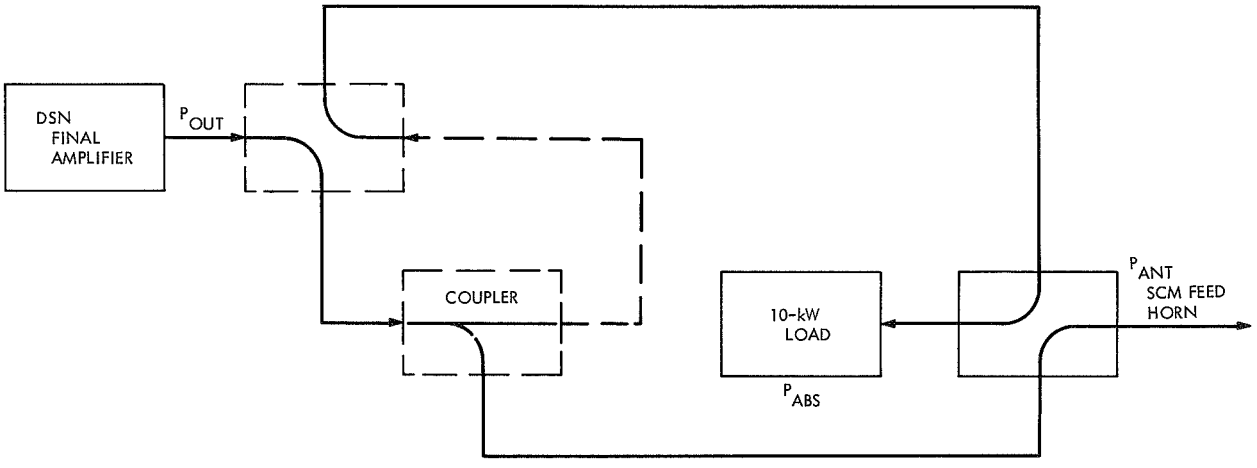
a. Fixed high-power directional couplers. In an 85-ft-diam antenna mutual station there are many possible configurations with directional couplers. These different configurations will give various combinations of antenna output power. Figures 1 and 2 illustrate two possible configurations of a mutual station, and Fig. 3 is one of a standard station. It should be noted that the 85-ft mutual station has a 20-kW transmitter and the standard station has only a 10-kW capability.

It can be seen that the output power range is dependent upon the use and coupling factor of the directional coupler. In all of the illustrations, it is assumed that the final amplifier output power is held constant at its maximum rated level.

b. High-power microwave power divider. A major problem with the use of directional couplers just described is the lack of flexibility in adjustment of output power. A directional coupler is designed to have a fixed coupling factor, and the antenna output power may vary in a continuous fashion by varying the final amplifier output power. This may lead to a degradation of transmitter stability and bandwidth characteristics.

A new device is described[1] where the power is split in any proportion between two output ports. This new device may be installed in place of the fixed directional

[1]Kolbly, R. B., *High Power Microwave Divider Concept*, May 1970 (JPL internal document).



— — — ADDITIONAL EQUIPMENT

MAXIMUM P_{ANT} = 10 kW

P_{OUT} = 10 kW (CONSTANT)

COUPLER FACTOR, dB	P_{ABS}, W	P_{ANT}, W
30	9990	10
20	9900	100
10	9000	1000

Fig. 3. Directional coupler in a standard 85-ft-diam antenna station

couplers previously described. If the device performs as described in the article, its most useable range would certainly be limited to approximately 60 deg of rotation which yields a maximum attenuation of −6 dB. Reference to Fig. 4 shows that the resolution (decibels of attenuation per degree of rotation) is degraded when the rotation exceeds 60 deg. That is, the sensitivity of the device is such that it should be limited to an operational maximum of 0.25 dB/deg of rotation. Given a maximum useable attenuation range of −6 dB, clearly the device is incapable of providing the degree of attenuation necessary to reduce 10 or 20 kW to 10 or 20 W. However, the device should prove to be very useful when used in conjunction with a directional coupler as illustrated in Fig. 5. The combination of a fixed coupler plus the power divider should provide a very accurate and variable (over a 6-dB range) method of obtaining the low power level desired at the antenna.

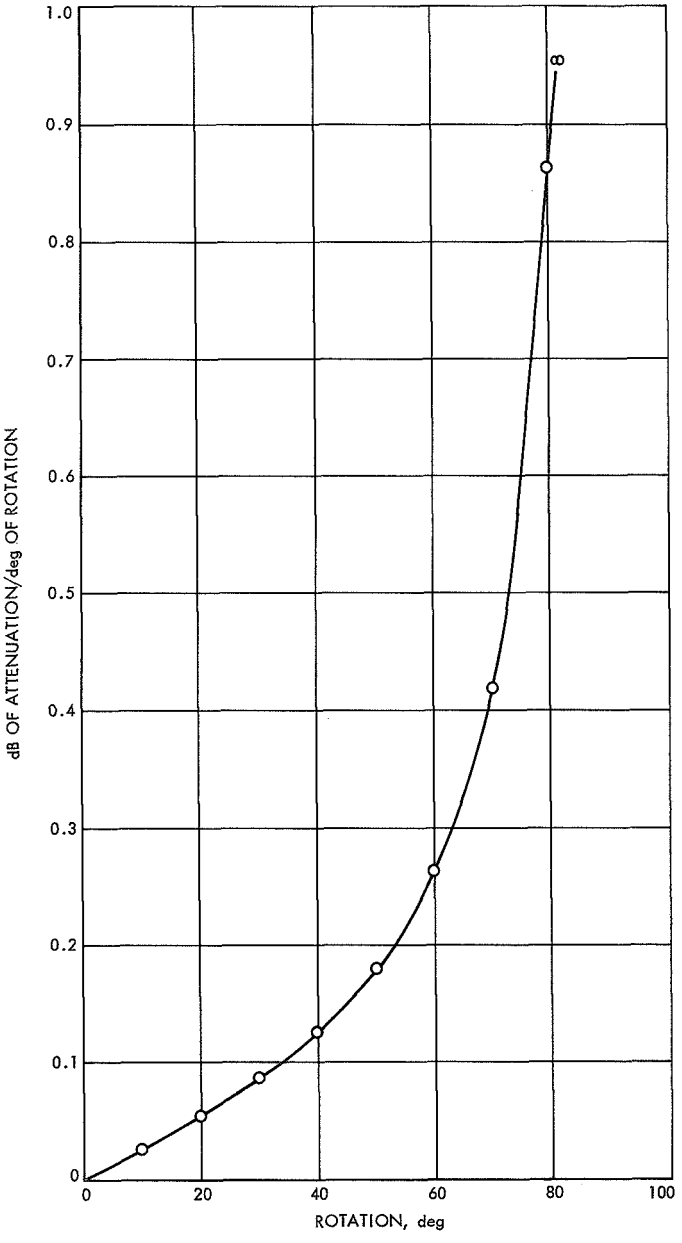


Fig. 4. Decibels of attenuation per degree of rotation for high-power microwave power divider

3. Summary

The configurations proposed above may be summarized in Table 1, which shows the power radiated from the horn using two different directional couplers; with and without the microwave power divider. In Table 1, it is assumed that the maximum attenuation obtained from the power divider is −6 dB. In actual practice, however, this may be increased to as much as −10 dB at 72 deg of rotation, but with a corresponding decrease in resolution, as the curve in Fig. 6 indicates.

It would appear that the optimum configuration for either the 85-ft mutual station or the standard station would be the use of a 20-dB coupler, plus a power divider, as shown in Figs. 1 and 3, respectively. This gives a 25- to 100-W variable output or, with a corresponding degradation in tolerance a variable output of 10 to 100 W.

Consideration should be given to the development of the high-power microwave power divider as described in the attached article. With the proper electromechanical design, the capability to accurately control transmitted power by remote control or automatically by computer would be possible.

Table 1. Station configuration summary

Station configuration	Power at antenna feed horn, W			
	Coupler only		Coupler plus divider	
	30 dB	20 dB	30 dB	20 dB
85-ft-diam antenna mutual station (20 kW)				
Fig. 1 configuration	10	100	2.5 to 10	25 to 100
Fig. 2 configuration	20	200	5 to 20	50 to 200
85-ft standard station (10 kW)				
Fig. 3 configuration	10	100	2.4 to 10	25 to 100

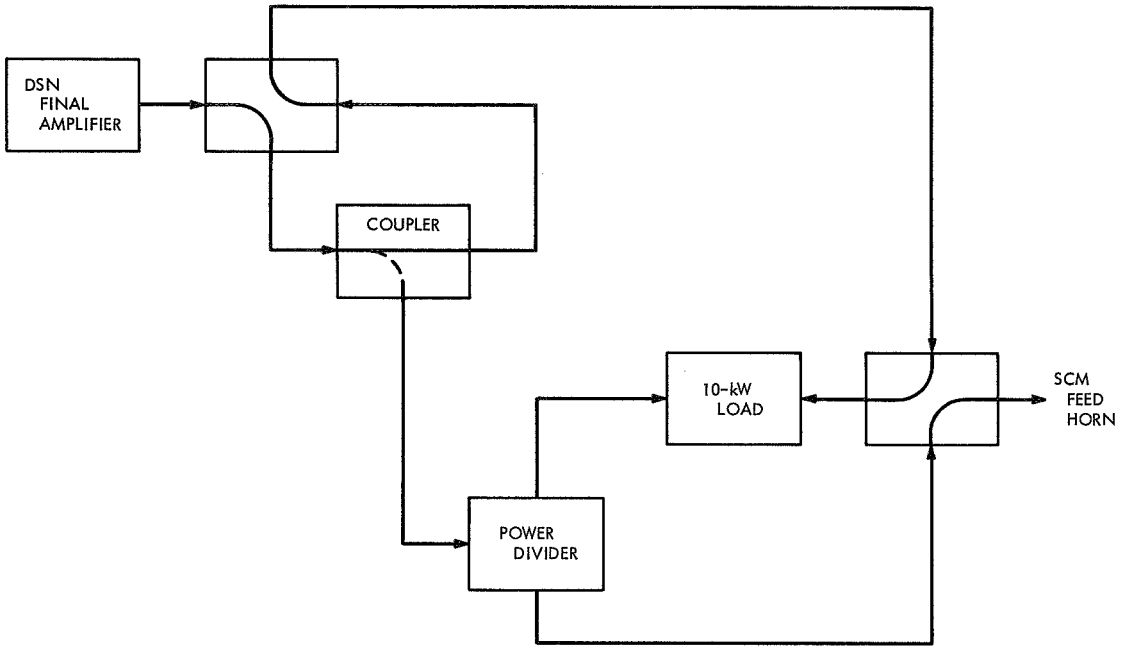


Fig. 5. Directional coupler plus power divider

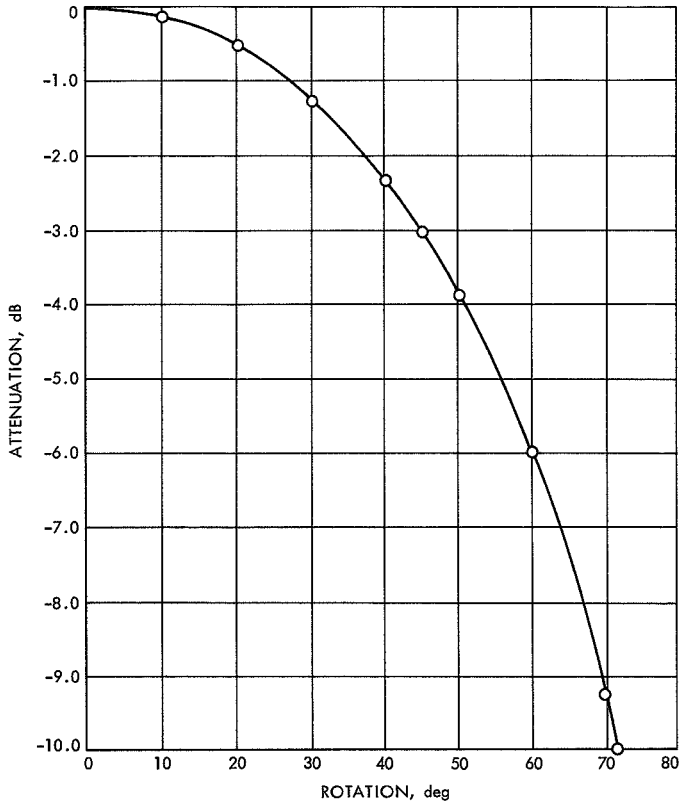


Fig. 6. Attenuation of input power versus degrees of rotation for high-power microwave power divider

D. A New Tropospheric Range Refraction Model,

A. L. Berman

1. Introduction

Atmospheric (tropospheric) refraction can be easily separated into two components. Most analysis lumps the two together; in this report they will be dealt with separately:

$$N = \text{refractivity}$$

$$N = N_d + N_w$$

where

$$N_d = \text{dry component}$$

$$N_w = \text{wet component}$$

Previously, the dry component of range and range rate refraction (signal retardation component) has been assumed to be a linear function of surface refractivity. This appears to be a result of assuming an isothermal atmospheric model and then failing to consider the functional dependence required of the "scale height" by that assumption. Consideration of the hydrostatic equation and the gas law lead inescapedly to the conclusion that range refraction (at least at reasonable elevation angles) is linearly dependent upon surface pressure and not surface refractivity. The derivation of the functional basis of the dry component is presented in *Subsection 2*; in *Subsection 3*, a method of integrating the wet component is presented; and finally, *Subsection 4* describes the errors which would result from the use of the new model—these errors being then compared to current model errors.

2. Dry-Term Derivation

In general, two tropospheric range refraction effects are recognized, namely, decrease in signal propagation velocity and increase in geometrical refracted signal path. The ratio of effect of the latter to the former can be roughly represented as (Ref. 1):

$$\frac{(\Delta R)_R}{(\Delta R)_v} \approx B \cos\theta$$

where

$$(\Delta R)_R = \Delta_R \text{ due to geometrical path}$$

$$(\Delta R)_v = \Delta_R \text{ due to propagation velocity}$$

$$B = \text{inverse scale height, } \text{km}^{-1} \approx 0.1$$

$$\theta = \text{elevation angle}$$

Since the value of $(\Delta R)_R$ can be seen to vary from 0 to 0.1 of the value of $(\Delta R)_v$, it seems reasonable to limit the discussion to the effect of signal velocity retardation only.

The index of refraction for the troposphere is generally considered to be comprised of two separate components, sometimes referred to as the dry (main) and wet components:

$$n = 1 + 10^{-6}N$$

$$N = N_d + N_w$$

$$N_d = 77.6\,\frac{P}{T}$$

$$N_w = 77.6\,(4810)\,\frac{(RH)e_s}{T^2}$$

where

$$T = \text{temperature, °K}$$

$$n = \text{index of refraction}$$

$$N = \text{refractivity}$$

$$P = \text{pressure, mbar}$$

$$RH = \text{relative humidity (\% of 1.0)}$$

$$e_s = \text{saturation vapor pressure, mbar}$$

A cursory examination of the two expressions shows N_w to be in the vicinity of 10% of the value of N_d; for example, in the following case where

$$T = 273°\text{K}$$

$$P = 1000 \text{ mbar}$$

$$RH = 1.0$$

$$e_s(T) = 6.11 \text{ mbar}$$

$$\frac{N_w}{N_d} = 0.1077$$

Thus, it would appear advantageous initially to limit the analysis to the dry component of refractivity.

We begin by considering the physical signal path range as the troposphere is traversed:

$$R = \int_0^s cdt = \int_0^s ds$$

We then consider an apparent refracted range R' with signal velocity v:

$$v = c/n$$

$$R' = \int_0^s cdt' = \int_0^s c\,\frac{ds}{v} = \int_0^s nds$$

so that the range correction ΔR is

$$\Delta R = R' - R = \int_0^s nds - \int_0^s ds$$

$$= \int_0^s (n-1)\,ds$$

$$= 10^{-6}\int_0^s Nds$$

Relating ds to height above the observation point z and elevation angle θ (at least for high elevation angles) we obtain

$$ds \approx \frac{dz}{\sin\theta} = \csc\theta\,dz$$

so that

$$\Delta R \approx 10^{-6}\int_0^z N \csc\theta\,dz$$

However, since what we are really interested in is the dependence of ΔR upon $N(z)$ we will ignore spherical geometry considerations:

$$\Delta R \approx 10^{-6}\csc\theta\int_0^z N(z)\,dz$$

and define

$$\Delta r = \frac{\Delta R}{10^{-6}\csc\theta} \approx \int_0^z N(z)\,dz$$

so that we can evaluate the functional dependence of Δr upon $P_0, T_0, \cdots$, etc.

Typically, in the past this has been handled by assuming that

$$N(z) = N_s \exp(-Bz)$$

where

N_s = surface refractivity

B = constant, inverse scale height in $km^{-1} \approx 0.1$

The total (elevation-independent) refraction effect (considering that the tropospheric-type refraction is nil by 75 km) becomes

$$\Delta r = \int_0^{75} N_s \exp(-Bz)\,dz$$

$$\approx \int_0^{\infty} N_s \exp(-Bz)\,dz$$

$$\approx N_s/B$$

However, this does not seem to be an entirely adequate treatment of the problem. The assumption of an exponential decay refractivity is rooted in the assumption of an isothermal, or constant temperature, troposphere. If we consider this to be true, we would have:

$$T(z) = T_0$$

$$P(z) = P_0\,e^{-Bz}$$

and indeed

$$N(z) = A\frac{P(z)}{T(z)} = \left(A\frac{P_0}{T_0}\right)e^{-Bz}$$

$$= N_s\,e^{-Bz}$$

However, it is easy to show (Ref. 2) that for this to be true, we require

$$B = g/RT_0$$

where

g = gravitational force

R = gas constant

$$\frac{g}{R} = 34.1°\text{C/km}$$

Integrating this expression, we have

$$\Delta r = \int_0^{75} N(z)\,dz \approx \int_0^{\infty} N(z)\,dz$$

$$\Delta r \approx \int_0^\infty A \frac{P_0}{T_0} \exp(-gz/RT_0)\, dz$$

$$\approx \left(A \frac{P_0}{T_0}\right)\left[-\left(\frac{RT_0}{g}\right)\exp(-gz/RT_0)\right]_0^\infty$$

$$\approx A\left(\frac{R}{g}\right)P_0$$

which would lead to the unmistakable conclusion that range refraction is basically linearly dependent upon surface pressure and not surface refractivity.

However, an isothermal atmosphere is a rather poor assumption; a much better idea would be to assume a standard atmosphere profile. For instance, a typical profile would be for a station at a height of h_0 above sea level:

$$0 \leq z \leq (11 - h_0)\ \text{km}$$

$$T(z) = T_0 - \gamma z$$

$$P(z) = P_0\left(\frac{T(z)}{T_0}\right)^{g/R\gamma}$$

$$\gamma = \gamma(T_0) = \left(\frac{T_0 - T_1}{11 - h_0}\right)$$

and

$$(11 - h_0)\ \text{km} \leq z \leq (21 - h_0)\ \text{km}$$

$$T(z) = T_1 = 216.65°\text{K}$$

$$P(z) = P_0\left(\frac{T_1}{T_0}\right)^{g/R\gamma}\exp\left(-\frac{z - [11 - h_0]}{RT_1}g\right)$$

We would then have for the refraction contribution

$$\Delta r = \int_0^z N(z)\, dz$$

$$= \int_0^{11-h_0} A\frac{P_0}{T_0}\left(\frac{T_0 - \gamma z}{T_0}\right)^{g/R\gamma - 1} dz$$

$$+ \int_{11-h_0}^{21-h_0} A\frac{P_0}{T_0}\left(\frac{T_1}{T_0}\right)^{g/R\gamma}$$

$$\times \exp\left(-\frac{z - [11 - h_0]}{RT_1}g\right) dz$$

We allow $(21 - h_0)\text{km} \to \infty$ (because of the small contribution above 21 km) so that

$$\Delta r \approx \frac{AP_0}{T_0}\int_0^{11-h_0}\left(\frac{T_0 - \gamma z}{T_0}\right)^{g/R\gamma - 1} dz$$

$$+ \frac{AP_0}{T_0}\int_{11-h_0}^\infty \left(\frac{T_1}{T_0}\right)^{g/R\gamma}\exp\left(-\frac{z - [11 - h_0]}{RT_1}g\right) dz$$

$$\approx AP_0\left(\frac{R}{g}\right)\left[1 - \left(\frac{T_1}{T_0}\right)^{g/R\gamma}\right] + AP_0\left(\frac{R}{g}\right)\left(\frac{T_1}{T_0}\right)^{g/R\gamma}$$

$$\approx AP_0\left(\frac{R}{g}\right)\left[1 - \left(\frac{T_1}{T_0}\right)^{g/R\gamma} + \left(\frac{T_1}{T_0}\right)^{g/R\gamma}\right]$$

$$\approx AP_0\left(\frac{R}{g}\right)$$

Finally, it can be shown quite easily that this result will occur regardless of the temperature/pressure distribution by merely examining the hydrostatic equation and the perfect gas law:

$$dp = -\rho g\, dz \qquad \text{(hydrostatic equation)}$$

and

$$P = \rho RT \qquad \text{(gas law)}$$

$$N(z) = A\frac{P_{(z)}}{T_{(z)}} = A\frac{\rho RT}{T} = A\rho R$$

$$= A\left(-\frac{1}{g}\frac{dp}{dz}\right)R$$

$$\Delta r \approx \int_0^\infty N(z)\, dz$$

$$\approx -\int_0^\infty \left(A\frac{1}{g}\frac{dp}{dz}\right)R\, dz$$

$$\approx -\int_{P_0}^0 \frac{AR}{g}\, dp = \frac{AR}{g}P_0$$

Or, once again:

$$\Delta r \approx A\left(\frac{R}{g}\right)P_0$$

3. Integration of Wet Term

In *Subsection 2*, it was indicated that a scaled range refraction Δr could be calculated as

$$\Delta r = \int_z^0 N(z)\, dz$$

where N could be split into a wet and dry component:

$$N = N_d + N_w$$

The quantity $(\Delta r)_d$ was then evaluated according to

$$(\Delta r)_d = \int_0^z N_d(z)\, dz$$

We now wish to evaluate a similar expression for the wet component:

$$(\Delta r)_w = \int_0^z N_w(z)\, dz$$

From SPS 37-45, Vol. IV, pp. 270–276, we have the following expression for N_w:

$$N_w = \left(\frac{(RH)C_1C_2}{T^2}\right) 10^{[(K_1T-K_2)/(T-C)]}$$

$$= \left(\frac{(RH)C_1C_2}{T^2}\right) \exp\left[\frac{(K_1T - K_2)\ln(10)}{T - C}\right]$$

$$= \left(\frac{(RH)C_1C_2}{T^2}\right) \exp\left(\frac{AT - B}{T - C}\right)$$

where

$C_1 = 77.6$

$C_2 = 29341.0$

T = temperature, °K

$A = K_1 \ln(10) = 7.4475 \ln(10)$

$B = K_2 \ln(10) = 2034.28 \ln(10)$

$C = 38.45$

RH = relative humidity (% of 1.0)

To perform (approximately) the integration

$$(\Delta r)_w = \int_0^z N_w(z)\, dz$$

we need two relationships, namely:

$$T = T(z)$$

$$RH = RH(z)$$

and an upper bound for z. The most difficult task is to determine how relative humidity might depend upon the surface value and altitude. Perhaps the best assumption would be that

$$RH(z) = RH_s$$

where

RH_s = the relative humidity at surface

At least this would be reasonable at low heights, where by far the greatest contribution to refraction (wet component) occurs. As concerns temperature, we can assume a constant negative lapse rate to the (approximate) tropopause ($z \approx [11 - h_0]$km):

$$T(z) = T_0 - \gamma z$$

$$\gamma = \gamma(T_0) = \frac{T_0 - T_1}{11 - h_0}$$

where

T_0 = surface temperature, °K

z = height above station, km

h_0 = station height, km

$T_1 = 216.65$°K

We can now transform our integral accordingly:

$$dt = -\gamma\, dz$$

$$(\Delta r)_w = \int_0^z N_w(z)\, dz$$

$$= \int_{T_0}^{T_1} \frac{(RH_s)C_1C_2}{T^2} \exp\left(\frac{AT - B}{T - C}\right)\left(-\frac{dT}{\gamma}\right)$$

$$= -\left\{\frac{C_1C_2(RH)_s}{\gamma(T_0)}\right\} \int_{T_0}^{T_1} \frac{1}{T^2} \exp\left\{\frac{AT - B}{T - C}\right\} dT$$

Thus we wish to evaluate

$$F(T_0) = \int_{T_0}^{T_1} \frac{dT}{T^2} \exp\left(\frac{AT - B}{T - C}\right)$$

If we make the substitution

$$Y = \frac{AC - B}{T - C}$$

we would then have

$$\exp\left(\frac{AT - B}{T - C}\right) = \exp(A + Y) = \exp(A)\exp(Y)$$

$$T = \frac{AC - B + CY}{Y}$$

$$dT = \frac{B - AC}{Y^2}\, dY$$

$$\frac{dT}{T^2} = \frac{B - AC}{Y^2}\, dY \left(\frac{Y^2}{[AC - B + CY]^2}\right)$$

and letting

$$Y_0 = \frac{AC - B}{C}$$

$$\frac{dT}{T^2} = \frac{1}{(B - AC)}\left(\frac{dY}{(1 + Y/Y_0)^2}\right)$$

so that

$$F(T_0) = \int_{T_0}^{T_1} \frac{dT}{T^2} \exp\left(\frac{AT - B}{T - C}\right)$$

$$= \left(\frac{\exp(A)}{B - AC}\right) \int_{Y(T_0)}^{Y(T_1)} \frac{\exp(Y)\, dY}{(1 + Y/Y_0)^2}$$

For the most extreme possible range of temperatures, we would have for the variation in the y-dependent terms:

$$216 < T < 316$$

$$-22.67 \simeq Y \simeq -14.50$$

$$1.4 \times 10^{-10} \lesssim \exp(Y) \lesssim 5 \times 10^{-7}$$

$$1.22 \lesssim (1 + Y/Y_0) \lesssim 1.15$$

Since the variation in the denominator is almost nil compared to the variation in the exponential term, we can approximate the integral:

$$\int_{Y(T_0)}^{Y(T_1)} \frac{\exp(Y)\, dY}{(1 + Y/Y_0)^2} \approx \frac{\exp(Y)}{(1 + Y/Y_0)^2}\Bigg|_{Y(T_0)}^{Y(T_1)}$$

Furthermore, since

$$\frac{\exp[Y(T_1)]}{\exp[Y(T_0)]} \approx 3 \times 10^{-4}$$

we can ignore the upper limit:

$$F(T_0) \approx -\frac{\exp[Y(T_0)]}{[1 + Y(T_0)/Y_0]^2}\left(\frac{\exp(A)}{B - AC}\right)$$

$$\approx -\frac{\exp\left(\frac{AC - B}{T_0 - C}\right)}{\left(1 + \frac{C}{T_0 - C}\right)^2}\left(\frac{\exp(A)}{B - AC}\right)$$

$$\approx -\frac{\left(1 - \frac{C}{T_0}\right)^2}{(B - AC)} \exp\left(\frac{AT_0 - B}{T_0 - C}\right)$$

and finally

$$(\Delta r)_w \approx \left[\frac{C_1 C_2 (RH)_s}{\gamma(T_0)}\right]\left[\frac{\left(1 - \frac{C}{T_0}\right)^2}{B - AC}\right] \exp\left(\frac{AT_0 - B}{T_0 - C}\right)$$

This result, when combined with the expression for $(\Delta r)_d$ presented in *Subsection 2* should give the total functional dependence of range (and range rate) refraction upon the surface variables P_0, T_0, and $(RH)_s$.

4. Error Analysis

a. Combined refraction model. In *Subsection 2* it was shown that the current so-called exponential refractivity model was predicated upon a faulty assumption, and a more correct formulation for range refraction was detailed. We are here concerned with the two separate, additive components of refractivity, the so-called wet and dry terms:

$$N(z) = N_d(z) + N_w(z)$$

with

$$N_d(z) = -\frac{AR}{g}\left(\frac{dP}{dz}\right)$$

and

$$N_w(z) = \frac{(RH)C_1C_2}{T^2}\,10\left(\frac{K_1T - K_2}{T - C}\right)$$

and where

$A = 77.6$

$R =$ gas constant

$g =$ gravity

$p =$ pressure

$RH =$ relative humidity

$C_1 = 77.6$

$C_2 = 29341.0$

$T =$ temperature

$K_1 = 7.4475$

$K_2 = 2034.28$

$C = 38.45$

Integration of the major, or dry, term, as shown in *Subsection 2*, is quite simple, namely:

$$\int_0^\infty N_d(z)\,dz = AP_0\left(\frac{R}{g}\right)$$

where

$P_0 =$ surface pressure

Whereas, integration of the wet term is far more complicated, but, with certain assumptions, can be shown (*Subsection 3*) to be approximately equal to

$$\int_0^\infty N_w(z)\,dz \cong \left[\frac{C_1C_2(RH_s)}{\gamma}\right]\left[\frac{\left(1 - \frac{C}{T_0}\right)^2}{B - AC}\right] \times \exp\left(\frac{AT_0 - B}{T_0 - C}\right)$$

where

$\gamma =$ temperature lapse rate

$(RH)_s =$ surface relative humidity

$T_0 =$ surface temperature

When combined, we have for the total zenith range refraction, as a function of surface meteorological parameters and temperature lapse rate:

$$(\Delta r) \cong AP_0\left(\frac{R}{g}\right) + \left[\frac{C_1C_2(RH)_s}{\gamma}\right]\left[\frac{\left(1 - \frac{C}{T_0}\right)^2}{B - AC}\right] \times \exp\left(\frac{AT_0 - B}{T_0 - C}\right)$$

which can be compared to the current expression:

$$(\Delta r) = \frac{N_s}{B} = AP_0\left(\frac{1}{T_0B}\right)$$

The assumptions required to perform the integration of the wet term are

$$T(z) = T_0 - \gamma z$$

and

$$RH(z) = (RH)_s$$

Thus it should be kept in mind that the expression for the integrated wet term is accurate only to the extent that these assumptions prove true. The accuracy ramifications embodied in the above assumptions will be examined in detail. Preparatory to this, however, it seems timely to present a description of prevailing meteorological conditions.

b. Description of temperature profiles. The region for which we are interested in defining the functional dependence of temperature upon altitude is from the station height to the tropopause, or roughly, 0–11 km. At the tropopause, the temperature is approximately −55°C, and by this temperature wet refractivity is nil. It is fortunate that many of the station sites are meteorologically similar, i.e., temperate, semi-arid, and not in proximity to any large bodies of water. For climates

such as these, we can identify two basic types of temperature profiles; a summer-type profile (warm and usually clear) and a winter-type profile (cold and often cloudy). The summer-type profile is the simpler of the two, and is shown in Fig. 1. Generally, over a period of days, the slope and T_0 remain reasonably constant, although near the ground, there is a diurnal swing from night inversion to day ground heating and back. The winter-type profiles are more complicated in that inversions occur up to a far greater altitude, and instead of consisting of just one main segment, they are typically comprised of several segments (Fig. 2).

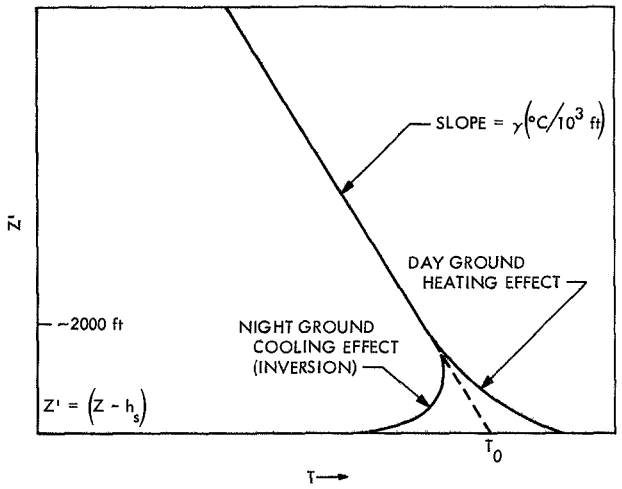


Fig. 1. Summer-type profile

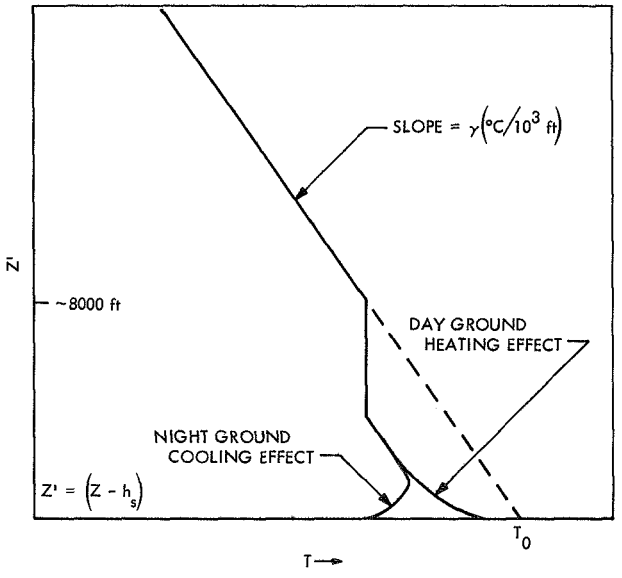


Fig. 2. Winter-type profile

c. Description of relative humidity profiles. After examining a number of relative humidity profiles, one is forced to conclude that there appears to exist no particular functional dependence of relative humidity upon altitude (such as exists with temperature); perhaps the most that can be inferred from the profiles viewed in this report, is that relative humidity seems (but only tenuously!) to remain more constant with altitude during summer-type weather than during winter-type weather (although this would have a good implication, as, in general, the summer wet-term contribution is two to three times the size of the winter contribution). Some examples of relative humidity profiles encountered are shown in Fig. 3. As one might guess, the uncertainties connected with relative humidity profiles produce the largest model errors; however, the use of the assumption:

$$RH(z) = RH_s$$

does not lead to errors as bad as one might at first be inclined to think, because most of the wet-term contribution is concentrated at low altitudes where often the relative humidity has not yet changed as appreciably as it frequently has at the higher altitudes.

d. Meteorological data necessary to implement the proposed model. We have derived the expression for the total (zenith) range refraction as:

$$(\Delta r) = AP_0\left(\frac{R}{g}\right) + \left[\frac{C_1C_2(RH)_s}{\gamma}\right]\left[\frac{B - AC}{\left(1 - \frac{C}{T_0}\right)}^2\right] \times \exp\left(\frac{AT_0 - B}{T_0 - C}\right)$$

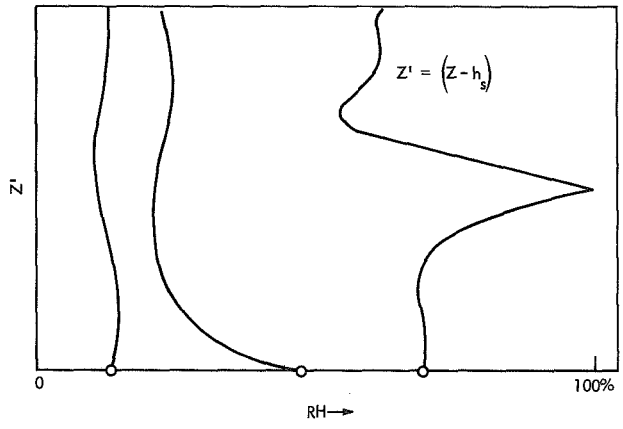


Fig. 3. Relative humidity profiles

To evaluate this expression we require the meteorological inputs:

P_0 = surface pressure

γ = temperature lapse rate

T_0 = extrapolated surface temperature

$(RH)_s$ = surface relative humidity

Basically, the proposed implementation being suggested here is that P_0 and $(RH)_s$ be supplied from frequent (several times daily) samplings at the station, while γ and T_0 be extracted from less frequent (perhaps every several days) atmospheric soundings from any available source in the vicinity (within 100 miles) of the station. When considering this proposal, it should be borne in mind that the dry term, which depends only on surface pressure, represents about 90% of the total tropospheric refraction; thus, even if atmospheric soundings do not prove readily accessible to most stations, γ and T_0 could probably be modeled on a seasonal basis without resulting in too great a loss of accuracy. More specifically then, the data requirements to support the model are defined as follows:

Dry component. Surface pressure measurements at the station would be used to compute the dry-term refraction contribution; due to the slowly varying nature of pressure (barring tornadoes, etc.), measurements every 6 h would be the most frequent that could possibly be needed.

Wet component. Atmospheric soundings would be the best source for the determination of γ and T_0; however, this requirement can be very loosely structured. For instance, the upper troposphere appears to change only slowly, such that soundings might only be required every several days; secondly, since the upper troposphere varies (relatively) slowly horizontally,[2] any source of soundings within perhaps 100 miles should be adequate. For instance, for the Goldstone DSCC, the atmospheric soundings taken at Edwards A.F.B. would seem to amply satisfy this requirement. As already mentioned, if soundings are just not available, T_0 and γ can probably be seasonally modeled without too great a degradation in accuracy. Finally, surface relative humidity measurements at the station would be required. Although this is a rapidly changing parameter, the assumption of constant relative humidity (with altitude) is already quite inaccurate, such that perhaps one "average" surface relative humidity sample per 6-h period would suffice.

*e. **Error analysis.*** When considering the implementation of the scheme proposed here, four major sources of error are in evidence: temperature lapse rate, extrapolated surface temperature, local ground temperature effects (night inversions and day ground-heating), and a non-constant relative humidity. To test the scheme, five days were chosen which represented both seasonal types (two winter-type profiles, two summer-types, and one in-between type). For each day, atmospheric sounding data from Edwards A.F.B. was used to construct two separate sets of temperature and relative humidity profiles, one in the late night and one in the afternoon, so as to maximize distortions due to local ground effects. The days chosen were in December, February, April, August, and September. See Fig. 4 a through j, which will hereafter be referred to as cases 1 through 10. A discussion of the various error magnitudes follows:

Errors in γ and T_0. To approximate the magnitude of the possible errors in slope γ and in T_0, the differences in the graphically computed values of γ and T_0 over the roughly 12-h separation for each day were calculated. These values are given in Table 1.

Table 1. Errors in $\Delta\gamma$ and ΔT_0

Error	Dec.	Feb.	April	Aug.	Sept.
$\Delta\gamma$, °C/10³ ft	0.100	0.025	0.073	0.166	0.017
ΔT_0, °C	1.0	1.0	3.0	2.0	1.0

The resulting average differences are

$$\Delta\gamma \approx 0.076°C/10^3 \text{ ft}$$

$$\Delta T_0 \approx 1.6°C$$

Enlarging these figures somewhat, e.g.,

$$\Delta\gamma \approx 0.1°C/10^3 \text{ ft}$$

$$\Delta T_0 \approx 2.0°C$$

we might consider that these values approximate the uncertainties inherent to the determination of γ and T_0 from atmospheric soundings taken every several days.

[2]Smyth, J. B., *Radio Meteorology at JPL Goldstone Station* (final report), p. 5, July 1965 (JPL internal document).

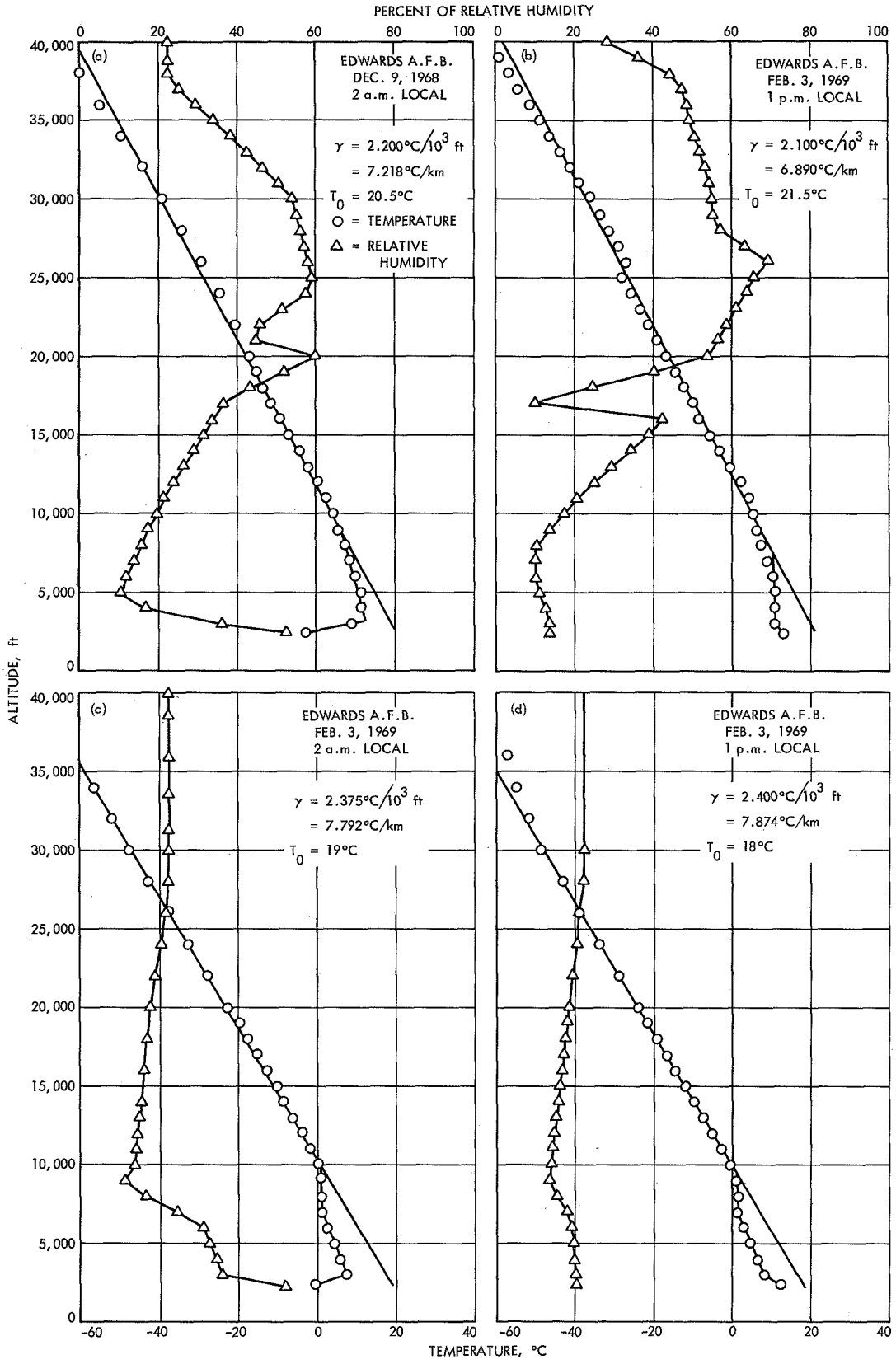


Fig. 4a to d. Temperature and relative humidity versus altitude, cases 1–4

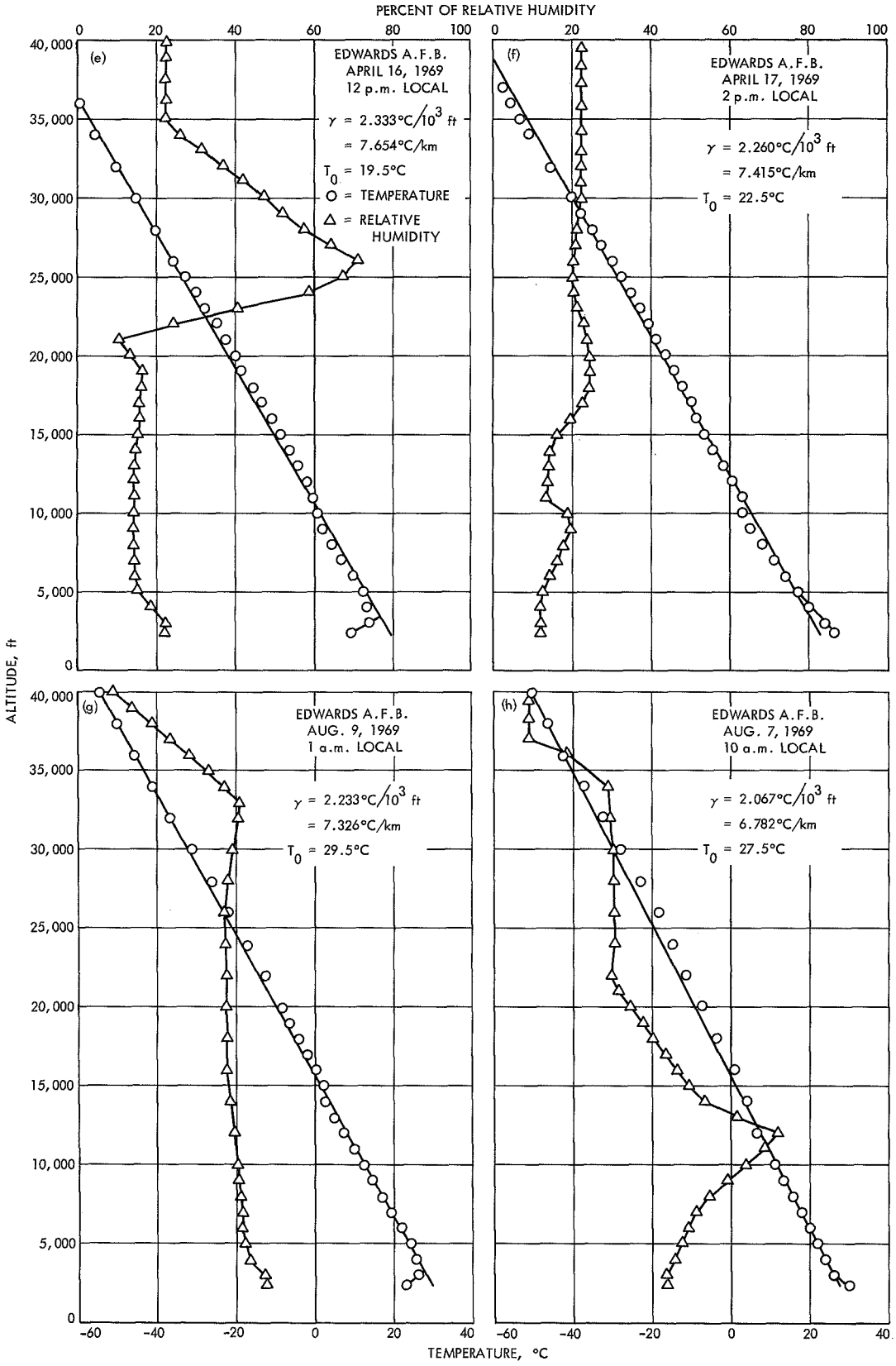


Fig. 4e to h. Temperature and relative humidity versus altitude, cases 5–8

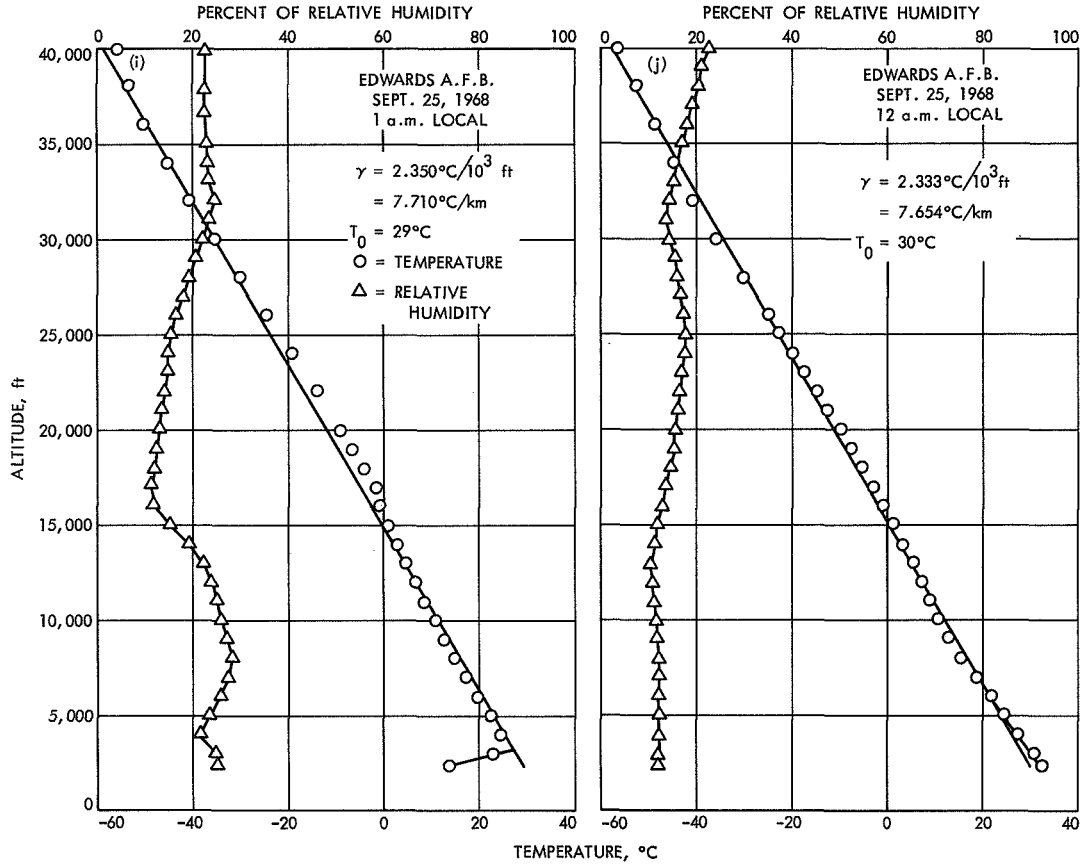


Fig. 4i to j. Temperature and relative humidity versus altitude, cases 9–10

The above values would have produced uncertainties in Δr_w (wet-term refraction, in centimeters) for each of the ten cases as shown in Table 2.

The average uncertainties are

$$\Delta T_0 \approx 1.20 \text{ cm}$$

$$\Delta\gamma \approx 0.48 \text{ cm}$$

Combining these uncertainties in an rms sense, we would have a composite uncertainty due to errors in γ and T_0 of

$$\sum_{\text{rms}}^{\gamma, T_0} \text{errors} \approx 1.30 \text{ cm}$$

Errors due to local ground temperature effects. Using the temperature profiles in cases 1 to 10, the differences between the actual near-ground temperature profile and the extrapolated temperature profile to T_0 were calculated. These differences, in centimeters of wet-term refraction, are given in Table 3. This gives an average error magnitude of

$$|\text{Error}| \text{ ground temperature effects} \approx 0.66 \text{ cm}$$

Errors due to assumption of constant relative humidity. Using the relative humidity profiles in cases 1 to 10, the differences in wet-term refraction due to the differences between the actual relative humidity profile versus an assumed constant relative humidity (equal to the surface relative humidity) were calculated. These differences are given in Table 4. The average error magnitude is

$$|\text{error}| \text{ constant relative humidity assumption} \cong 2.32 \text{ cm}$$

Thus, the assumption of constant relative humidity leads to the largest model errors.

Combined errors of proposed scheme. Considering all the errors to be independent (which to varying degrees

Table 2. Uncertainties in Δr_w

Uncertainty in Δr_w	Case No.									
	1	2	3	4	5	6	7	8	9	10
Due to $\Delta\gamma$	0.66	0.18	0.60	0.23	0.27	0.18	1.05	0.85	0.54	0.27
Due to ΔT_0	1.74	0.48	1.56	0.58	0.68	0.44	2.52	2.04	1.32	0.64

Table 3. Differences in Δr_w due to ground temperature effects

Case No.									
1	2	3	4	5	6	7	8	9	10
1.5 cm	0.7 cm	2.3 cm	0.9 cm	0.4 cm	−0.1 cm	−0.2 cm	0.0 cm	0.3 cm	−0.2 cm

Table 4. Differences in wet-term refraction due to constant RH

Case No.									
1	2	3	4	5	6	7	8	9	10
7.0 cm	−1.6 cm	5.8 cm	0.7 cm	1.2 cm	−1.0 cm	2.7 cm	−2.3 cm	0.8 cm	−0.1 cm

they are not), we would have a total rms error of

$$\sum_{\text{rms}}^{\text{all}} \text{errors} \approx 2.74 \text{ cm}$$

Or simply summing the absolute values of the various errors, we would have

$$\sum^{\text{all}} |\text{errors}| \approx 4.66 \text{ cm}$$

Both figures are considerably less than what is seemingly being obtained in the way of accuracy with the current model.

Comparison with current models. Although the above figures point to the magnitude of the errors which might result if the proposed model was implemented in conjunction with the data sources already described, perhaps the best way to illustrate the increase in accuracy would be to compare the errors which would have been obtained via the new scheme to those which would result from using the current (incorrect) model. First, we will assume an average pressure (at Edwards A.F.B.) of

$$P_0 = 900 \text{ mbar}$$

which will produce a dry-term contribution (to zenith range refraction) of

$$AP_0\left(\frac{R}{g}\right) = 204.8 \text{ cm}$$

Graphically, the (actual) wet-term contributions have been calculated for each of the ten cases, and, when added to the aforementioned dry-term contribution, produce the resultant total refraction shown in Table 5.

The wet-term and total refraction which would have resulted using the proposed scheme (i.e., exhibiting errors due to local ground temperature effects and assumption of constant relative humidity) is shown in Table 6.

As we have seen in *Subsection 2*, the current "exponential decay refractivity" model leads to an expression for total refraction of

$$(\Delta r) = N_s/B$$

Presently, computer programs such as the Orbit Determination Program (ODP), PREDICTS, etc., allow only

Table 5. Total refraction of wet-term and dry-term contributions

Refraction	Case No.									
	1	2	3	4	5	6	7	8	9	10
Wet, cm	4.8	4.6	3.8	2.9	3.7	4.6	18.1	19.3	9.7	5.7
Total, cm	209.6	209.4	208.6	207.7	208.5	209.4	222.9	224.1	214.5	210.5

Table 6. Wet-term and total refraction using proposed scheme

Refraction	Case No.									
	1	2	3	4	5	6	7	8	9	10
Wet, cm	13.3	3.7	11.9	4.5	5.3	3.5	21.0	17.0	10.8	5.4
Total, cm	218.1	208.5	216.7	209.3	210.1	208.3	225.8	221.8	215.6	210.2

Table 7. Surface refractivity and resultant refraction

Refractivity	Case No.									
	1	2	3	4	5	6	7	8	9	10
Actual N_s, cm	272	253	271	259	259	250	293	305	262	252
N_s/B, cm	191.5	178.5	190.9	182.1	182.5	176.3	206.0	214.7	184.4	177.7

Table 8. Comparison of refraction scheme errors

Model	Case No.									
	1	2	3	4	5	6	7	8	9	10
Proposed Scheme	8.5	− 0.9	8.1	1.6	1.6	− 1.1	2.9	− 2.3	1.1	− 0.3
Current and actual N_s	−18.1	−30.9	−17.7	−25.6	−26.0	−33.1	−16.9	− 9.4	−30.1	−32.8
Current and fixed N_s	−31.4	−31.2	−30.4	−29.5	−30.3	−31.2	−44.7	−45.9	36.3	−32.3

fixed values for surface refractivity and inverse scale height. The values typically in use are

$$N_s = 253 \quad \text{(Goldstone DSCC yearly average)}$$

$$B = 0.142$$

leading to a total refraction of

$$N_s/B = 178.2 \text{ cm}$$

However, we can take the current (incorrect) model one step further and calculate the actual surface refractivity and resultant refraction for each of the ten cases (Table 7).

Finally, differencing the current model with constant refractivity, the current model with actual surface refractivity, and the proposed scheme, all with the actual refractivity (graphically/numerically obtained), we would have the results listed in Table 8.

The following average error magnitudes would result:

Proposed scheme	= 2.8 cm
Current model with N_s actual	= 24.1 cm
Current model with N_s fixed	= 34.3 cm

If the theoretical basis for the new formulation as derived in *Subsections 2* and *3* is correct, then use of the here-proposed scheme would seem to offer the following advantages:

(1) Decrease in total (zenith) range refraction error of up to an order of magnitude.

(2) Less need for station meteorological data samplings than is presently envisioned in Tracking System Analytical Calibration (TSAC) documentation.

References

1. Rowlandson, L. G., *Simple Analytical Functions Which Provide Magnitude of Range and Angle Errors for Propagation in an Exponential Atmosphere*, U.S. Air Force, Hanscom Field, Bedford, Mass., Jan. 19, 1968.

2. Hess, S. L., *Introduction to Theoretical Meteorology*, Holt, Rinehart & Winston, Inc., New York, 1959.

XIV. Facility Engineering

OPERATIONS AND FACILITIES

A. Overseas 210-ft-diam Antenna Project,

R. D. Casperson and W. W. Lord

1. Introduction

Two existing 85-ft-diam antenna stations at overseas locations are being upgraded to 210-ft-diam antenna stations. The project includes the construction of two 210-ft antenna instruments, support facilities, and provisions for sufficient electronic equipment to meet the minimum performance requirements of spacecraft in the 1973 era. In addition to the antenna instrument, modifications must be made to the existing station facilities in order to adapt them for simultaneous use of the existing 85-ft antennas and the new 210-ft antennas. Figures 1 and 2 show the locations of the projects.

A fixed-price contract for the antenna instrument was executed with the Collins Radio Company on August 14, 1969, to:

(1) Review the detail engineering specifications, drawings, procedures, etc.

(2) Fabricate, transport, erect, and test two 210-ft antennas, one to be erected at Ballima (DSS 43) near Canberra, Australia, and the other at Robledo (DSS 63) near Madrid, Spain.

The facilities adaptation in Australia is being conducted jointly by the Australian Department of Supply, Network Support Facility, and the Australian Department of Works. In Spain, the facilities adaptation is under the supervision of the U.S. Navy, Naval Facilities Command, with an independent firm doing the actual engineering.

The tasks to be accomplished by the overseas agencies are as follows:

(1) Modifications of the operations and engineering building.

(2) Modifications and addition to the utilities support–power plant building.

(3) Modifications to and extension of site utilities.

(4) Modifications and extension of power generation and distribution systems.

(5) Modifications to site grading and paving.

(6) Site modifications and preparation for installation of high-power transmitter facilities.

(7) New RF protection tunnel and covered access walkway from the 210-ft antenna pedestal to the existing operations and engineering building.

(8) Fire protection.

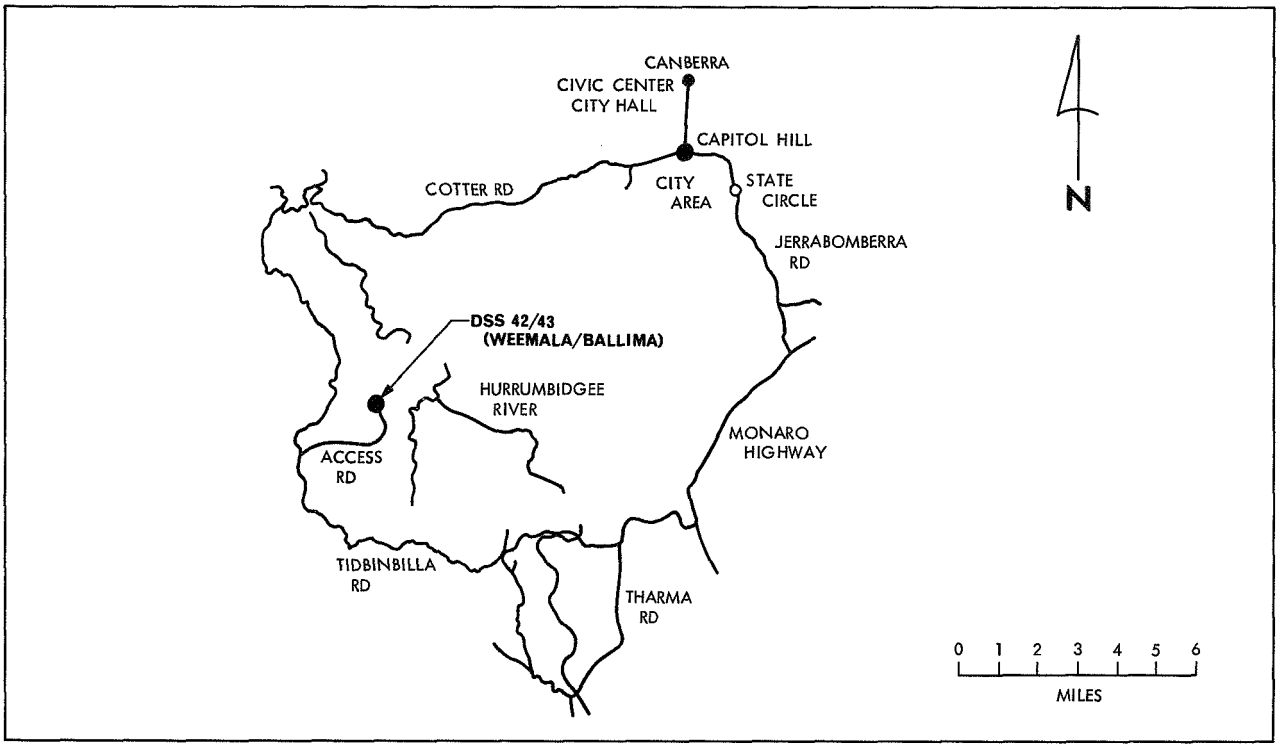


Fig. 1. Site location in Australia

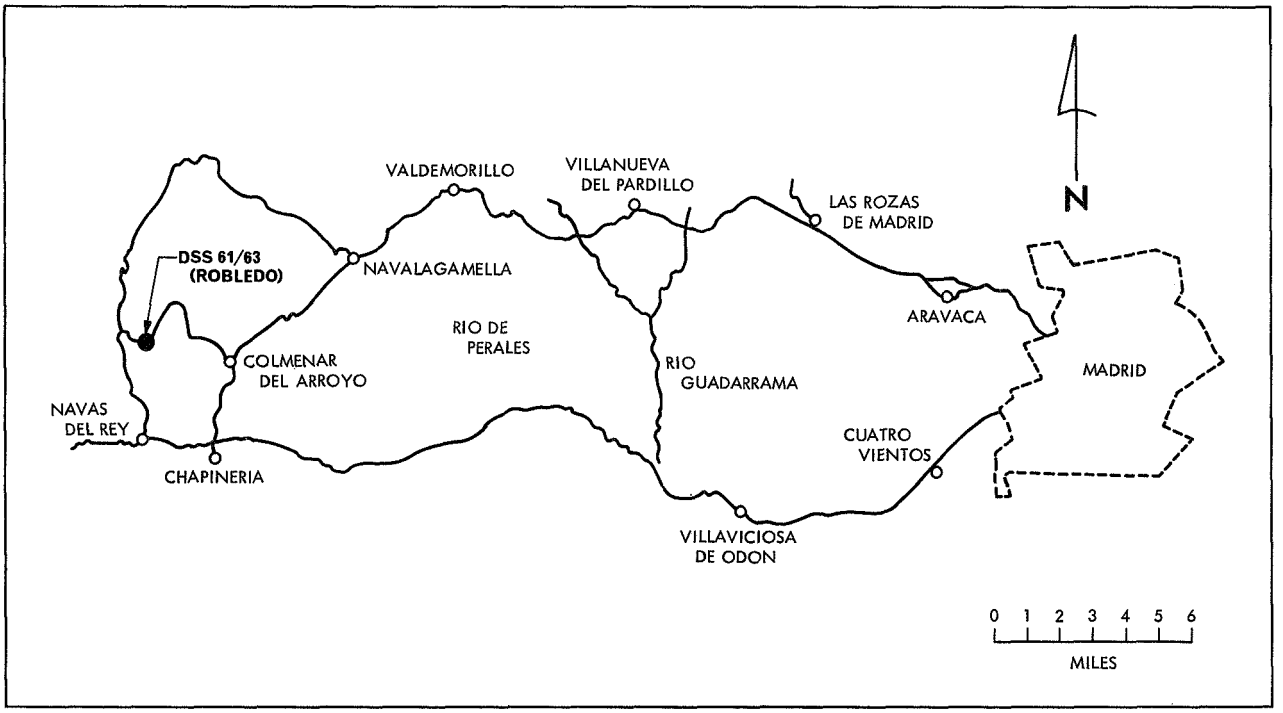


Fig. 2. Site location in Spain

In both Australia and Spain, the cognizant overseas agencies are responsible for the construction, by subcontract, of all work under their supervision; however, overall project responsibility remains with JPL. The planned configurations for the Canberra and Madrid sites are shown in Figs. 3 and 4. All work outside the dashed line is to be done by the overseas agencies. The high-power transmitter facilities are also their responsibility.

All antenna subsystems, other than the antenna mechanical subsystem and the facilities subsystem, will be furnished and installed by JPL.

2. Progress

a. Engineering review. Review of engineering specifications, drawings, procedures, etc., has been completed and accepted as correct by the contractor (Collins Radio Co.).

b. Fabrication. Collins Radio Company has subcontracted the following antenna components:

Component	Subcontractor
Bull gears and drive assemblies	Philadelphia Gear Corp.
Elevation bearing assembly	National Steel and Ship Building Co.
Azimuth radial runner	Westinghouse Electric Corp.
Reflector tipping assembly	Coeur d'Alenes Steel Co.
Surface panels	Radiation Systems, Inc.
Alidade, quadripod, apex, instrument tower, and alidade building	Precision Fabrication, Inc.
Hydrostatic bearing and azimuth bearing trucks	Western Gear Corp.

Fabrication to date is progressing in a satisfactory manner and on schedule. To insure against improper fabrication and costly construction delays, extensive trial erections of major components are being conducted in the fabricators' facilities prior to shipment. Figures 5, 6, and 7 show examples of these trial erections.

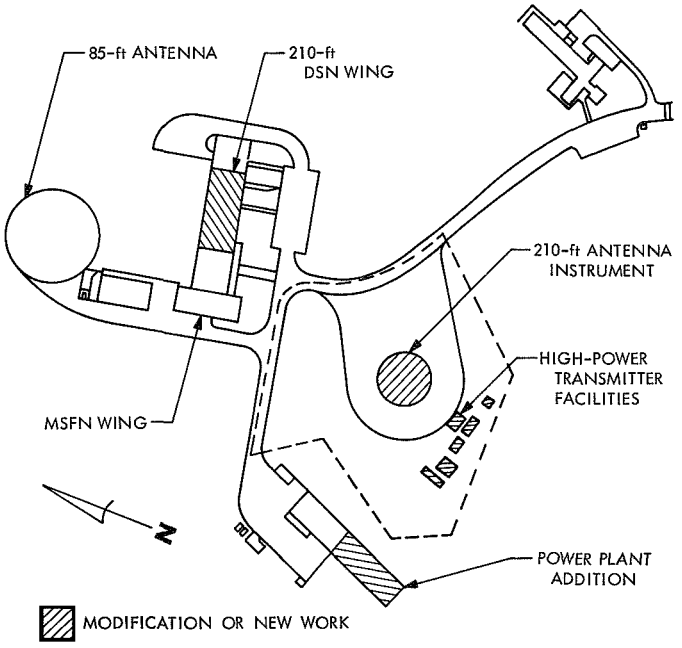


Fig. 3. Plot plan of Canberra complex

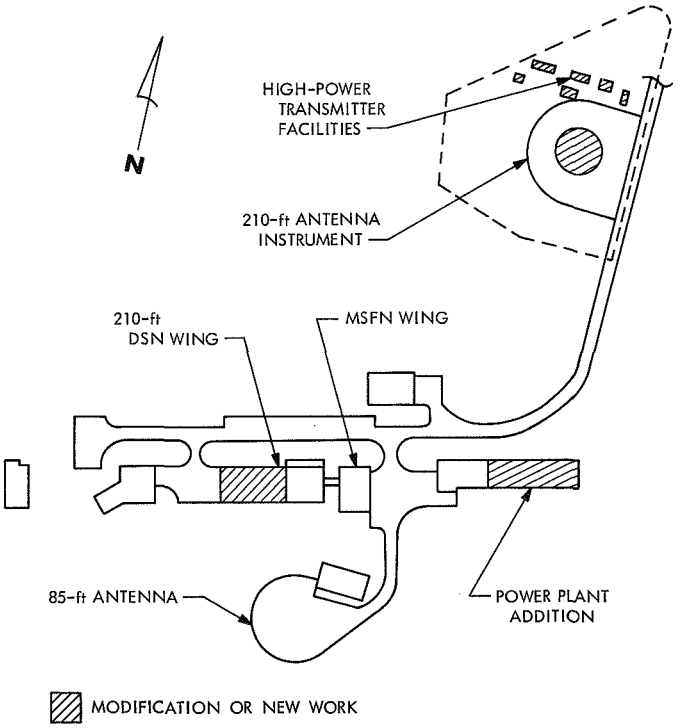


Fig. 4. Plot plan of Madrid complex

The tricone and rotary hyperbola were first erected and tested at DSS 14, Goldstone, California (SPS 37-63, Vol. II, pp. 56–60). After successful operation at Goldstone, these subsystems were accepted as standard for

Fig. 5. Alidade

Fig. 6. Elevation wheel

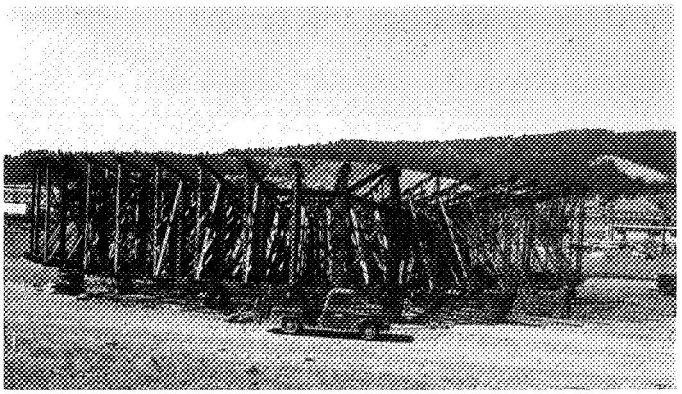


Fig. 7. Reflector backup structure

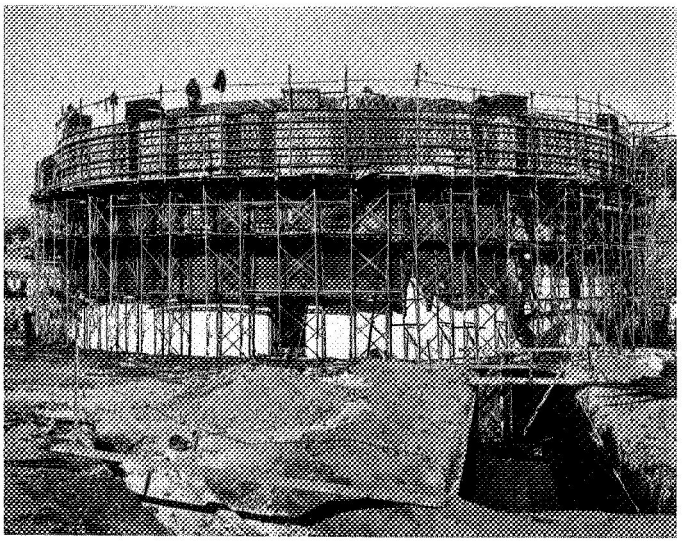


Fig. 8. Pedestal construction at DSS 43

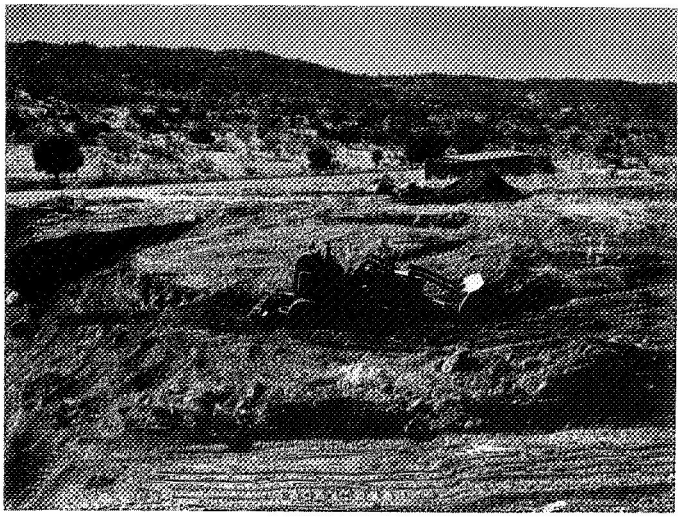


Fig. 9. Pedestal excavation at DSS 63

the 210-ft-antenna subnet. After installation on the overseas antennas, all of the DSIF 210-ft antennas will be of the same structural–mechanical configuration.

A contract has been placed with Philco-Ford for the fabrication, installation, and test of tricone systems for each of the overseas 210-ft antenna projects. This work is a part of the project and will be completed after the work of the antenna prime contractor.

c. Construction. Construction began at DSS 43 on November 20, 1969. The instrument tower and pedestal concrete are virtually complete and progressing on schedule. In Australia, Collins is acting as the construction contractor.

At DSS 61A, construction began on June 22, 1970. Excavation, site preparation, and utility work are progressing on schedule. The pedestal subcontractor in Spain is Laing-Ibecia, an English–Spanish company. The steel erection subcontractor has not yet been selected by Collins Radio Company. Figures 8 and 9 show the construction sites and the progress that has been made.

d. Facility design. Design of facility adaptations by the overseas agencies is progressing satisfactorily. In both Australia and Spain, the 30% design review is complete. At this time, effort is being made to coordinate modifications of the control rooms to permit support of presently scheduled spacecraft projects.

Subject Index

Subject Index (contd)

NASA — JPL — Coml., L.A., Calif.

N71-14951

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No. 37-65, Vol. II	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle SPACE PROGRAMS SUMMARY, VOL. II THE DEEP SPACE NETWORK		5. Report Date September 30, 1970 6. Performing Organization Code
7. Author(s) JPL Staff		8. Performing Organization Report No.
9. Performing Organization Name and Address JET PROPULSION LABORATORY California Institute of Technology 4800 Oak Grove Drive Pasadena, California 91103		10. Work Unit No. 11. Contract or Grant No. NAS 7-100 13. Type of Report and Period Covered Space Programs Summary July 1 to August 31, 1970
12. Sponsoring Agency Name and Address NATIONAL AERONAUTICS AND SPACE ADMINISTRATION Washington, D.C. 20546		14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The Space Programs Summary is a multivolume, bimonthly publication that presents a review of technical information resulting from current engineering and scientific work performed, or managed, by the Jet Propulsion Laboratory for the National Aeronautics and Space Administration. The Space Programs Summary is currently composed of four volumes:

Vol. I.	Flight Projects (Unclassified)
Vol. II.	The Deep Space Network (Unclassified)
Vol. III.	Supporting Research and Advanced Development (Unclassified)
Vol. IV.	Flight Projects and Supporting Research and Advanced Development (Contents Confidential)

17. Key Words (Selected by Author(s))	18. Distribution Statement
Not applicable for this type of report	Unclassified -- Unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	160	