1431

(NASA-CR-164939) THE TELECOMMUNICATIONS AND DATA ACQUISITION REPORT Progress Report, Jul. - Aug. 1981 (Jet Propulsion Lab.) 244 p HC A11/MF A01 CSCL 17B

N82-11279 THRU N82-11310 Unclas G3/31 08289

# The Telecommunications and Data Acquisition Progress Report 42-65

## July and August 1981

N. A. Renzetti
Editor

October 15, 1981

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

# The Telecommunications and Data Acquisition Progress Report 42-65

## July and August 1981

N. A. Renzetti
Editor

October 15, 1981

National Aeronautics and
Space Administration

**Jet Propulsion Laboratory**
California Institute of Technology
Pasadena, California

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

# Preface

This publication was formerly entitled *The Deep Space Network Progress Report*. Although the practice of reporting progress in the development and operations of the Deep Space Network continues, the report has been expanded to include developments in Earth-based radio technology as applied to other research programs. These programs are:

(1) Geodynamics: For several years, the laboratory has been developing radio interferometry at microwave frequencies for application to geodetic measurements. This branch of telecommunications technology is now being applied to the study of geodynamics.

(2) Astrophysics: The deep space stations, individually and in pairs as an interferometer, have been used by radio astronomers for astrophysics research by direct observations of radio sources.

(3) An activity closely related to radio astronomy's use of the deep space stations is NASA's continuing program of radio search for extraterrestrial intelligence in the microwave region of the electromagnetic spectrum.

Each succeeding issue of this report will present material in some, but not all, of the following categories:

Radio Astronomy
Search for Extraterrestrial Intelligence
Radio Interferometry at Microwave Frequencies

Geodetic Techniques Development
Spacecraft Navigation
Orbiting Very Long Baseline Interferometry

Deep Space Network

Description
Program Planning
Planetary and Interplanetary Mission Support
Advanced Systems
Network and Facility Engineering and Implementation
Operations
Spacecraft Radio Science
Planetary Radar
Energy

In each issue, there will be a report on the current configuration of one of the seven DSN systems (Tracking, Telemetry, Command, Monitor and Control, Test Support, Radio Science, and Very Long Baseline Interferometry).

The work described in this report series is either performed or managed by the Telecommunications and Data Acquisition organization of JPL.

# Contents





PRECEDING PAGE BLANK NOT FILMED

Station Control and System Technology



Network Data Processing and Productivity



## NETWORK AND FACILITY ENGINEERING AND IMPLEMENTATION

Network



Deep Space Stations

N82-11280



# Radio Astronomy

R. D. Shaffer and J. T. McCluskey
Control Center Operations Section

S. Gulkis, M. Klein, and T. Kuiper
Planetary Atmospheres Section

*This article reports on the activities of the Deep Space Network in support of Radio Astronomy operations during the third quarter of 1981. Some results of the use of a low-noise maser on loan from NRAO are presented.*

## I. Introduction

Deep Space Network (DSN) 26-, 34-, and 64-meter-antenna stations are utilized in support of these categories: NASA Office of Space Science (OSS) and Radio Astronomy Experiment Selection (RAES).

## II. Radio Astronomy Operations

### A. NASA OSS Category

As was reported in TDA Progress Report 42-63 (March–April 1981) a K-band (18- to 25-GHz) reflected-wave ruby maser (Ref. 1) has been borrowed from the National Radio Astronomy Observatory for use on the 64-meter (DSS-43) antenna at the Tidbinbilla Tracking Station, near Canberra, Australia. The successful operation of this system is the result of an international collaboration involving scientists and engineers from JPL, National Radio Astronomy Observatory (NRAO), and Commonwealth Scientific and Industrial Research Organization (CSIRO; e.g., see Ref. 2). The first observations with this new system were made by the Planetary Radio Astronomy (OSS 196-41-73) and Interstellar Microwave Spectroscopy (OSS 188-41-55-12-55) Program. Preliminary results, which have been presented to the Astronomical Society of Australia, concentrate on water vapor and ammonia sources (Ref. 2).

During February and March 1981, spectral line observations were carried out near 22 GHz for water vapor sources and near 24 GHz for ammonia sources. For the water vapor observations, beam switching between the main beam and the reference beam at a 4-Hz rate was used. The beam-switching technique was not used for ammonia because of its wide angular distribution throughout the Galaxy; instead, each 4-min on-source observation was followed by a 4-min reference observation at a position several minutes of arc away and clear of any obvious continuum sources.

**1. $H_2O$ Sources.** The water vapor observations were made in the direction of known southern OH and $H_2O$ maser sources. All of the previously detected water line sources examined were detected. In addition, two new water vapor maser sources were discovered, G301.1+1.1 and G308.9+0.1.

The spectrum of G301.0+1.1 is presented in Fig. 1. Observations of the Parkes source G305.4+0.2 presented in Fig. 1 show it to be spatially resolved. The high-velocity feature at -90 km $s^{-1}$ that was found by Batchelor, et al. (Ref. 3) appears in the Tidbinbilla reference beam (and hence appears as a negative feature), accompanied by a component at -40 km $s^{-1}$. These spectra are reproductions of the data produced at the telescope and have not had baselines removed.

2. **Ammonia Sources.** Six ammonia sources were found: G291.3-0.7, G305.4+0.2, G322.2+0.6, G327.3-0.5, G333.6-0.2, and G268.4-0.8. Spectra of two of these sources, G291.3-0.7 (RCW 57) and G305.4+0.2, are presented in Fig. 2. Both show clearly the presence of the quadrupole splitting satellite lines that will allow the determination of $NH_3$ optical depths in these clouds. An analysis of the conditions in the ammonia clouds is presently under way.

## References


1. Moore, C. R., and Clauss, R. C., *IEEE Trans. Microwave Theory Tech., MTT-27*, 249, 1979.
2. Jauncey, D. L., Batty, M. J., Gay, G. J., Moore, C. R., Batelaan, P., Clauss, R., Dickinson, D., Gaulkis, S., Klein, M., Kuiper, T., Morris, G., Neff, D., Ricketts, W. B., and Swanson, P., "18-25 GHz Low-Noise Line Receiver for the Tidbinbilla 64-m Antenna." Prepared for submission to *Proceedings of the Astronomical Society of Australia*, June 1981.
3. Batchelor, R. A., et al., *Lett. Astron. J.*, USSR Acad. Sci., *2*, 467 (1976).

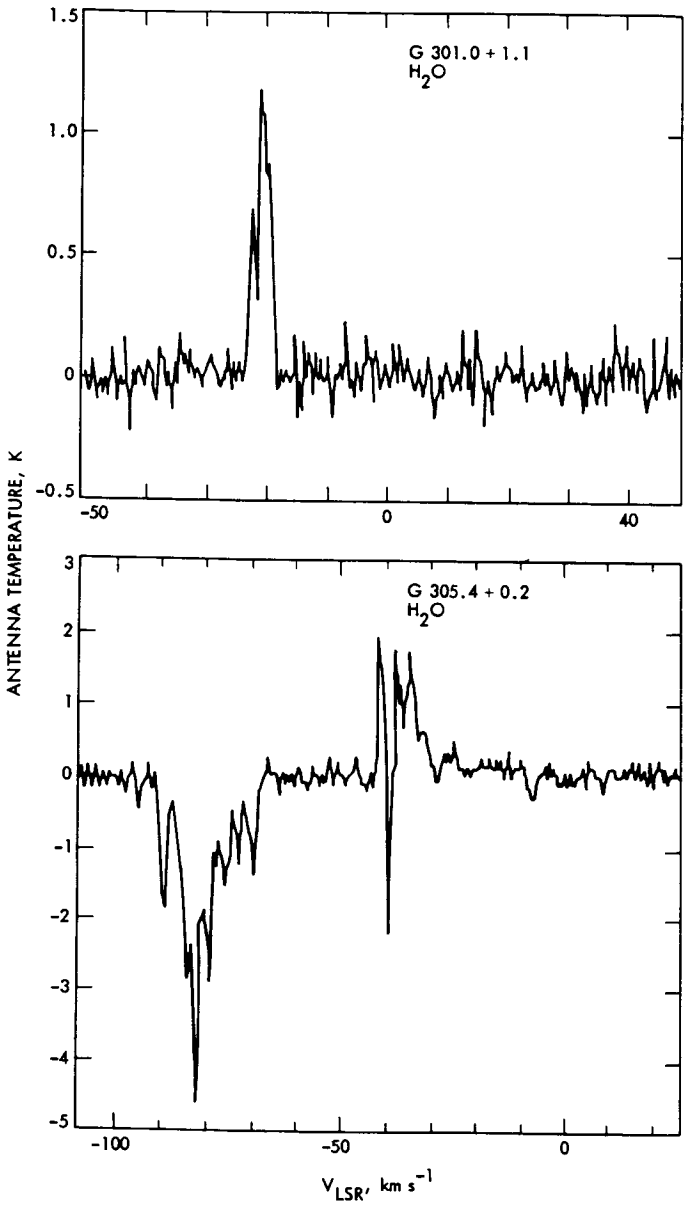



**Fig. 1. Spectra of the ($6_{16}$ - $5_{23}$) transition of $H_2O$($\nu_0$ = 22.235 GHz) taken with the 64-m antenna in March 1981 in the direction of two galactic sources. Galactic coordinates of sources are indicated on the figure.**

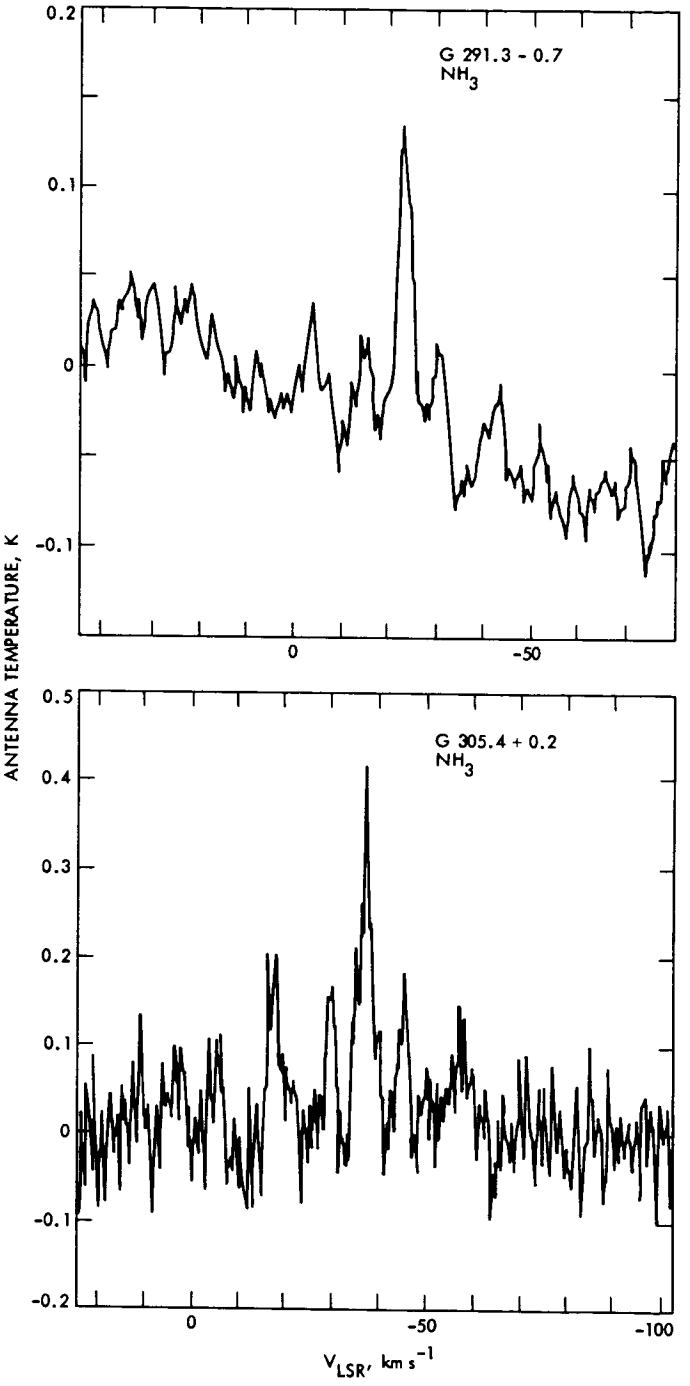



**Fig. 2. Spectra of the (1, 1) transition of $NH_3$($\nu_0$ = 23.694 GHz) taken with the 64-m antenna in March 1981 in the direction of two galactic sources. Galactic coordinates of sources are indicated on the figure.**

N82-11281



# A Dedicated 26-m SETI Sky Survey Instrument Facility: A Feasibility Study

A. L. Berman
TDA Mission Support Office

R. D. Shaffer
Control Center Operations Section

*During February 1985, the Ground Spaceflight Tracking and Data Network (GSTDN) 26-m radio telescope facilities at Goldstone, California, and Canberra, Australia, are planned for deactivation. It is here proposed that these existing facilities, including large diameter (26-m) radio telescopes, control rooms, office buildings, and technical support facilities would provide an economical framework for a standalone SETI Instrument Facility. Additional benefits that would accrue from the proposed utilization of these facilities are the capability for remote monitor and control from the Deep Space Communications Complex (DSCC) Signal Processing Center, combined with a reduction in ground-induced radio frequency interference due to the physical distance to the other DSCC front end areas and the Signal Processing Center.*

*This article examines the feasibility of converting a deactivated GSTDN 26-m facility into a dedicated SETI Instrument Facility.*

## I. Introduction

The Search for Extraterrestrial Intelligence (SETI) Program is a NASA Supporting Research and Technology Program with primary intent to search the microwave region of the spectrum for signals of extraterrestrial intelligent origins. During the next several years, the foremost goal of the SETI Program is the development of a SETI breadboard instrument (Refs. 1 and 2). Ultimately, the goal of the SETI Program will be the exploration of a well-defined volume of multidimensional microwave search space using large existing radio telescopes and a new technologically sophisticated data acquisition and analysis system (Refs. 3 and 4). A major component of this search is the SETI "Sky Survey," which proposes to observe the entire celestial sphere over a wide range of frequencies. Primary requirements to conduct such a search at high sensitivity ($<10^{-22}$ W/m$^2$) are a low noise, large diameter radio telescope, available to SETI for long periods of time.

In 1985, the current Ground Spaceflight Tracking and Data Network (GSTDN) 26-m facilities are expected to be deactivated as part of the Networks Consolidation Program. This

article explores the feasibility of converting one of these to-be-deactivated 26-m GSTDN facilities into a dedicated, SETI Sky Survey Instrument Facility.

## II. SETI Sky Survey Requirements and Modes of Operation

The proposed SETI Sky Survey is intended to be a high resolution, high sensitivity search of the entire sky between the frequencies 1.0 and 10.0 GHz, with spot coverage to 25 GHz. The survey will be conducted by sweeping out swatches of the sky that are a half-power beamwidth wide, at a constant telescope drive rate. The minimum detectable flux for a constant telescope drive rate is:

$$\phi = (4\alpha kT_s/\pi\epsilon)\sqrt{\omega b\nu/70\,cD^3}$$

where

$\phi$ = flux, W/m$^2$

$\alpha$ = signal-to-noise ratio

$k$ = Boltzman's constant

$T_s$ = system noise temperature, K

$\epsilon$ = aperture efficiency

$\omega$ = angular tracking rate, deg/s

$b$ = binwidth, Hz

$\nu$ = frequency, Hz

$c$ = speed of light, m/s

$D$ = diameter, m

Hence, for a given total survey time $t$ (where $t \propto D\,\omega^{-1}$), sensitivity scales directly with system noise temperature and inversely with diameter:

$$\phi \propto T_s D^{-1}$$

Therefore, the fundamental requirements to perform the sky survey can be stated as:

(1) High radio telescope availability.

(2) Large radio telescope diameter.

(3) Low noise radio telescope front end.

The telescope drive rate will be set to achieve sensitivites in the range of $10^{-22}$ to $10^{-23}$ W/m$^2$ (a "range" of sensitivity results with a constant telescope drive rate because the sensitivity decreases with the square root of the frequency). This sensitivity range, coupled with the proposed (Ref. 2) 256-MHz bandwidth, 32-Hz binwidth Multichannel Spectrum Analyzer (MCSA), then determines the total amount of time necessary to complete the SETI Sky Survey.

The SETI Sky Survey System is intended to be automated to the fullest extent possible, with the largest deviation from the desire for automation resulting from the necessity to frequently "look-back," or check areas that appear promising. It is estimated that the look-back time will consume 25 to 50% of the total time of the survey. The time required to perform the SETI Sky Survey with look-back time, on a low noise front end, 26-m diameter telescope with MCSA is five years, assuming a dedicated telescope with an effective duty cycle of 16 hours per day.

## III. The Standalone 26-m SETI Instrument Facility

The current NCP plan for combining the DSN and the GSTDN indicates that the still existing GSTDN 26-m facilities will be deactivated in approximately February 1985. It is here suggested that these 26-m GSTDN facilities at Goldstone, California, or Canberra, Australia, would provide an excellent and economical basis for a standalone SETI Instrument Facility. The economic utility of this concept stems directly from the existence of a complete, in-place radio telescope facility, including telescope structure, control room, office space, and technical facilities, such as power, water, air conditioning, and lighting. Additional benefits accrue in that the facilities are reasonably isolated from other radio telescope activities so as to lessen ground-induced radio frequency interference (RFI) problems, but are close to a Deep Space Communications Complex (DSCC) remote monitor and control facility (the DSCC Signal Processing Center). It is here considered that, following deactivation, the former GSTDN site would be stripped of all equipment and capabilities not directly related to or required by a SETI Instrument, and only those portions of the existing buildings required by SETI would be maintained.

Figure 1 is a functional block diagram that illustrates how the existing 26-m GSTDN facility would be reconfigured as a SETI Instrument. The existing system design of the SETI Sky Survey Instrument includes the following general subsystems and assemblies:

(1) Antenna Mechanical Subsystem: Antenna Control Assembly.

(2) Antenna Microwave Subsystem.

(a) Orthomode Feed Assembly.

(b) Parametric Upconverter Assembly.

(c) Maser Assembly.

(3) Receiver Subsystem.

(a) Noise Adding Radiometer Assembly.

(b) Receiver Assembly.

(c) A/D Converter Assembly.

(4) Spectrum Analyzer Subsystem

(a) Multichannel Spectrum Analyzer Assembly.

(b) Signal Detector Assembly.

(5) Signal Processor Subsystem

(a) Signal Processor Computer Assembly.

(b) Signal Processor Software Assembly.

(6) Monitor and Control Subsystem.

(a) Monitor Assembly.

(b) Controller Assembly.

(7) Frequency and Timing Subsystem.

(8) Communications Monitor and Formatter Subsystem.

(9) Technical Facilities Subsystem.

(a) Air Conditioning Group.

(b) Power Generation Group.

(c) Power Distribution Group.

(d) Site Protection Group.

(e) Facility Lighting Group.

(f) Civil Structures Group.

It is assumed that the Technical Facilities Subsystem will be transferred from the existing facilities essentially intact and will require little modification to meet SETI requirements. In the Antenna Mechanical Subsystem, the telescope surface may require modification to allow frequency coverage to 10 (or 25) GHz. The SETI Instrument equipment, from the Microwave Subsystem to the Signal Processor Subsystem, will be all new equipment – designed, fabricated, and implemented by the SETI Program.

## IV. Operations Concept

The Operations of the SETI Instrument will be automated to the fullest degree possible. However, there will still be operator action required in terms of routine operations, such as changing of magnetic tapes, reconfigurations, and computer reinitializations. The concept proposed here is that routine monitor and operation of the SETI Instrument will be performed remotely from and by the personnel of the DSCC Signal Processing Center (SPC). In fact, this concept fits nicely into one of the roles of the Signal Processing Center, which is to monitor and operate the various DSCC front end areas (FEAs). In this regard, the SETI Instrument will appear just as the other remote, automated FEA facilities; the operators will monitor the SETI Instrument operation, and will perform some actions directly from the SPC (SETI monitor and control inputs), or will travel to the SETI Instrument Facility to perform other actions (such as mounting a magnetic tape).

## V. Maintenance and Repairs

Maintenance and repairs of the SETI Instrument will be handled by the same DSCC personnel and in the same manner as would be the various DSCC FEAs. Routine maintenance will be on a scheduled basis. Equipment failures will be monitored at the SPC, with repair personnel dispatched from the SPC in response to indicated equipment malfunctions.

## VI. A SETI Instrument Facility for Off-Line Operations

During the life of the SETI Sky Survey, it is assumed that one member of the SETI Science Team will be engaged on a full-time basis in both maintenance of the SETI Sky Survey configuration and in off-line data processing and analysis activities, such as:

(1) Analysis of false alarms.

(2) Adjustment of detection algorithm parameters concerned with threshold, baseline, system gain, and RFI recognition.

(3) Reconfiguration to look back at interesting or possible signal areas of the sky.

(4) Preliminary processing of magnetic tape recorded data prior to shipment to other Science Team investigators.

The availability of existing office space adjacent to the SETI Instrument itself combines to make an attractive concept as an all-inclusive SETI Instrument Facility, including both the Instrument and the off-line data processing analysis area, manned full time (i.e., 40 hours/week) by a member of the Science Team. Additional existing areas would be used for SETI data storage, working space for other SETI Science Team members who might choose to work directly at the SETI Instrument Facility at various times, and SETI conferences.

## VII. Radio Frequency Interference (RFI) Considerations

The advantages of a deactivated 26-m GSTDN as a stand-alone SETI Instrument Facility are further enhanced by the site's (relatively) remote location from other transmitters and electronics within the DSCC. The site is (in the case of Goldstone) located approximately 12 kilometers from the other FEAs, with several ridges or hills interspersed between the two areas. The current GSTDN facility thus accrues the advantages of being sufficiently remote from the rest of the DSCC so as to reduce ground-induced RFI, but sufficiently close to the SPC to be monitored and operated remotely and conveniently.

## Acknowledgment

The authors thank Dr. M. J. Klein and Dr. E. T. Olsen for many excellent suggestions and comments.

## References

1. Berman, A. L., and Crow, R. B., "The SETI Program Plan and Instrument Development Status," in *The TDA Progress Report 42-64*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 3-9, August 15, 1981.

2. Crow, B., "The SETI Instrument Development Plan," in *The TDA Progress Report 42-59*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 15-22, October 15, 1980.

3. Berman, A. L., "The SETI Observational Plan," in *The TDA Progress Report 42-57*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 9-15, June 15, 1980.

4. Gulkis, S., Olsen, E. T., and Tarter, J., "A Bimodal Search Strategy for SETI," in *The TDA Progress Report 42-59*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 23-32, October 15, 1980.

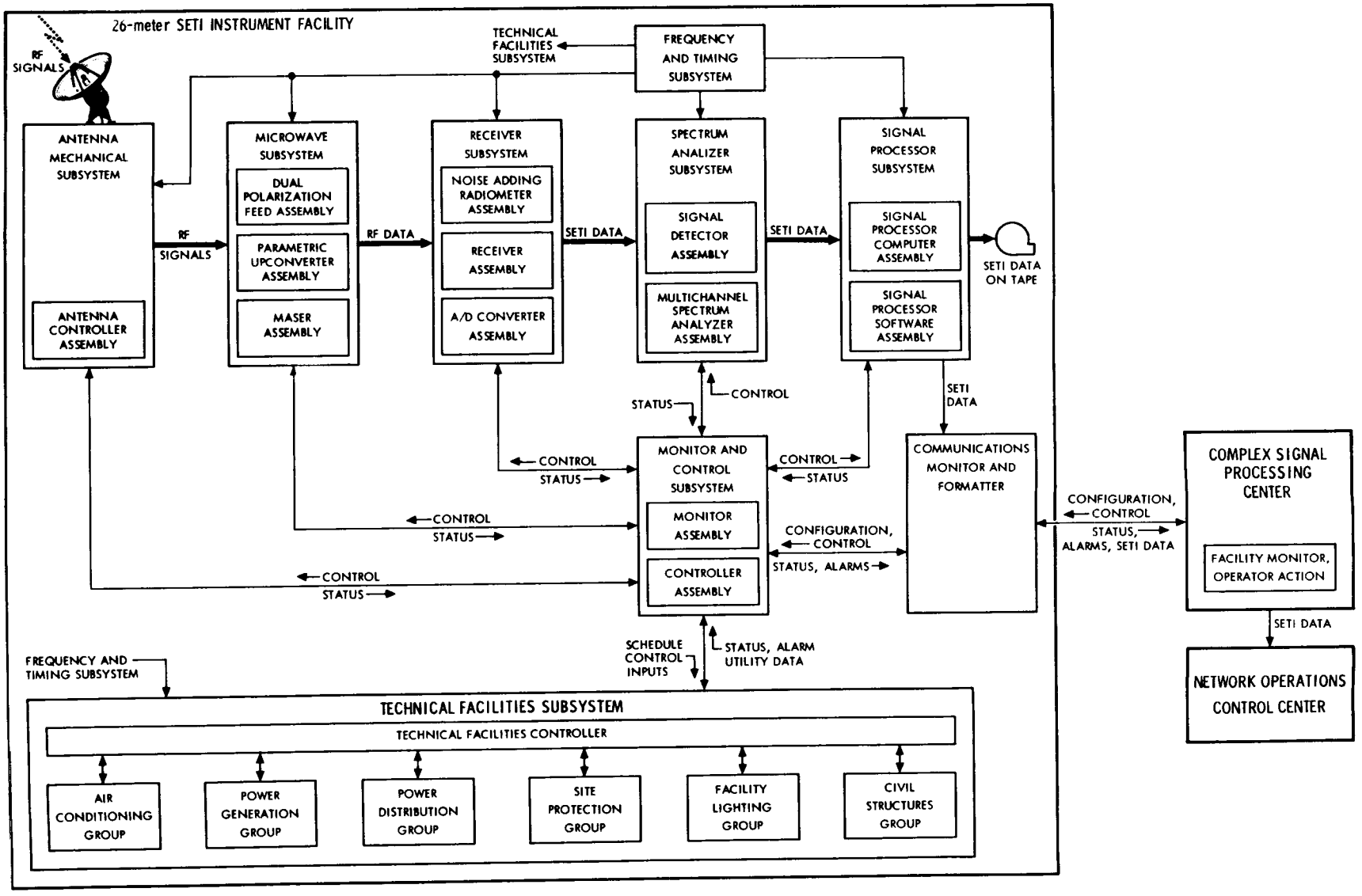



**Fig. 1. 26-meter SETI Instrument Facility functional block diagram**

N82-11282 D3



# DSN Ground Communications Facility

R. H. Evans
TDA Engineering Section

*A functional description of the GCF and its relationships with other elements of the DSN and NASCOM is presented together with development objectives and goals and comments on implementation activities in support of flight projects.*

## I. Introduction

The Ground Communications Facility (GCF) is one of the three elements of the DSN. The GCF provides for transmission, reception, and monitoring of Earth-based point-to-point communications between the Deep Space Stations(DSSs, one of the DSN elements), the Network Operations Control Center (NOCC, the other element) located at JPL, Pasadena, and the Mission Operations Control Center (MOC) at JPL. Voice, teletype, high-speed data, and wideband data circuits of the worldwide NASA Communications Network (NASCOM) are utilized for all long-distance circuits, except those between JPL and the Goldstone Deep Space Communications Complex (GDSCC). Goddard Space Flight Center (GSFC) NASCOM Engineering has delegated the responsibilities for planning, budgeting, design, implementation, operation, and maintenance of the communications requirements between Goldstone and JPL to the DSN GCF. Additionally, the GCF provides communications services between the DSSs at each geographic communications complex (Madrid, Australia and Goldstone, Calif.) via intersite microwave systems and between separated areas of the NOCC at JPL via 230 kbit/s wideband data channels. Also, voice communications are provided within the stations, between the stations, within the complexes, and within the NOCC. The GCF is comprised of seven subsystems; Voice, Teletype, High-Speed Data, Wideband Data, GCF Monitor and Control, Data Records, and Network Communications Equipment. The DSN Telecommunications and Data Acquisition Engineering Office of JPL provides the technical direction and systems management of the GCF and acts as the representative of NASCOM for switching and interconnect functions on the west coast.

## II. GCF-NASCOM Interrelationships

The interrelationships at the programmatic level between JPL's DSN GCF and the NASCOM network, which is managed, engineered, and controlled at GSFC, are characterized as follows:

NASCOM

(1) Provides long-haul operational ground communications in support of all NASA projects and mission activities including those supported by the DSN.

(2) Accepts and supports communications requirements established by the DSN and validated through continuing consultation and review.

(3) Establishes in consultation with the users the basic characteristics of the NASCOM systems, such as tele-

type line rate and block header formats for switching, and the user electrical interfaces.

GCF

(1) Provides ground communications for all DSN missions and uses the services of NASCOM.

(2) Establishes additional characteristics of all GCF subsystems on an end-to-end basis such as block multiplexing, error correction, GCF monitoring and control, and data records capabilities.

## III. Objectives and Goals

The primary objectives of the GCF are to provide highest quality point-to-point transfer of operational data within the DSN and provide simple user and NASCOM electrical and operational interfaces. These objectives are being met by:

(1) Providing automatic message switching and routing.

(2) Providing data transmission subsystems that are as transparent to the user as possible.

(3) Minimizing project-dependent equipment within the GCF.

(4) Providing a centralized common user data records capability.

The goals of the GCF are to provide highly reliable and cost-effective data transmission while continuing an adequate capability balance for multiple mission users. These goals include the following:

(1) Equipment and routing redundancy to minimize single-point-of-failure impact.

(2) End-to-end performance which provides essentially block-error-free throughput.

(3) Design coordinated and consistent with the NASCOM Development Program.

## IV. Configuration and Functional Subsystem

The current GCF configuration, including the related NASCOM interfaces and functions, is illustrated in Fig. 1. This configuration illustrates the long-haul communication circuit services external to JPL and Deep Space Communications Complexes (except circuits between the Goldstone Complex and JPL) which are the responsibility of NASCOM. The Voice, Teletype, High-Speed Data, Wideband Data, and GCF Monitor and Control Subsystems point-to-point communications are serviced by this Fig. 1 configuration.

### A. High-Speed Data Subsystem

This subsystem shall consist of GCF assemblies that switch, transmit, record, process, distribute, test, and monitor digital data and is used for transmission of:

(1) All digital data for the DSN Command, Tracking, and Monitor and Control Systems.

(2) All low or medium rate data of the DSN Telemetry, Radio Science, Very Long Baseline Interferometry (VLBI), and the DSN Test Support System.

The High-Speed Data Subsystem provides a capability for transmitting and receiving the serial bit stream block formatted data over a properly conditioned full duplex alternate voice/data channel having a 3.0-kHz bandwidth or over time division multiplexed 56 kb/s satellite circuit. This serial bit stream is impressed on communication circuits at a continuous line bit rate divided into message segments referred to as high-speed data blocks.

Two types of data blocks are used:

(1) Data blocks containing user data bits to be transmitted.

(2) Filler blocks containing filler data bits provided by the GCF when the user data rate is insufficient to maintain contiguous blocks on line required for continuous line monitoring and error control.

Current capabilities for the GCF Mark III period provide the functional capabilities illustrated in Fig. 2. The GCF High-Speed Data Subsystem is standardized on a 1200-bit block size (message segment) and a line bit rate of 7200 b/s. The subsystem capabilities include a 22-bit error detection encoding/decoding polynomial code, two error status bits, and error control using two eight-bit fields. The error control field facilitates numerical serialization and acknowledgement numbers for error correction by retransmission.

The error correction capability has significantly reduced the post-pass time required for non-real-time replay of blocks received in error.

Figure 3 illustrates the High-Speed Data Subsystem and the GCF Subsystems located at the Central Communications Terminal.

### B. Wideband Data Subsystem

The Wideband Data Subsystem consists of assemblies that switch, transmit, receive, process, distribute, test and monitor data requiring the use of bandwidths greater than those provided by standard high-speed data channels. The GCF Wideband Data Subsystem functionally illustrated in Fig. 4, together

with a listing of functional capabilities provided, includes standard wideband circuits as well as all intersite microwave (area microwave) capabilities. The Wideband Data Subsystem is used for the transmission of:

(1) All DSN Telemetry System high-rate data that exceed High-Speed Data Subsystem capabilities.

(2) Data interchange between the NOCC and GCF Communications Terminal at JPL.

(3) Data interchange between DSSs within a complex via intersite microwave, including critical timing signals and receiver baseband signals for antenna arraying and signal combining systems support.

(4) Simulation System Data from the Mission Control and Computing Center/Mission Operations Center to the DSSs.

(5) DSN Test Support System data from the Network Operations Control Center to the DSSs.

The wideband data circuits for interchange of data between the DSSs and JPL are impressed with serial bit streams at a continuous line rate, typically 56 or 230.4 kb/s, divided into 4800-bit message segments (data blocks). Similar to the High-Speed Data Subsystem, the blocks are either data blocks or filler blocks inserted when the user data load is insufficient to maintain contiguous data blocks.

## C. Voice Subsystem

The Voice Subsystem consists of GCF assemblies that switch, transmit, receive, distribute, test, and monitor transmissions originally generated in vocal form, and includes internal voice communications within the Deep Space Station Communications Complexes, DSSs, and the NOCC. The subsystem service provides capabilities between those areas and to non-DSN area interfaces as follows:

(1) NOCC and DSS.

(2) NOCC and MCCC/MOC (or remote MOC).

(3) MOC and DSS for Command System backup.

The Voice Subsystem functional capabilities and key characteristics include:

(1) Standard voice-data grade circuits for all traffic.

(2) Conferencing capability on one intercontinental circuit during noncritical periods for all Deep Space Stations supporting a single project (individual circuits for each DSS during critical periods, resources permitting).

(3) User-controlled intercomm switching.

(4) Circuits used for high-speed data transmission (backup) if required.

(5) Voice traffic recording in the central communications terminal upon request.

## D. Teletype Subsystem

The teletype (TTY) subsystem uses an eight-level ASCII (American Standard Code for Information Interchange) national standard.

The subsystem consists of assemblies that switch, transmit, receive, distribute, test and monitor digital signals at a TTY line rate of 100 words per minute. The operational use of teletype continues to be deemphasized and is used primarily for emergency backup operational transmissions and administrative communications. Service functions and key characteristics include:

(1) Handling Air Force Eastern Test Range (AFETR) generated predicts for DSN initial acquisition.

(2) Transmitting nonoperational messages between the JPL Message Center and other locations.

(3) Use of standard NASCOM format and the NASCOM communications processor for message switching.

(4) Employment of time division multiplexing techniques to reduce trunk circuit costs.

## E. GCF Monitor and Control Subsystem

The GCF Monitor and Control Subsystem consists of assemblies that collect, process, and display the status and performance of the GCF Subsystems in real-time. The GCF Monitor and Control Subsystem functional capabilities are illustrated in Fig. 5. Functions are implemented in minor subassemblies located at each DSS to interface station GCF status and performance indicators to the CMF for monitor block formatting and transfer to the Central Communications Monitor (CCM) Processor at JPL. The CCM also receives real-time status and performance information from local GCF subsystems. All real-time status and performance information received by the CCM is processed and displayed relative to preset standards and limits to facilitate operations monitoring and technical control. Information and alarms are displayed on continuous line performance and data flow throughput including error control.

## F. Data Records Subsystem

The DSN requirements for the data record processing and production functions are implemented in the GCF Data Records Subsystem. The Data Records Subsystem consists of assemblies in the CCT that log in real-time, monitor, iden-

tify gaps, provide for processing and editing of data gap lists, control data gap recalls from the DSSs and the generation and accounting for Intermediate Data Records (IDRs) and fill-data records selected from records of the GCF real-time log.

The Data Records Subsystem maintains accountability of high-speed and wideband data, performs automatic recall (under operator control) of missing data, and generates near-real-time intermediate data records for delivery to the Mission Control and Computing Center.

### G. Network Communications Equipment Subsystem

The Network Communications Equipment (NCE) Subsystem consists of GCF minicomputers and peripheral I/O assemblies that switch, transmit and receive data. The NCE Subsystem assemblies are located in the Network Data Processing Area (NDPA). The NCE assemblies comprise a GCF Area Communications Terminal located adjacent to the NDPA.

The NCE interchanges multiplexed block formatted data with the ECS assembly located in the CCT over three full-duplex 230-kb/s wideband data channels (see Fig. 6). This GCF-NDPA interface function provides for:

(1) Processing data for transmission to and accepting data from the GCF CCT.

(2) Multiplexing/demultiplexing and buffering data for all NDPA processors.

(3) Routing data to and from NDPA processors.

## V. Typical Configuration

The DSN GCF is designed for multiple mission support. Improvements and additions are integrated to meet new era and project requirements (Voyager and Pioneer-Venus requirements were completed in CY 1978) while continuing to support the Viking, Helios, and Pioneer 6 through 11 Projects. Plans for west coast switching center support of the IRAS Project are completed. Requirements for the ISPM and Galileo projects are being defined and included in ongoing plans. Figure 7 illustrates, in general, the GCF configuration for support of these projects. Additionally, remote information centers and other non-DSN NASCOM-serviced installations on the west coast are serviced through the NASCOM West Coast Switching Center, an integral part of the GCF 20/Central Communications Terminal at JPL.

## VI. Implementation Activities

### A. GCF-NOCC Reconfiguration

The task to reconfigure the CCT was begun with requirements specifications established in March 1979 and a responsive design review in October 1979. The reconfigured CCT was completed in the second quarter of 1981. The reconfigured CCT has:

(1) Reduced the number of computers in use.

(2) Provided a more timely IDR production capability.

(3) Reduced the complexity of the man-machine interface.

(4) Reduced manpower required at the CCT.

(5) Released the wideband data subsystem coded multiplexer/demultiplexers for reinstallation at the 34-meter DSSs.

(6) Increased the capacity to handle simultaneous wideband data lines at the CCT.

(7) Provided a central console at the CCT for control and monitoring of all GCF subsystems.

### B. Mark IVA GCF

Implementation planning and design to meet the requirements of the DSN's Mark-IV Long Range Plans was begun in late 1979. The priority design and implementation effort will be to consolidate the data communications interfaces to a single communications processor interface at each of the three Deep Space Communications Complexes (DSCC). This single interface processor will accommodate the DSN's Networks Consolidation Program plans for a single signal processing center located at each DSCC. Other less significant GCF changes and additions required within the DSCCs along with changes at the CCT are included in this effort, which is to be completed in 1985. With the consolidation of the networks, the GCF will see a significant increase in data communications requirements as both deep space and high elliptical near-earth orbiting spacecraft are tracked, and acquired data will pass through the single communications processor at the DSCC. Modifications will be made in the CCT to accommodate the new DSCC interfaces and the new type of remote operations control centers.

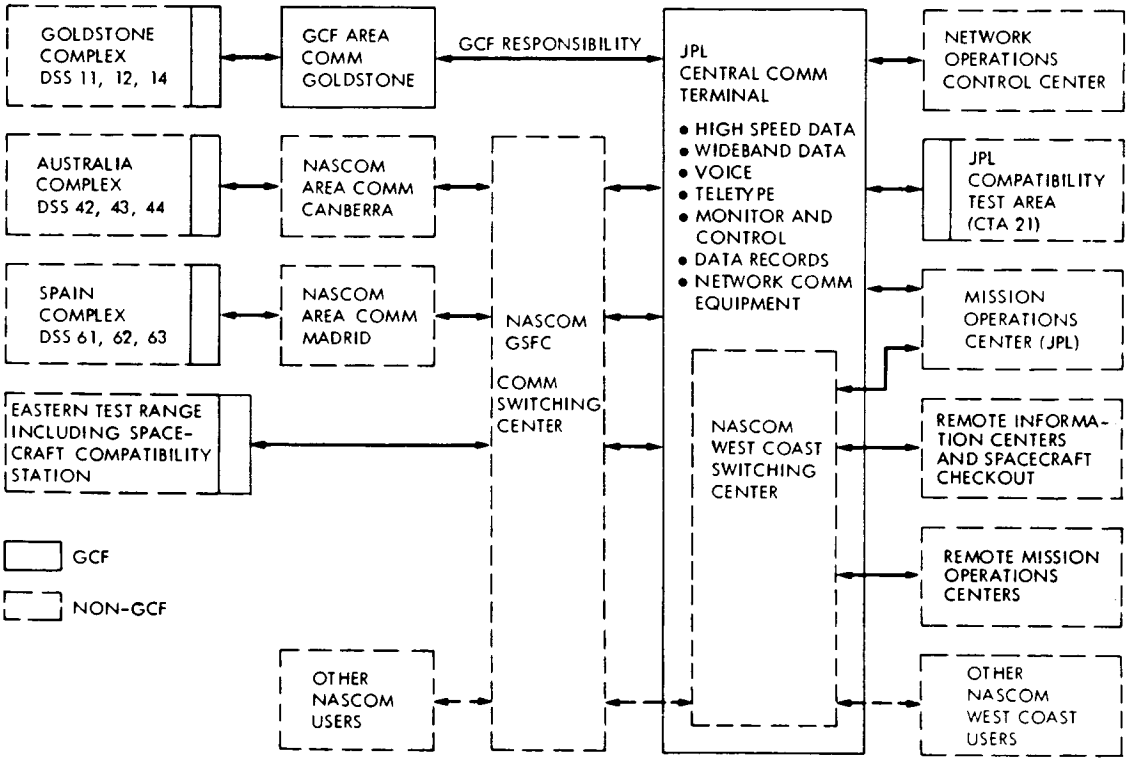



**Fig. 1. GCF configuration**

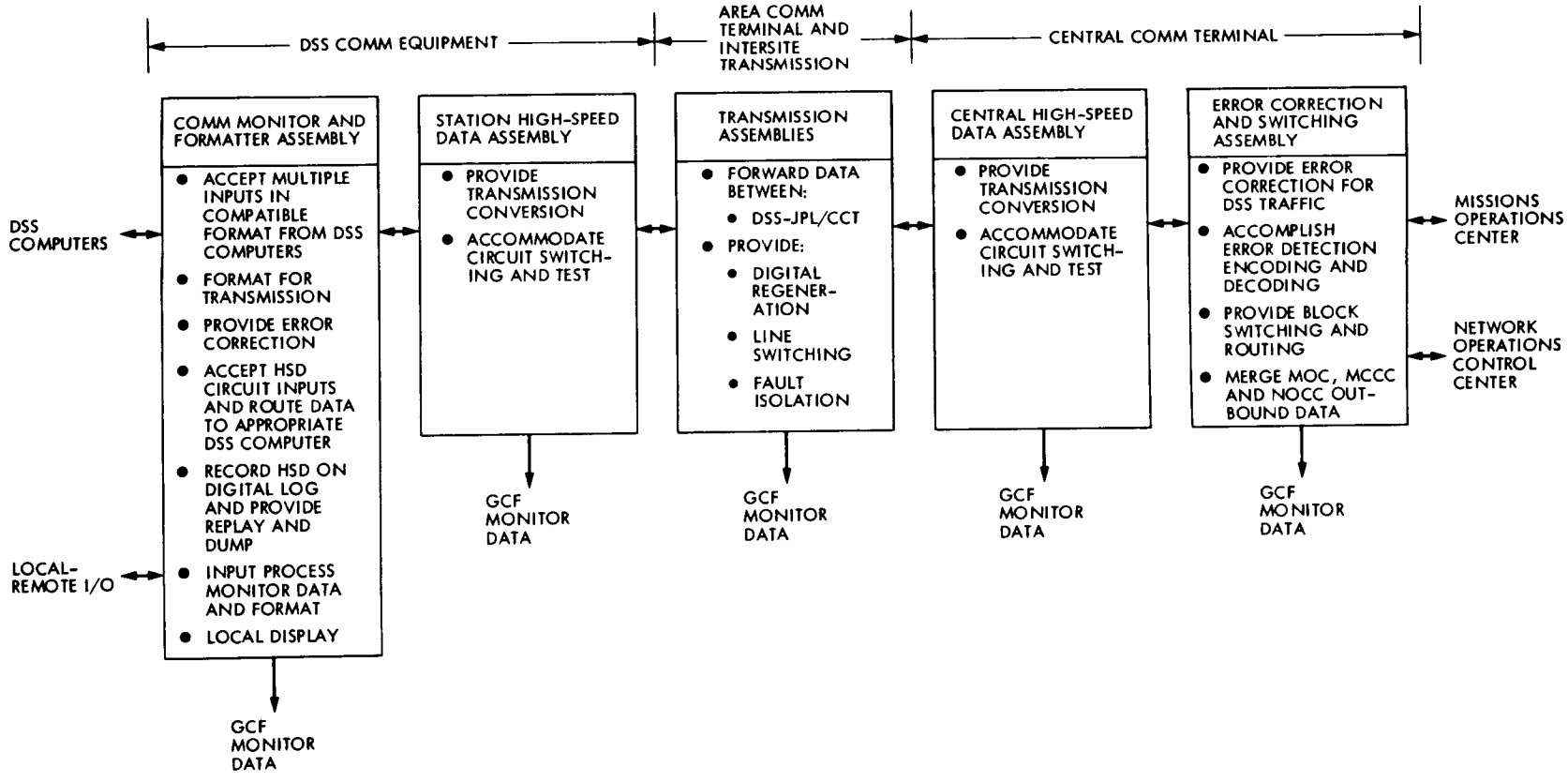



**Fig. 2. GCF High-Speed Data Subsystem functional capabilities**

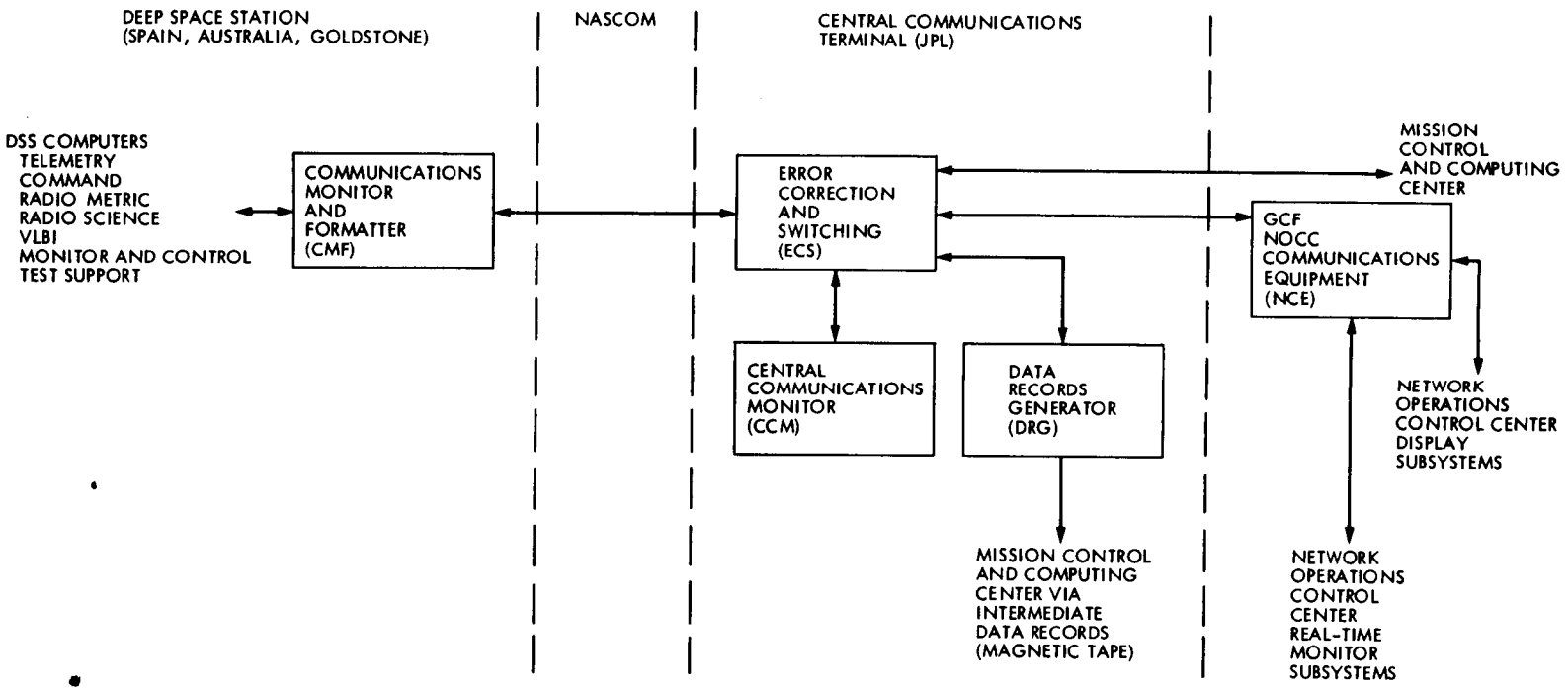



**Fig. 3. GCF High-Speed Data Subsystem and Central Communications Terminal Subsystems and interfaces**

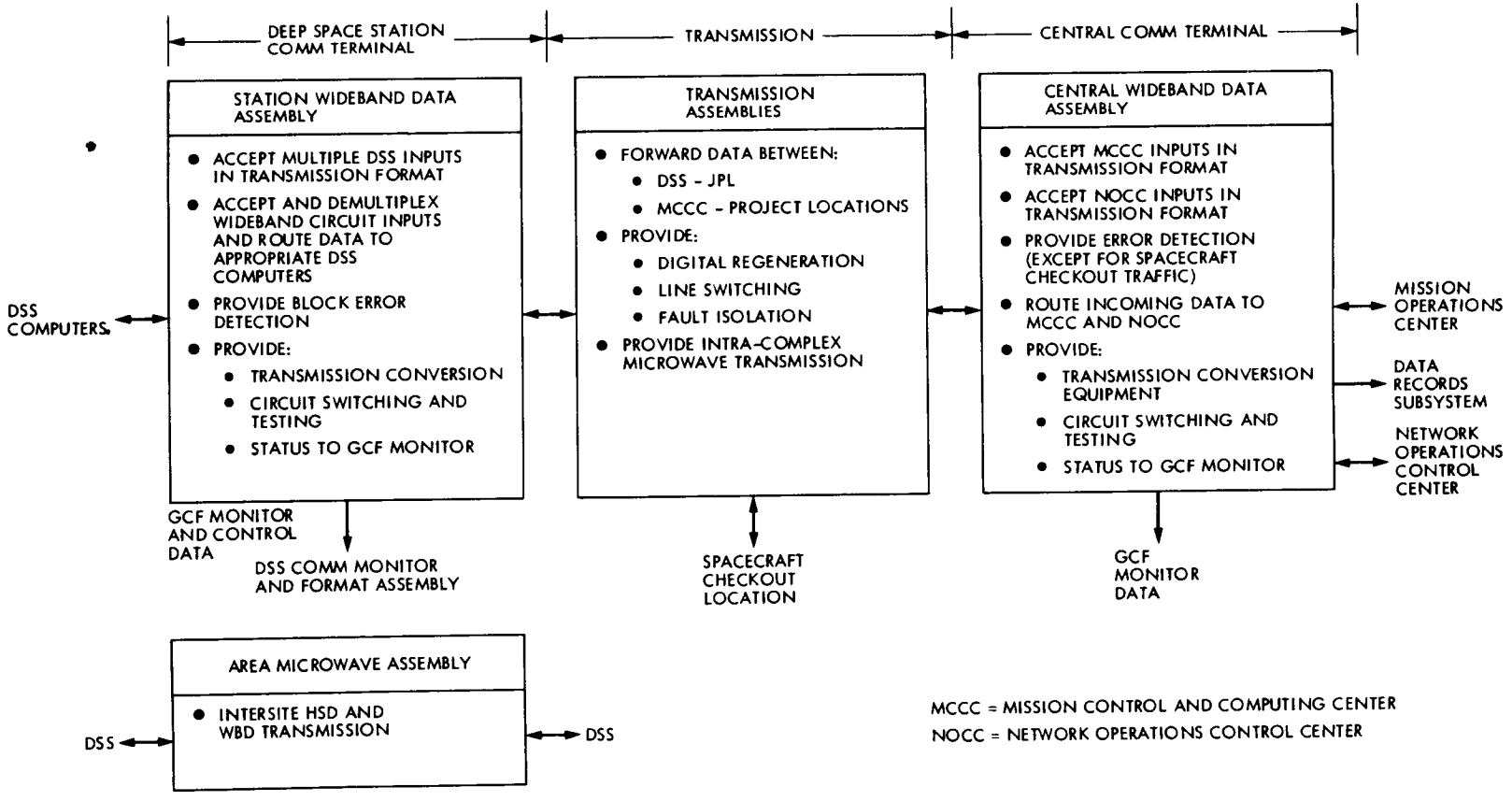



**Fig. 4. GCF wideband subsystem**

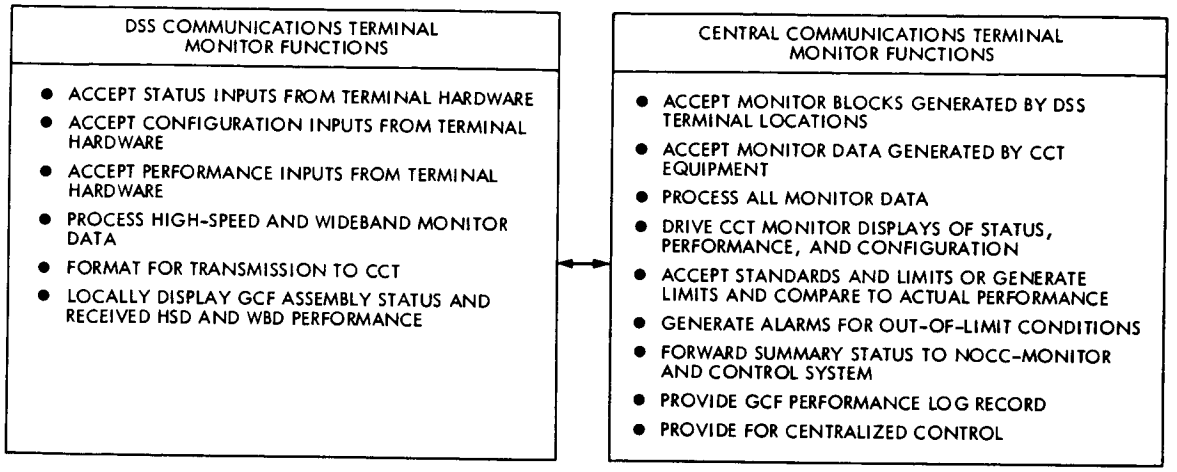



**Fig. 5. GCF Monitor and Control Subsystem functional requirements**

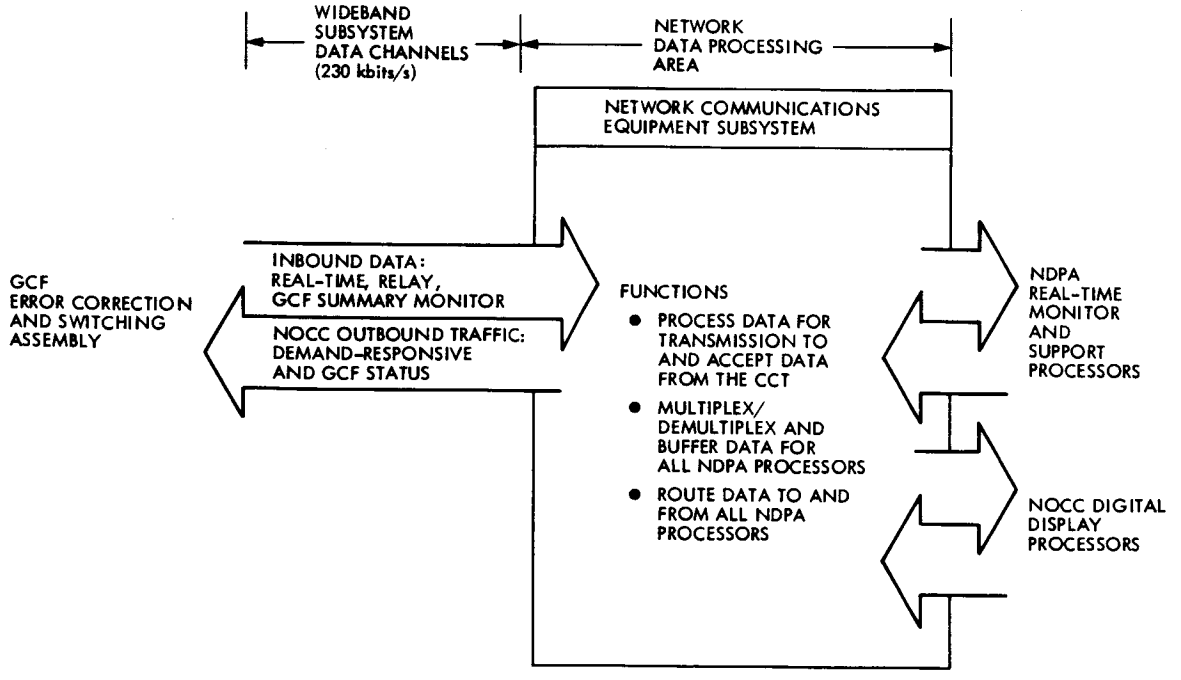



**Fig. 6. GCF Network Operation and Control Communications Equipment Subsystems**

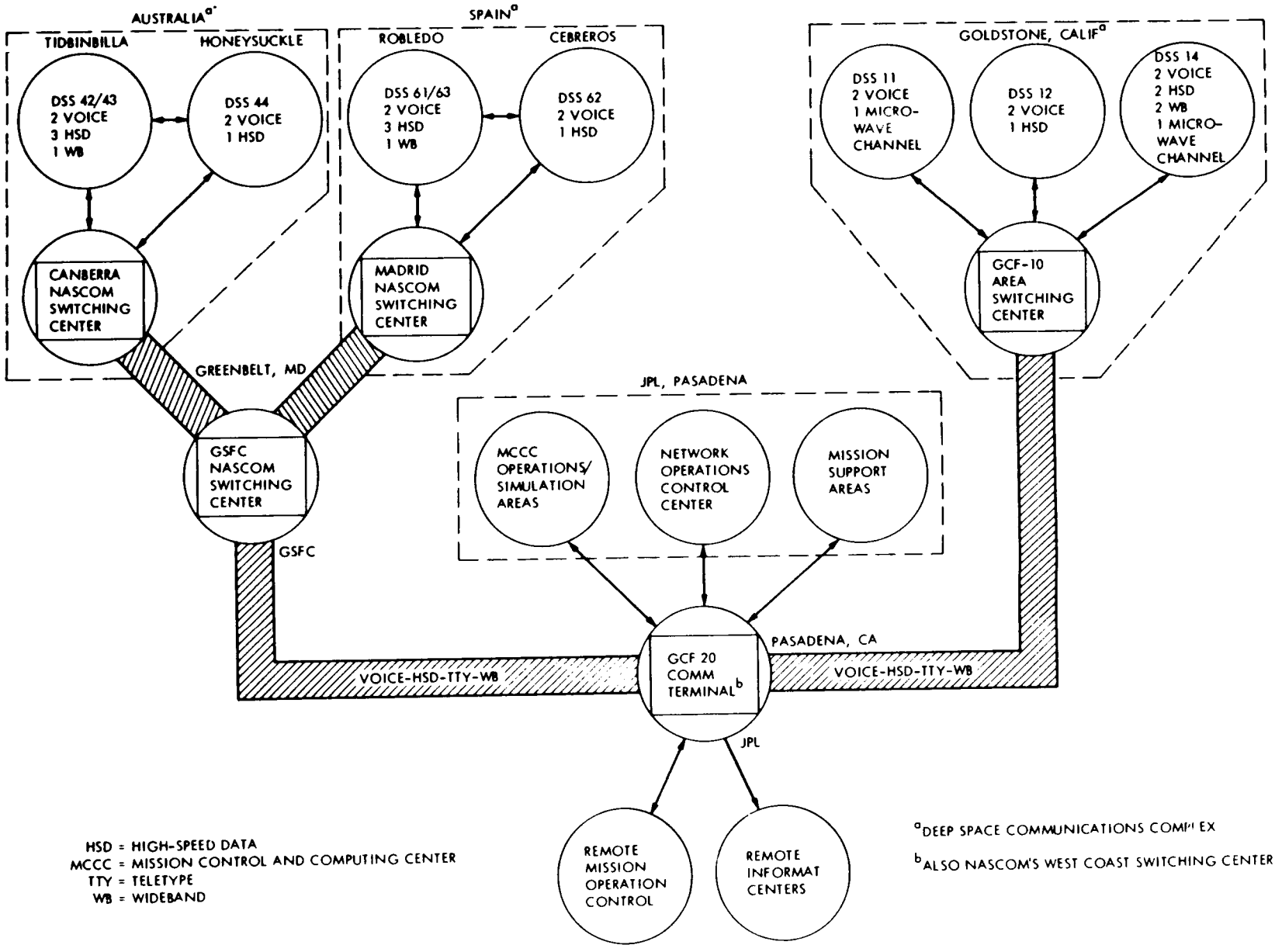



**Fig. 7. DSN support locations and GCF-NASCOM circuit requirements**

N82-11283



# Networks Consolidation Program

M. L. Yeater and D. T. Herrman
Office of Telecommunications and Data Acquisition

*The Networks Consolidation Program (NCP) was established by NASA in 1979 to implement recommendations of the Networks Planning Working Group which were presented at NASA Headquarters in October of that year. The goals of the NCP were defined to include the planning, designing and implementing of a single consolidated network of ground tracking stations to replace the present two ground-based spacecraft tracking networks.*

*The proposed consolidated network will make use of some facilities that are now included in the Ground Spaceflight Tracking and Data Network (GSTDN), operated by Goddard Space Flight Center (GSFC), as well as the existing Deep Space Network (DSN), operated by Jet Propulsion Laboratory (JPL). These facilities will be combined and modified to provide a consolidated network managed and operated by JPL and capable of supporting the set of planetary and Highly Elliptical Earth Orbiter (HEEO) missions planned for the Tracking and Data Relay Satellite System (TDRSS) era.*

*The consolidated network will come into being after the TDRSS becomes operational. It will continue to support planetary missions that are now supported by the DSN and also provide support to a broad class of other spacecraft missions which are not compatible with or which, for other reasons, cannot be supported by the TDRSS.*

*A previous report in the TDA Progress Report series traced the history of activities and events that led to the decision to consolidate the NASA ground tracking and data networks into a single network. It included a summary of the NCP management and planning activities that have taken place at Jet Propulsion Laboratory, Goddard Space Flight Center, and at the NASA Headquarters Office of Space Tracking and Data Systems (OSTDS), from the October 1979 decision to proceed to August 1980. This report shows the progress of the NCP since the previous report to the present with special emphasis on the planning and budgeting activities that have occurred.*

# I. Introduction

The first progress report on the Networks Consolidation Program (NCP) appeared in *The Telecommunications and Data Acquisition Progress Report 42-59* (Ref. 1). It included a history of the development of the NCP from the initial meetings of the Networks Planning Working Group through activities completed in August 1980. This report addresses management, design, and implementation activities that have occurred since the previous article.

# II. Program Accomplishments

## A. NCP Organization

When the decision was made to proceed with the NCP as a major NASA program, two teams were formed at JPL to do planning and design tasks. A Planning Team, with representatives of involved organizations at JPL, was chartered to develop program guidelines and plans. A System Design Team, with representatives from the implementation and operations organizations, was formed to develop system functional requirements, system designs, implementation alternatives and plans. A third team was formed for execution of the implementation of the plan when the first two teams completed their work. This team, the Implementation Team, has absorbed the functions of the System Design Team, and operates under the direction of a manager who is responsible for all NCP implementation activities. In addition, a Steering Committee was formed to provide support, review, and advice to the NCP Manager on a continuing basis. It reviews progress of planning, design, and implementation tasks and makes recommendations to the program manager. It also conducts special studies to help the NCP Manager make informed decisions when considering proposals regarding implementation alternatives.

## B. Formal Reviews

A series of formal reviews was established at which the program management reported on planning, design and implementation activities of the program. Two of the reviews were reported in Ref. 1. Two others were held during the period of this report. The third review covered the functional designs of the systems and subsystems that will make up the consolidated network. This review was rated as "very well done" by the review board, although there were many "requests for action" submitted by attendees and board members.

Two points were raised and discussed. One was that NASA OSTDS had directed NCP management to consider the addition of requirements for backup support to the Space Shuttle, TDRSS, and certain HEEO missions. Since these extensions had not been considered in the baseline mission set, the review board recommended that a special team be formed to study the impact to NCP plans of adding these requirements. They further advised that this effort should be kept apart from, and operate in parallel with, the main design activities and should not interfere with the baseline effort. The second point was that the overall plan appeared to be success oriented, with little visible schedule contingency.

The fourth formal review covered the Networks Consolidation Implementation Organization and Plan. The board was asked if the plan was adequate to accomplish the networks consolidation goals and if it would recommend that the program should proceed.

The review board noted that the review was excellent in format and detail; they recommended that the program should proceed according to the implementation plan. They did take special note of the continuing growth of requirements beyond the original baseline and the continuing level of schedule contingency visible in the plan. It was agreed that a fifth program review should be scheduled after completion of the special parallel study, initiated during the third review. It was felt that this added review should lead to recommendations regarding the inclusion or rejection of proposed extensions to the NCP baseline mission set. The fifth review, which will follow the format and agenda of the fourth review, is scheduled for October 21-22, 1981.

## C. Budgets

The budgetary picture of the NCP came to a first level of completion in a presentation to the NASA Comptroller in July 1980. In preparing for that presentation, all of the major tasks for the NCP were identified and costed in such a manner as to be consistent with the overall budget that Headquarters had allocated to the program. It should be noted that the cost structure was based on materials presented in Ref. 2. These included assumptions regarding management organization, station equipment availability, mission support requirements and schedules for all affected missions.

In recent months the program has experienced a succession of changes in most of the areas in which assumptions had been made. Overall funding levels have been modified to reflect changes in mission schedules – some have slipped and some have been cancelled – and budgetary reductions brought about by decisions of the new Administration. Overall, NCP costs remain within the boundaries estimated in July 1980.

## D. Antenna Array Alternatives

The report of the working group (Ref. 1) included the Large Advanced Antenna System (LAAS), an array of new 40-meter antennas, as an element of the consolidated network. This antenna array, which was to provide increased telemetry

capability for support of the Voyager spacecraft at Uranus encounter, was to be at the Goldstone, California, DSN complex. Since the LAAS will not be developed, the NCP must initiate a plan to array antennas at each complex and to develop the additional antennas needed to provide the effective aperture to support Voyager Uranus data requirements.

Initial plans for antenna arraying included colocation of all the antennas at each complex to the locations of the present 64-meter antennas and conversion of all 26-meter antennas to 34-meter antennas. Colocation has been questioned from two major points of view. First is the question of overall costs. Studies show the higher front-end costs of moving the antennas to a central point at each complex are quickly offset by savings in the continuing maintenance and operations costs when compared to the costs of continuing staffing and support at the present diverse locations. Thus from a cost standpoint the better course of action is to colocate antennas. A second concern was radio frequency interference (RFI) among the colocated antennas. This question has been studied by systems and engineering personnel. The study shows that the potential for RFI is not sufficient to deter the decision to colocate the antennas at a complex.

The types and number of antennas planned for each complex after the networks consolidation is completed are:

| | |
|---|---|
| 1 64-meter antenna: | the existing antenna in its present location |
| 1 34-meter antenna: | the existing antenna which has been converted from a 26-meter antenna, relocated to the site of the 64-meter antenna |
| 1 9-meter antenna: | the existing 9-meter antenna, relocated to the site of the 64-meter antenna |
| 2 34-meter antennas: | representing the residual 26-meter antennas from the GSTDN and the 26-meter subnet of the DSN, each converted to 34-meter antennas and relocated to the site of the 64-meter antenna |

There is a question of whether it is better to convert 26-meter antennas to 34-meter antennas or to procure new antennas which would be installed directly at the colocated sites. This is now being studied and includes questions of life-cycle costs and comparison of technical capabilities of the competing antenna alternatives. No firm decision on this point has been made.

## E. Mission Set

The initial networks consolidations planning showed TDRSS, once it became operational, providing support to all low earth-orbital missions. The remaining or ongoing high earth-orbiting or libration-point missions would remain the responsibility of the residual GSTDN stations until those stations were closed or converted for inclusion in the consolidated network. It was assumed that only the Origins of Plasma in the Earth's Neighborhood (OPEN), the Active Magnetospheric Particle Tracer Experiment (AMPTE), and the International Sun-Earth Explorer (ISEE-3) missions from this class would continue to require support beyond that time. The DSN would continue to support all planetary or deep space missions. The new consolidated network would accept continuing responsibility for support of planetary and a limited set of high earth-orbiting missions. It would not support any low earth-orbital missions. Any such missions which were not compatible with TDRSS were expected to terminate. The mission set to be supported by the consolidated network was restricted to the prime mission periods for the missions identified in Fig. 1-3 of Ref. 2 and no other missions were included in the baseline set.

As previously noted, OSTDS has directed program management to evaluate the cost, schedule, and capability impacts associated with the addition of some missions to the baseline. This effort is now underway and will be addressed in future program reviews. A copy of the current NCP Mission Set is shown in Fig. 1.

## F. Transition Plan

The first edition of the Transition Plan was completed in January 1981 and has been used in initial transition planning activities. Owing to the large number of actual and proposed changes in the scope of NCP mission coverage and in the dates of coverage that will be required, the original version of the plan has been revised and was published in June 1981.

The accomplishment of the Networks Consolidation Program will require an extensive effort to permit the transfer of responsibility for support of individual projects from the GSTDN to either the TDRSS or to the new consolidated network in such a way as to assure that there will be no adverse impact on the flight projects. The NCP Transition Plan, Volume II of the series of NCP plans, will control the transition

process. It will encompass the phasing over of management functions and technical support, as well as the transfer of equipment from the GSFC-managed STDN stations to the JPL-managed Consolidated Network.

### G. Mission Support Plan

The development of a Mission Support Plan for the NCP commenced during the second quarter of FY 1980. This task has involved a major effort to determine the network load that must be supported by the consolidated network. The state of flux that has affected the NCP Baseline Mission Set has had a resulting effect in the determination of network loads which cannot be considered firm until the mission set stabilizes.

Based on all available information, the first draft of the Mission Support Plan was released for review and comment. The document is now in the final stages of the approval cycle and will be released in September 1981.

### H. Implementation Plan

The development of the NCP Implementation plan commenced in the third quarter of FY 1980 under the auspices of the NCP System Design Team. This effort progressed through the end of the fiscal year and resulted in the release of the first draft of the plan in January 1981. This draft represented the basic implementation concepts for the consolidated network as they were presented at the fourth formal review of the program. A second draft of the plan has been prepared and is being circulated for internal review and comments. This document is subject to further changes as decisions are made with respect to antenna alternatives and mission support requirements. Figure 2 represents the current status of the NCP planning and design schedule as developed in the implementation plan.

## References


1. Layland, J. W., "OSTDS Networks Planning for the TDRSS Era," OSTDS Report, Office of Space Tracking and Data Systems, NASA, Jan. 1980.
2. Yeater, M. L., Herrman, D. T., and Sanner, G. E., "Networks Consolidation Program," *TDA Progress Report 42-59, July and August 1980*, p. 107, Jet Propulsion Laboratory, Pasadena, Calif., Oct. 15, 1980.

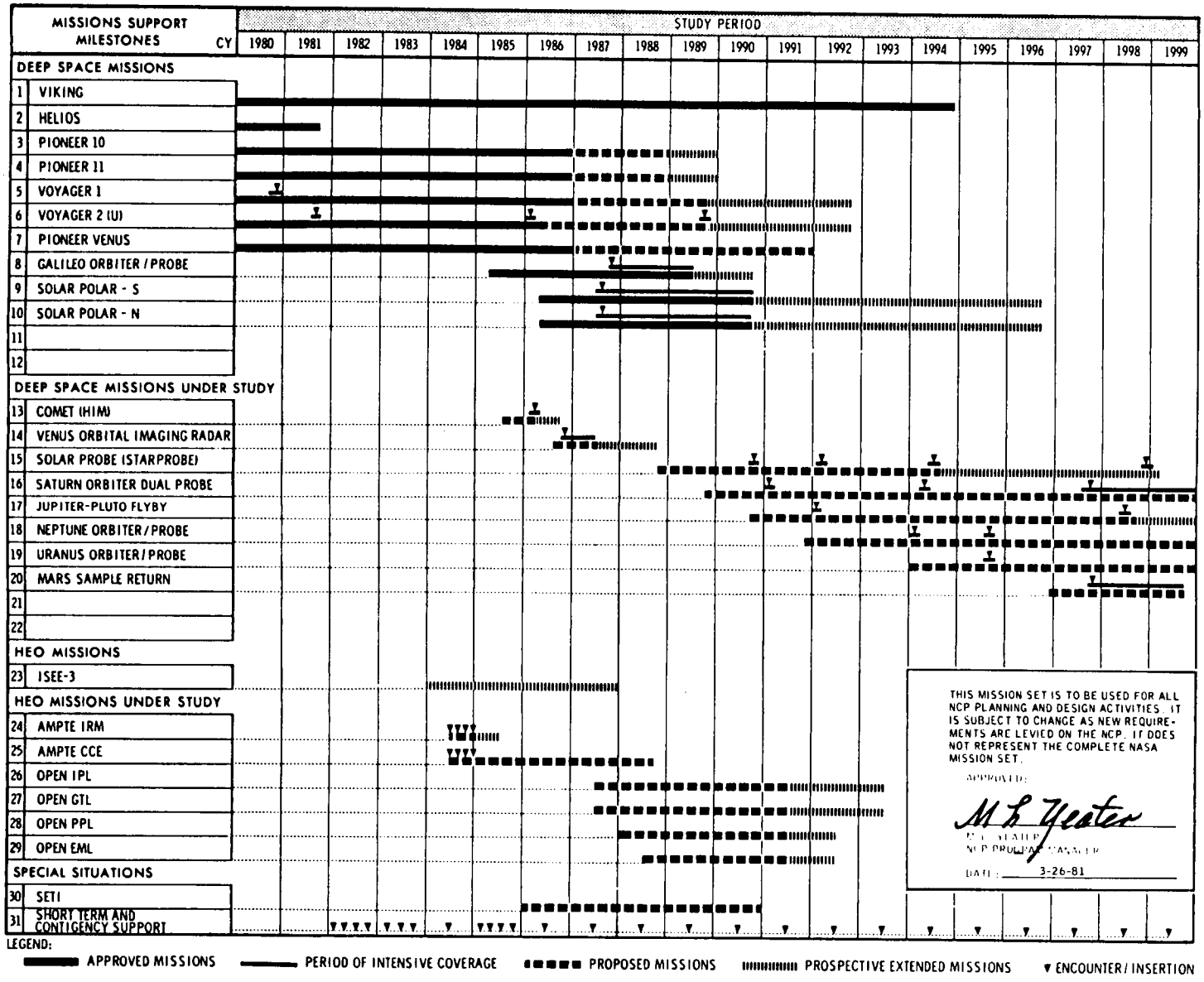



**Fig. 1. NCP flight mission set**

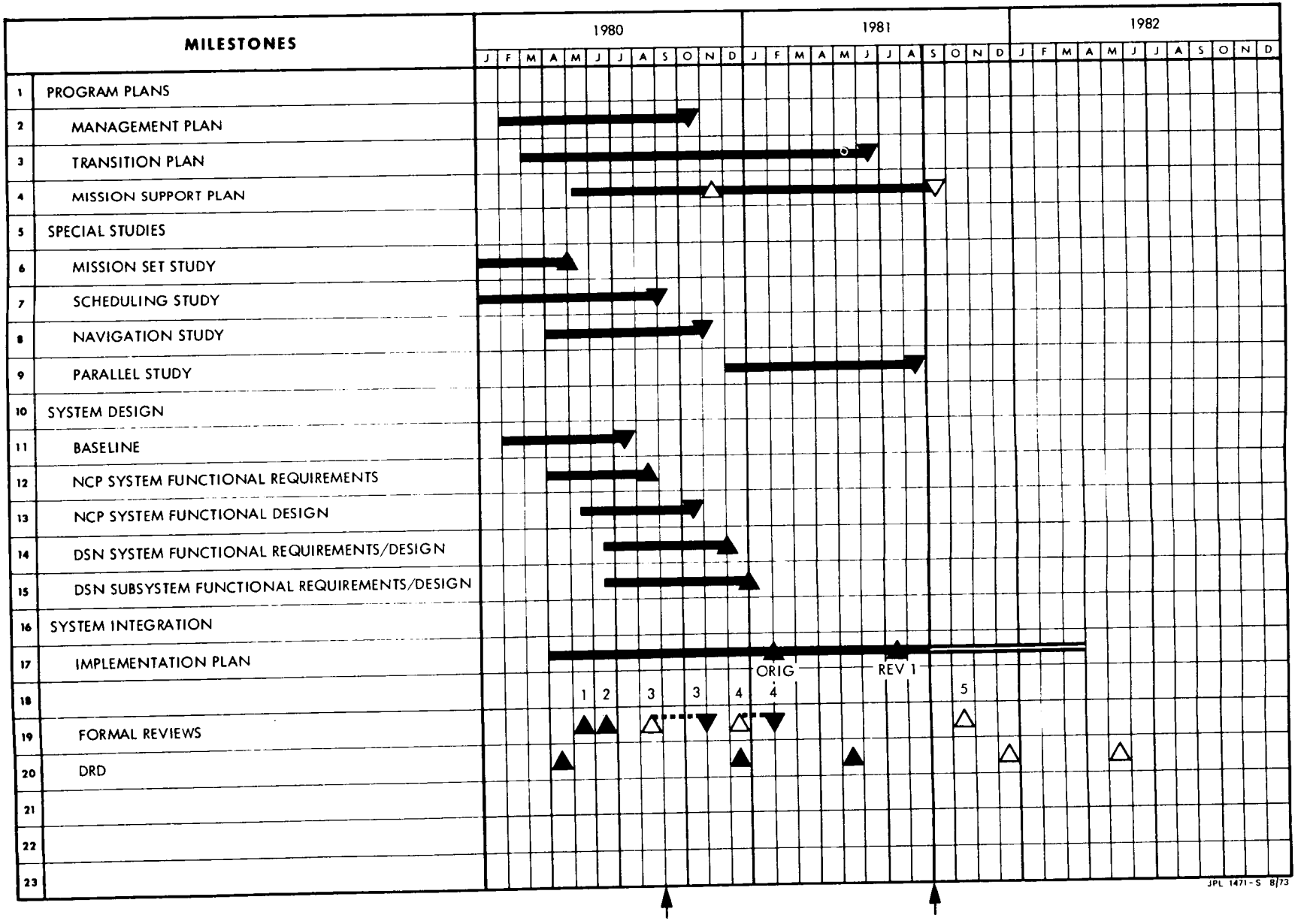



**Fig. 2. Planning and design schedule**

N82-11284



# DSN Model for Use In Strategic Planning

K. C. Kelly
TDA Program Control

C. Y. Lin and M. McKenzie
Telecommunications Science and Engineering Division

*A System Dynamics Model of the DSN is being developed to support strategic planning for the Network. Applications for the model are described, as well as the foundations of System Dynamics and the methodology being used to develop the model. Activities to date and plans for future work are also discussed.*

## I. Introduction

In the course of planning, some of the more difficult questions for managers to address are of the following kind:

What behavior patterns does our system exhibit, and what causes them?

What effect would a new policy have on the behavior of our system?

What policy, decision, or action will create a given desired change in our system's behavior?

What would be the effect on our system of a major outside event?

These are strategic questions – questions that address "should" more than "how" and that appear throughout the process of long-range planning. Such questions, and the difficulties involved in answering them, prompted the formation of the TDA Network Dynamic Model task.

This task provides for the development, testing, and validation of one or more models of the Deep Space Network system. These DSN Models (DM) will be designed to support TDA managers in strategic planning for the Mark IV and Mark V eras (i.e., planning for the next 20 years). The modeling technique chosen for DM is System Dynamics; its strength lies in the ability to model the *structure* of a complex system, thus enabling a manager to better understand the forces producing the system behavior and effect desired changes thereto.

The intent of this article is to describe the System Dynamics technique and its application to the DSN. Section II discusses the foundations of System Dynamics and Section III the methodology that will be used to produce, test, and validate the DSN Model. A summary of progress to date and a discussion of future work are given in Sections IV and V.

## II. Fundamentals of System Dynamics

The System Dynamics technique is a tool for policy design and analysis in the environment of a complex system. Although a relatively new discipline, System Dynamics has been successfully applied to the strategic planning process in various major corporations and R&D organizations.

System Dynamics was developed at M.I.T. under the direction of J. Forrester. Its foundations lie in the study of feedback system behavior in physical systems; this approach was then extended to the analysis of social systems. In System Dynamics, cybernetic theory is combined with perceptions of the system structure, derived primarily from managerial experience. Using computer simulation, a model of system structure is developed in which policy can be designed, and the consequences of that policy evaluated.

## A. Characteristics of Complex Systems

The DSN is representative of a large, complex system. One approach to solving problems in such a system is to institute a new policy and try it out "in vivo." Among the difficulties associated with this approach are: (1) the policy may not work and the system is left in worse condition than before the intervention, (2) if the policy is effective, it may not be possible to determine why, and (3) it may be very difficult to assess how the same policy might work under changed conditions or in other parts of the system (Ref. 1).

Instead of testing policies in the real system, it is possible to use System Dynamics to model the dynamic, or time varying behavior of complex systems. The interrelationships between policies and system state are analyzed, and high and low leverage policies identified.

There are some very compelling reasons to try out new policies "in vitro," i.e., in a simulation environment. These reasons are based upon the behavioral characteristics of complex systems:

(1) The nature of complex systems renders them resistant to most policy changes. Only 5 to 10% of the attempted changes are effective, as much as 50% of the changes are detrimental, and the rest have no effect (Ref. 2).

(2) Complex systems exhibit counterintuitive behavior. The manifestation of a symptom may be far removed in time and space from its cause. Unrelated conditions resulting from the dynamics of the system structure may have an apparent cause and effect relationship due to their juxtaposition in time and location (Ref. 3).

(3) Complex systems can be controlled through leverage points: while they are quite insensitive to changes in most of their parameters, it is possible to identify points where a small policy change can produce a large change throughout a system. There are relatively few such high leverage points, and they frequently are in unexpected places (Ref. 3).

(4) The short-term response to change in a complex system frequently occurs in the opposite direction to the long-term outcome (Ref. 3). This necessitates policy tradeoffs such as whether to implement an improvement in the present at the expense of the future.

(5) External assistance applied to the system may be internally counteracted (Ref. 3). The burden of system improvement is transferred to the intervening agent because the assistance in effect becomes part of the system.

(6) Complex systems can appear to exhibit lower performance than expected. This is due to the detrimental effects of design changes suggested by the counterintuitive nature of the system, and by the opposing directions of short- and long-term responses to change (Ref. 3).

Complex systems do not lend themselves to traditional linear mathematical modeling techniques. As implied above, these systems exhibit nonlinear behavior and changing behavioral modes. For example, such a system may move from a condition of equilibrium, through a period of exponential growth, then proceed to exhibit oscillatory behavior. System Dynamics, unlike other modeling techniques, accommodates these nonlinear behavioral modes. It tracks the complex, dynamic relationships of system variables, demonstrating the trends of system levels over time. This, because of the subtlety and complexity of the interactions, is impossible to do over an extended period of time on an intuitive or mental basis. It should be noted that System Dynamics is not intended to do quantitative, point predictions of the future.

## B. Basic Precepts of System Dynamics

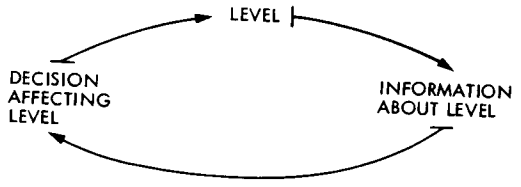
The fundamental construct of System Dynamics is: *system structure determines the behavior of a system.* While other techniques explain events as caused by other events, or behavior as part of a time series, System Dynamics views events as part of behavior patterns that are themselves caused by the underlying system structure. Structure, in System Dynamics parlance, is defined as the system states and their interrelationships as determined by policy. Policy is defined as the rules that describe how available information is used to determine action.

Another primary construct of System Dynamics is: *the model of system structure is drawn largely from mental data.* Most of the world is run on the basis of mental information; here resides most of the knowledge about policy and the reasons for why things have happened in the past. By the time mental information is filtered down into the written record, much of the time-sensitive rationale for policy formulation is lost. The pressures that existed causing a particular decision to be made are not recorded. Furthermore, when written information is further condensed into numerical data, it loses the peripheral information that influenced the decision.

Consequently, the modeler should *determine the structure of the system by using the information that most of the decisions and policies are based upon, the mental information, rather than relying upon numerical data to infer structure* (Ref. 2). For this reason, managerial participation is imperative in the conceptualization phase of the DSN model development effort.

### C. Elements of System Structure

Complex systems consist of multiple, interlocking feedback loops. A feedback loop is a closed path that connects (1) a system level (state) to (2) information about the level to (3) a decision that controls action concerning the level. Information is the observed or apparent level of the system and may not represent the true level.




Feedback loops, the organizing principle of system structure, account for both the growth and stability of the system. Positive loops are goal divergent, moving exponentially from some point of equilibrium (thus accounting for growth or decay). Negative loops are goal seeking, or stabilizing structures. In a negative loop, any disparity between an apparent state and a desired goal generates inormation feedback that influences the system in a direction that reduces the discrepancy (Ref. 2).

It was previously stated that typically only 5 to 10% of those policies instituted to improve a given system are effective. This is because a great many of the attempted interventions get trapped in these negative, or compensatory, feedback loops. Here, a change applied to move the system in one direction produces a system change in the opposite direction.

The changing behavioral modes exhibited by a complex system may be accounted for by the nonlinear coupling of the various feedback loops within it. This condition can allow one loop to dominate the system for awhile, then cause control to shift to a loop elsewhere in the system. Thus, the same structure is producing behavior so different that it seems unrelated.

The System Dynamics modeling process involves transforming all elements of the real system into a feedback structure containing just two types of variables: levels and rates. Levels are the accumulations within the system – they describe the system state at any point in time. Rates are the system policies – the rules that determine what decisions are made affecting the levels. They reflect goals, observed conditions, the discrepancies between the two, and action.

Once the system structure is defined in these terms, a computer simulation technique (finite forward difference) is applied. The essence of this technique is, quite simply, that the value of a system level at time $t$ is equivalent to its value at time $(t - 1)$ modified by the difference between the growth rate and the loss rate over the intervening time period.

The foregoing principles of System Dynamics are included in the model building methodology used for the development of the DSN Model. This methodology is described in the following section.

## III. Methodology for Developing the DSN Models

Essentially, the process of building a System Dynamics model involves the same definition, design, construction, and testing/validation phases involved in building any good product. These phases form the generic classes from which the model builder will carry out the following specific activities:

(1) Definition Phase.
   (a) Identification of problem areas within the system.
   (b) Identification of questions that the model is to address.
   (c) Definition of the model boundary.

(2) Design Phase: development of the model feedback structure.

(3) Construction Phase: formulation of a computer program simulating the system structure.

(4) Testing/Validation Phase.
   (a) Performing simulation runs.
   (b) Testing and validation of the model structure.
   (c) Performing policy analysis and implementation.

Each activity is expanded in the following paragraphs.

### A. Identification of Problem Areas

One cannot model a system without understanding its structure and recognizing its problem areas. Thus, the DSN Model builders will first investigate the DSN operational structure and dynamic behavior by talking to managers, and will then identify the problem areas with which managers are concerned. Identification of problem areas within a system

is not only the starting point for constructing a model, but also helps to achieve the purpose of the model: to solve system problems and enhance the effect of future policies.

### B. Identification of Questions

Questions that managers would like answered are a major consideration during model construction. The ability of a model to process management queries and supply viable answers is a function that will drive the model design. Thus, the modeler must identify and focus on managers' questions as central points of the model concept (structure).

### C. Definition of Model Boundaries

One of the preliminary activities involved in building a System Dynamics model is to identify the elements that are significant to the dynamic behavior of a system being modeled. These elements are included within the model boundaries. Thus, the DSN Model builders will define these elements prior to developing the model feedback structure.

### D. Development of Model Feedback Structure

The preceding activities have been devoted to model definition. Having identified the organizational problem areas, obtained managers' questions, and defined the model boundaries, the model builders may then move forward to develop the feedback structure of the model. During this design phase, the modeler together with interested managers will formulate all the interrelationships that exist among the DSN elements. In so doing, the model feedback structure can represent the structure of the system (DSN) to be modeled. In addition, data collection activities should be performed prior to or during this phase.

### E. Formulation of a Computer Program

Once the model feedback structure has been established, the computer program corresponding to the model structure can be developed. In so doing, the modeler will be able to interact with the model via a computer and perform further development and refinement activities.

### F. Performing Simulation Runs

To investigate the behavior produced by a model, a number of simulation runs will be necessary. Simulation runs allow the modeler to exercise the model and see what system behavior it produces.

### G. Testing and Validation of the Model Structure

The DSN Model structure will be drawn from managers' preceptions and intuitions regarding the DSN. Since mistakes can easily be made while linking these pieces of information together, testing and validation of the model structure is a required activity.

Various techniques have been employed in System Dynamics modeling for the testing and validation of a model's structure and behavior. "There is no single test which serves to validate a System Dynamics model. Rather, confidence in a System Dynamics model accumulates gradually as the model passes more tests and as points of correspondence between the model and empirical reality are identified" (Ref. 4). Thus, a combination of tests will be employed for building confidence in the DSN Model. The tests fall into three categories: model structure tests, model behavior tests, and tests of policy implications of the model. The following subsections will discuss the concepts of these tests.

(1) *Tests of Model Structure*: Structure verification, parameter verification, and extreme condition are three widely used tests for examining the correctness of a model structure.

(a) Structure-Verification Test. This test examines whether the model structure coincides with that of the real system under study. One of the techniques that the DSN Model builder may employ to perform this test is to present the model structure to DSN managers. The managers will in turn comment as to the validity of the model structure based upon their perceptions of the real system.

(b) Parameter-Verification Test. When applying this test, the modeler will determine if the parameters incorporated within the model structure actually correspond to those perceived in the real system.

(c) Extreme-Condition Test. "Much knowledge about the real system relates to consequences of extreme conditions" (Ref. 5). For example, assuming that there is no food available for people (i.e., an extreme condition), one would expect the birth rate to approach zero. Structure in a System Dynamics model should, indeed, be able to demonstrate such extreme conditions since the model itself represents a real system. The DSN Model structure must be reexamined if extreme-condition tests are not met.

(2) *Tests of Model Behavior*: The process of comparing model-generated behavior to real-system behavior to evaluate the adequacy of a model structure is called the model behavior test. A series of model behavior tests may be employed to validate the DSN Model. They are: behavior reproduction, behavior prediction, surprise behavior, and behavior sensitivity.

(a) Behavior-Reproduction Tests. The behavior-reproduction tests are used to determine the model's

ability to replicate behavior observed within the real system. Behavior-reproduction tests include: symptom generation, frequency generation, relative phasing, multiple mode, and behavior characteristics. These tests serve the common purpose of examining whether or not a model recreates the symptoms of problems that were observed in the real system.

(b) Behavior-Prediction Tests. "Behavior-prediction tests are analogous to behavior-reproduction tests. Whereas behavior-reproduction tests focus on reproducing historical behavior, behavior-prediction tests focus on future behavior" (Ref. 6). In other words, behavior-prediction tests are used to examine whether a model will generate patterns that are expected to happen in the future.

(c) Surprise-Behavior Test. A more comprehensive model may well display behavior that exists within the system but has not been previously recognized. When such a situation occurs, the modeler must investigate the causes of the unexpected behavior within the model and determine if such characteristics are exhibited within the real system. By thus employing the surprise-behavior test, the DSN Model builder can investigate the model with respect to its full usefulness.

(d) Behavior-Sensitivity Test. This test is used to examine how changed parameter values within the model will affect the model's behavior. One will establish more confidence in the model structure if its generated behavior, resulting from the parameter changes, corresponds to that perceived in the real system.

(3) *Tests of Policy Implications.* Policy implication tests are used when one wishes to check the model's predictive ability with respect to policy analysis. There are two major difficulties in using the results of policy tests. First, when implementing a new policy (found beneficial from exercising the model), a fair amount of time is required to ascertain the results. It is time consuming. Second, implementing a new policy that has been recommended by the model can be risky.

The changed behavior-prediction test (one of the policy tests) can be used to circumvent the difficulties discussed above. Thus, this test (discussed in the following paragraph) will be employed to examine the usefulness of the DSN Model as a policy analysis tool.

"The changed-behavior prediction test asks if a model correctly predicts how behavior of the system will change if a governing policy is changed. The test can be made by examining responses of a model to policies which have been made in the real system to see if the model responds to a policy change as the real system responded" (Ref. 7). Therefore, by employing this test, the risks that may occur when performing the other policy tests can be avoided.

### H. Performing Policy Analysis and Implementation

Once the DSN Model has been built, and confidence in its structure and behavior have been established through various tests, it can then be utilized as a policy analysis tool. The system structure, represented in the model, includes the existing organizational policies. Thus, one can simulate changes in policy by judiciously tweaking the relevant parts of the model structure and observing the effects. New policies may be evaluated and if desirable results are obtained, these new policies may be implemented in the real system. For example, policy analysis may be performed to determine the most effective allocation of limited resources. During this activity, the modeler will demonstrate policy analysis to the manager by exercising and interacting with the model.

### I. Sustaining

Future modifications to the model will be required and are desirable so as to maintain a model structure that corresponds to the dynamic nature of the real system. Therefore, once the DSN Model is completed, some sustaining activity will be necessary.

Figure 1 illustrates and summarizes the methodology explained above that will be used for the development of the DSN Model.

## IV. Progress to Date

A System Dynamics model of the DSN, called DSNMOD, was built in the early seventies. The first step, in the process of developing DM then, was to analyze and describe the existing model. This was done and the following conclusions were drawn. The older model is a thorough representation of the DSN from one person's perspective. The base of understanding built into such a model, however, must be broadened. Secondly, the model addresses itself to the structure of the DSN in the sixties and very early seventies. The DSN structure has been modified since that time, and the newer structure must be represented. Lastly, the model, as a tool for managers, must be widely discussed by current managers and must be directed toward today's planning questions. These conclusions were integrated into the implementation plan for the DSN Model.

The team working on the DM task next interviewed over 20 TDA and DSN managers to ascertain the uses and questions to which DM should be applied. Three products came from these interviews: (1) a list of questions that managers would like to see addressed by the model, (2) a conceptual model of what the DSN strategic planning structure is, and (3) the functional requirements for DM. The functional requirements were reviewed by a board of TDA and DSN managers, plus an expert in the field of System Dynamics.

To further support the development efforts, the task team received training in System Dynamics, analyzed other useful modeling techniques, performed preliminary analysis of data needs and availability, and made presentations to TDA and DSN personnel on the task and the technique. Most recently, a computerized demonstration model of the DSN, using System Dynamics, was developed, coded, run, analyzed, and presented.

## V. Future Work

Questions that concern DSN managers involved in strategic planning seem to fall into the formats given in Section I. However, each manager has a different perspective and a different segment of the DSN with which he/she must be concerned. Thus, the number of potential strategic planning questions that DM could address is quite large. An important step in the development of DM, then, is choosing a few questions for DM to address in its initial form. The remaining questions can be addressed as refinements or extensions to the base model.

The full methodology (discussed in Section III) will be employed to develop the base model for DM. This base model is expected to be available by January 1982. Subsequently, the high level design for all of DM will be written and reviewed, further extensions of the model will be developed, reviewed, and documented, and the complete DM set will be transferred to a sustaining state in the fall of 1983.

## References


1. Howard, J. E., and Gardiner, P. C., "Modeling the Civil Trial Court System," *Proceedings of the 1980 International Conference on Cybernetics and Society*, Cambridge, Mass., October 1980.
2. Lecture Notes on *Corporate and Economic Policy: The System Dynamics Approach*. Massachusetts Institute of Technology 1981 Summer Course in System Dynamics.
3. Forrester, J. W., *Industrial Dynamics*, M.I.T. Press, Cambridge, Massachusetts, 1961, pp. 107-114.
4. Forrester, J. W. and Senge, P. M., *Tests for Building Confidence in System Dynamics Models*, Alfred P. Sloan School of Management, Massachusetts Institute of Technology, December 16, 1978, p. 2.
5. Ibid., p. 8.
6. Ibid., p. 15.
7. Ibid., p. 23.

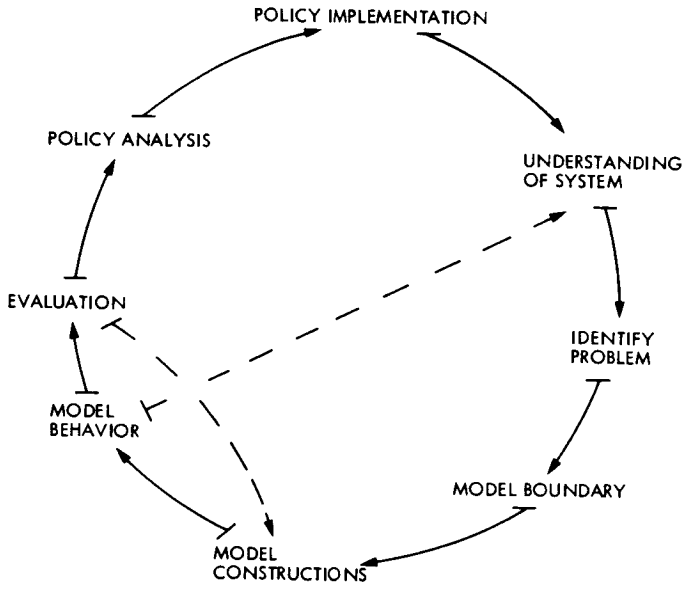



**Fig. 1. Methodology for developing the DSN Models (from Ref. 2)**

N82-11285



# Pioneers 6 Through 12 Mission Support

R. E. Nevarez
Control Center Operations Section

*This article discusses the activities of DSN Operations in support of Pioneers 6, 7, 8, 9, 10, and 11 and Pioneer Venus Orbiter (Pioneer 12) missions during April through June 1981.*

## I. Introduction

This article covers the continuing efforts of the Deep Space Network in support of Pioneer Mission Operations during the cruise phases of the Pioneers 6 through 11 spacecraft and Pioneer Venus orbital operations.

## II. Mission Operations and Status

### A. Pioneers 6 through 9

All spacecraft continue to function nominally. Pioneers 6 and 7 received increased Network support due to solar radial alignment with the Pioneer Venus Orbiter.

### B. Pioneer 10

The spacecraft continues to function nominally with onboard systems in good health. The Network has supported various signal strength tests to determine spacecraft receiver threshold at low to high transmitter powers. The results indicate conical scan can continue to be supported by the Network at low power for a few more years.

### C. Pioneer 11

The spacecraft continues in a state of good health and functioning nominally.

### D. Pioneer 12

The spacecraft continues in good health and functioning nominally.

The Project has programmed periapsis to occur at a fixed time of approximately 0200 (Universal time). The spacecraft went through superior conjunction of a duration of approximately 50 days. The 64-meter-antenna network supported with the high-power transmitter on a daily basis to ensure spacecraft command reception vital to its survivial.

## III. Special Activities

The Deep Space Network Compatibility Test Area (CTA-21) continues to support the Pioneer Project in concurrence with command software development and operations training.

Owing to the distance of Pioneer 11, the 34-meter-antenna network will require installation of 3-Hz loop filters. Steps have been implemented to provide the 34-meter-antenna network with 3-Hz loop filters and ensure continuous network support.

## IV. Summary

All spacecraft appear in good health and functioning nominally as the Deep Space Network continues its vigilant support. Tracking coverage from April through June 1981 is shown in Table 1.

**Table 1. Tracking coverage, April through June 1981**

| Month, 1981 | Pioneer spacecraft | Station type | Number of tracks | Tracking time, hr/min |
|---|---|---|---|---|
| Apr | 6 | 64-m | 2 | 10:15 |
| | 7 | 64-m | 1 | 7:00 |
| | 8 | 64-m | 1 | 7:30 |
| | 9 | 64-m | 2 | 10:55 |
| | 10 | 26-m | 2 | 8:20 |
| | | 34-m | 3 | 13:25 |
| | | 64-m | 53 | 378:50 |
| | 11 | 26-m | 38 | 269:05 |
| | | 34-m | 11 | 92:40 |
| | | 64-m | 16 | 96:45 |
| | 12 | 26-m | 5 | 41:45 |
| | | 34-m | 16 | 140:40 |
| | | 64-m | 68 | 408:55 |
| May | 6 | 64-m | 6 | 34:35 |
| | 7 | 64-m | 11 | 56:10 |
| | 8 | 64-m | 1 | 6:10 |
| | 9 | 64-m | 6 | 36:40 |
| | 10 | 64-m | 54 | 466:40 |
| | 11 | 26-m | 39 | 389:10 |
| | | 34-m | 8 | 80:00 |
| | | 64-m | 20 | 149:35 |
| | 12 | 26-m | 32 | 261:55 |
| | | 34-m | 56 | 494:50 |
| | | 64-m | 16 | 92:35 |
| Jun | 6 | 64-m | 3 | 13:50 |
| | 7 | 64-m | 18 | 81:00 |
| | 8 | 64-m | 1 | 4:25 |
| | 9 | 64-m | 3 | 14:55 |
| | 10 | 34-m | 2 | 6:45 |
| | | 64-m | 69 | 465:55 |
| | 11 | 26-m | 50 | 440:55 |
| | | 34-m | 26 | 212:40 |
| | | 64-m | 16 | 122:00 |
| | 12 | 26-m | 61 | 504:05 |
| | | 34-m | 55 | 444:30 |
| | | 64-m | 17 | 106:55 |

N82-11286



# Strategies for Weather-Dependent Data Acquisition

E. C. Posner
TDA Planning Office
and
Electrical Engineering Department,
California Institute of Technology

*A strategy for data acquisition from a very distant spacecraft is presented, when the system performance can be severely degraded by the Earth's weather due to the high microwave frequency being used. Two cases are considered, one in which there is a certain minimum data rate to be maintained and one in which there isn't. The goal is to maximize expected data return, where we assume that there is always new data, or a backlog of old data, that can be sent if conditions are favorable. When there is no minimum rate to be maintained, the optimum strategy is the greedy strategy, which always transmits at that single rate which maximizes the expected data returned. If there is a minimum data rate that we strive to maintain even in adverse conditions, the optimum strategy transmits simultaneously at the minimum or base data rate and at a bonus data rate. We use a coding system designed for the bandwidth-constrained degraded broadcast channel. The optimum version of this system can, under realistic assumptions, save on the order of 5 dB over the conservative strategy of just transmitting at a single lower data rate.*

## I. Introduction

The Solar System Exploration Program uses X-band (8.5 GHz) to get a narrow antenna beam from spacecraft to Earth and so to maintain a high data rate in clear weather. And someday $K_a$-band (32 GHz) may be used. When reception is limited to receivers buried in the Earth's atmosphere, we must suffer weather effects which reduce the data rate from the value that we can maintain in clear dry weather. How can we best cope with this?

We assume a spacecraft so distant that the data rate cannot be changed to take account of the actual weather at a receiving station. At most, the spacecraft knows a statistical prediction of the weather that will be experienced at reception on Earth, and chooses a data rate or data rates based on this. In this paper, we also assume that there are always data to be sent, either because new data are being acquired or because there is a buffer which stores data until the data can be sent. In the latter case, a feedback channel is involved.

The goal in this paper is to maximize the total expected data returned during a mission or portion of a mission. This is the time integral of the total data rate chosen for transmission times the probability that reception could be carried out. For, we assume Shannon link coding, in which either the signal-to-noise ratio can sustain the particular code chosen, or the word and bit error probability are substantial and so the decoded data must be discarded.

Here we have a risk-balancing problem. If we choose too *high* a data rate, we risk losing all, but if we choose too *low*,

we risk not getting as much out of the mission as we could have. We will consider the possibility of using simultaneous data rates, in an optimum joint coding scheme from the theory of broadcast channels. We shall also point out that mixed strategies, ones in which we randomly choose data rates, should not be used if the criterion is maximizing the expected data return, as it is here.

We may also be required to maintain a certain *minimum* data rate with high probability such as 0.99, for example, spacecraft health data and minimal imaging coverage. We call this the *base* data rate. The additional data that we may be able to receive is called *bonus* data.

Here there are three possibilities. The weather may turn out good enough to support the bonus data rate and the base data rate. Or the weather may not be that good, but may still support the base data rate. Finally, the weather may be so bad that we receive nothing. We still wish to maximize the expected total data returned. But we have the constraint that the probability that we receive nothing (this is the probability that the base data cannot be received) be held to a given small probability. This approach will be contrasted with the more traditional approach of using a single data rate only, but putting in a "weather pad" so that the chosen single data rate can be received with acceptably high probability.

## II. No Minimum Rate Required

Here we define $p(x)$ as the probability density that the weather is good enough to support data rate $x$. If the spacecraft were close enough to Earth, $p(x)$ would be a delta function because we know what the weather is. For a distant spacecraft, $p(x)$ represents the residual uncertainty after our best weather prediction is made. The density $p(x)$ is assumed known.

We will first choose a single rate deterministically depending on $p$. Later we will show that multiple rate strategies should not be considered here. This means no mixed strategies and no strategies employing more than one rate at once. But, if we need to retain a certain minimum rate with a high probability, we shall see in the next section that a dual-rate system should be employed. Here we ask, at what data rate $x_0$ should we choose to transmit so as to maximize the expected data return?

We receive $x_0$ bits per second with probability $P(x_0)$ given by

$$P(x_0) = \int_{x=x_0}^{x_{max}} p(x)\,dx \qquad (1)$$

Here $x_{max}$ is the data rate supportable in clear dry weather. We receive $x_0$ bits/sec with probability $P(x_0)$, and 0 bits/sec with probability $1 - P(x_0)$. This is because if the weather we actually have will not support the data rate that was transmitted, then according to Shannon we get nothing.

What is the expected data rate $E(x_0)$? It is just

$$E(x_0) = x_0 P(x_0) \qquad (2)$$

according to what we just said. So we want to choose an $x_0$ to maximize (2). This is the best we can do. In reasonable cases, we can use differentiation to maximize (2). The condition for a maximum is

$$P(x_0) - x_0 p(x_0) = 0 \qquad (3)$$

For example, let $p(x)$ be given by

$$p(x) = \frac{3}{2x_{max}^3}(x_{max}^2 - (x_{max} - x)^2) \text{ for } 0 \leqslant x \leqslant x_{max},$$

so that

$$P(x_0) = \frac{3}{2}\left(1 - \frac{x_0}{x_{max}}\right) - \frac{1}{2}\left(1 - \frac{x_0}{x_{max}}\right)^3.$$

This is the quadratic density function having its peak at $x_{max}$ and having the value 0 at $x = 0$. Equation (3) becomes

$$\frac{3}{2}\left(1 - \frac{x_0}{x_{max}}\right) - \frac{1}{2}\left(1 - \frac{x_0}{x_{max}}\right)^3 = \frac{3}{2}\frac{x_0}{x_{max}}\left[1 - \left(1 - \frac{x}{x_{max}}\right)^2\right]$$

We let $u = 1 - (x_0/x_{max})$ and solve this equation numerically to find $u \doteq 0.459$. Or,

$$\frac{x_0}{x_{max}} \doteq 0.541$$

The maximum expected data return is then

$$E(x_0) = x_0 P(x_0) = 0.346\, x_{max}$$

Suppose now that we were prescient and knew the actual weather $x$ in advance. We would transmit exactly at the rate $x$. The expected data return $\overline{E}$ with prescience is then

$$\overline{E} = \int_{x=0}^{x_{max}} xp(x)\,dx \qquad (4)$$

This of course must be greater than the value of $E(x_0)$ for the maximizing $x_0$. But, how much greater?

In the example, we can find $\bar{E}$. The result is

$$\bar{E} = (5/8)\,x_{max}$$

The gain in expected data return will be

$$\frac{\bar{E}}{E(x_0)} = \frac{(5/8)\,x_{max}}{0.346\,x_{max}} = 1.806 = 2.6 \text{ dB}$$

This is equivalent to a 2.6-dB power increase, at least if we have an infinite-bandwidth gaussian channel at our disposal. We lose 2.6 dB by not being able to predict the weather perfectly in this example. We lose in two ways. Some of the time we could have gotten rates higher than $x_0$ when we only got $x_0$, and some of the time we could have gotten something (less than $x_0$), instead of getting nothing.

How large can the loss be for not being able to predict the future perfectly? Appendix A shows that we can lose arbitrarily much. This is not surprising. Appendix A also shows that if the density $p(x)$ is nondecreasing in the interval of rates $[0, x_{max}]$, as is sometimes reasonable, then there is a maximum loss. It is 3 dB, attained only for the uniform density on $[0, x_{max}]$. For every other nondecreasing density, the loss will be less than 3 dB. The 2.6-dB loss in the quadratic density example is near the maximum possible for a nondecreasing density function.

Let us close this section by turning to another matter. We want to show that we do not need to consider multiple rates, either simultaneously or as part of a single-rate mixed strategy involving several rates.

It is easy to rule out mixed strategies. For, if we choose rates $x_i$ with probability $p_i$, the expected data return is $E(\{p_i\})$ given by

$$E(\{p_i\}) = \sum_i p_i \cdot x_i P(x_i) \tag{5}$$

We may as well always choose an $x_i$, say $x_i'$, maximizing $x_i \cdot P(x_i)$ over all $i$. This will yield a larger $E(\{p_i\})$ (unless all $x_i \cdot P(x_i)$ with $P_i \neq 0$ are equal). Mixed strategies do not pay.

Why should we not send data simultaneously at more than one rate if there is no minimum data rate required? We might think to do this to salvage something even if the weather turns out worse than we expected. Suppose we send data simultaneously at rates $x_i$, $1 \leqslant i \leqslant n$. We will use as our simultaneous coding scheme the degraded gaussian broadcast channel model of Ref. 1. Here increasing index means worse weather.

We see that all rates $x_j$ with $j \geqslant i$ can be received if weather $i$ holds at time of receipt. We shall let $u_i \geqslant x_i$ be the single-channel rate that could have been supported with weather $i$. Here the data $x_i$ are disjoint, so that we really do get credited for

$$\sum_{j>i} x_j$$

when weather $i$ occurs. If we want to maximize expected data return, there is no reason to repeat data on separate channels.

Actually, the total data returned if weather $i$ occurs cannot exceed the data that would have been returned had we put all our power into a single channel:

$$\sum_{j>i} x_j \leqslant u_i \tag{6}$$

This is because $u_i$ is the channel capacity when weather $i$ occurs. Reference 1 shows that the inequality in (6) is strict unless the bandwidth is infinite.

The probability that the weather is at least as good as $i$ is given by $P(u_i)$ from Eq. (1). The expected data returned when we use the multiple rates is then given by say $E_1(x_1, x_2, \ldots, x_n)$, where

$$E_1(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} x_i P(u_i) \tag{7}$$

This is because the probability that $x_i$ can be received is the probability $P(u_i)$ that the weather is at least as good as $i$.

Because increasing $i$ means worse weather,

$$P(u_i) \leqslant P(u_2) \leqslant \ldots \leqslant P(u_n)$$

So, Eq. (7) can be converted into an inequality:

$$E_1(x_1, x_2, \ldots, x_n) \leqslant P(u_n) \sum_{i=1}^{n} x_i \tag{8}$$

In view of (6) with $i = 1$, (8) becomes

$$E_1(x_1, x_2, \ldots, x_n) \leqslant u_n P(u_n) = E(u_n) \tag{9}$$

The right-hand side of (9) is, by (2), the expected data return if we use the single rate $u_n$. This means we assume the worst weather and use all our power with the single rate $u_n$ tailored for that weather. This shows that single rates are best when there is no minimum data rate to be maintained.

## III. Minimum Rate Required

Suppose there is a minimum or base data rate $x_2$ that we strive to receive even under adverse conditions. This means that for some small $\epsilon > 0$, we want the probability of not being able to receive at least this base data to be $\epsilon$ or less. Typically, $\epsilon$ may be 0.05 or 0.01, but may even be 0. This is because there generally is a worst weather loss. For example, at X-band (8.5 GHz), the loss due to attenuation and the resulting noise temperature increase is at worst 10 dB or so into a roughly 30 K receiving system. We should and do assume that $\epsilon$ and $x_2$ are such that, in weather corresponding to probability $\epsilon$, if we put all our power into a single rate, a rate $x_2$ can in fact be supported. We call $x_2$ the *base* data rate, and the data to be sent at that rate is called *base* data.

As in Section II, we need only consider the following strategies. We choose a single other level, say $\eta$, $1 \geqslant \eta \geqslant \epsilon$, of weather degradation (not precluding that it be the same as $\epsilon$-weather). We transmit at the maximum rate $x_1$ such that in $\eta$-weather, the simultaneous rates $(x_1, x_2)$ can be supported. We call $x_1$ the *bonus* data rate, and the data communicated by it *bonus* data.

Here the problem is to choose $x_1$ (or alternatively $\eta$, for $x_1$ and $\eta$ determine each other when $x_2$ and $\epsilon$ are given) so as to maximize the total expected data return, base and bonus. There are three possibilities. The weather may be worse than $\epsilon$-weather, and we receive nothing. The weather may be at least as good as $\epsilon$-weather, but worse than $\eta$-weather. We then receive only base data. Finally, the weather may be at least as good as $\eta$-weather. We then receive base and bonus data.

Let us use broadcast channel theory from, for example, Ref. 1. We let $B$ be the bandwidth available for signalling. $B$ may be infinite, but the notation assumes that $B$ is finite. Also let $P$ be the total available received signal power. Let $N_1$ be the receiver noise power spectral density corresponding to $\eta$-weather, the weather that determines $x_1$. Let $N_2 \geqslant N_1$ be the noise density corresponding to $\epsilon$-weather. Then from Ref. 1 the allowable simultaneous rates $(x_1, x_2)$ at which we can reliably communicate are given by

$$\begin{aligned} x_1 &= B \log_2 \left(1 + \frac{\alpha P}{N_1 B}\right) \\ x_2 &= B \log_2 \left(1 + \frac{\bar{\alpha} P}{\alpha P + N_2 B}\right) \end{aligned} \tag{10}$$

Here $\alpha$ is a parameter between 0 and 1, and $\bar{\alpha} = 1 - \alpha$. Because $N_2$ is given (it is determined by $\epsilon$), and $x_2$ is a given requirement, $\alpha$ is determined from the second equation of (10). We note that as $B \to \infty$, (10) just becomes the time-division or frequency-division multiplexing formula. We shall not explore this here. Reference 2 considered this case when the weather distribution consisted of a finite number of delta functions.

The meaning of (10) is that we devote a fraction $\alpha$ of the power to the bonus data at rate $x_1$, ignoring the base data. We code for the resulting gaussian channel of signal power $\alpha P$, bandwidth $B$, and noise density $N_1$. The codewords for the base data at rate $x_2$ are assigned the remaining average power, which is $\bar{\alpha} P$. These "base" codewords are centered as a "cloud" around the "bonus" codewords. The bonus codewords look like Gaussian noise of power $\alpha P$ to a receiver designed for $\epsilon$-weather. We cannot decode them, and they add $\alpha P$ to the $N_2 B$ noise power denominator of the second equation of (10). But, if the weather is at least as good as $\eta$-weather, we *can* reliably decode the bonus codewords, remove them as noise by subtraction, and certainly then decode the base codewords. In this way, we see that we can do at least as well as (10), and Ref. 1 shows we can do no better. Our assumption that we can at least get rate $x_2$ in $\epsilon$-weather if we let $x_1$ be 0 means that these equations are not contradictory – no $\log_2$'s of negative numbers.

What is $\alpha$? The second equation of (10) becomes

$$2^{x_2/B} = 1 + \frac{\bar{\alpha} P}{\alpha P + N_2 B}$$

from which we may derive

$$\begin{aligned} \bar{\alpha} &= \left(1 + \frac{N_2 B}{P}\right)\left(1 - 2^{-x_2/B}\right) \\ \alpha &= 1 - \left(1 + \frac{N_2 B}{P}\right)\left(1 - 2^{-x_2/B}\right) \end{aligned} \tag{11}$$

and,

$$\alpha P = P - (P + N_2 B)(1 - 2^{-x_2/B}) \qquad (12)$$

Finally, the supportable bonus rate $x_1$ is given in terms of $N_1$ as

$$x_1 = B \log_2 \left(1 + \frac{P - (P + N_2 B)(1 - 2^{-x_2/B})}{N_1 B}\right) \qquad (13)$$

Equation (13) allows us to translate the probability distribution of the weather, which we may think of as a distribution on $N_1$, into a probability distribution on the attainable bonus data rate $x_1$.

We now observe that the only rate varying is $x_1$, because $x_2$ is fixed. So as far as optimization problems go, this is really the same problem as in the preceding section. We have a probability density $p(x_1)$ on rates $x_1$, and want to maximize the total expected bonus data returned. For, by our assumptions, the expected base data returned will be $(1 - \epsilon) x_2$, no matter what $x_1$ we choose. (This really assumes some continuity in the density function $p$, but we shall say no more about this.)

Let us now determine the maximum penalty due to lack of prescience, as we did in the previous section when no base data was required. We shall again want to assume that the probability density $p(x_1)$ on the achievable bonus data rate $x_1$ is nondecreasing in $x_1$. However, since we really would be given a probability density $r(N_1)$ on the noise density $N_1$ that actually occurs, this condition may be hard to check. Appendix B shows that if $N^2 r(N)$ is nonincreasing in its interval of definition $[N_{min}, N_{max}]$, then the density $p(x_1)$ will indeed be nondecreasing in its interval of definition $[0, x_{max}]$.

Under this mild condition, the previous section shows that the *rate* penalty in $x_1$ for not being prescient is at most 3 dB. What is this as a *power* penalty? This will help tell us how hard we should try to reduce the weather uncertainty.

We are really asking, by how much does the received power $P$ have to increase the double $x_1$? We still must have base rate $x_2$ achieved with probability $1 - \epsilon$. So, Eq. (10) holds, and (13) holds, with $2x_1$ replacing $x_1$ and say $P'$ replacing $P$. Starting at various forms of (13), we quickly realize that under the proper circumstances, doubling $x_1$ can mean an arbitrarily large increase ratio for $P$, or an arbitrarily many dB power penalty for not being prescient.

But, let us assume, as is reasonable, that $x_1/B$ in bits/cycle is not too large before the prescience doubling. This is probably a reasonable implementation constraint for most coding schemes that would be adopted. Let us agree that $x_1$ was equal to $B$ before the doubling. What happens to $P$ now?

Equation (13) can be written

$$2^{x_1/B} - 1 = \frac{P}{N_1 B} \cdot 2^{-x_2/B} - \frac{N_2}{N_1}(1 - 2^{-x_2/B}) \qquad (14)$$

Here, of course, $N_2 > N_1$. If $x_1/B = 1$, so that $2^{x_1/B} - 1 = 1$, we will have $2^{2x_1/B} - 1 = 2^2 - 1 = 3$. From (14), we need to at most *triple* $P$ to double $x_1$, when $x_1/B$ originally equalled 1. There is never more than a 4.8 dB power penalty for lack of prescience when we have a base data rate $x_2$ to maintain with probability $1 - \epsilon$, if the bonus rate $x_1$ does not exceed the available bandwidth $B$.

However, there is a more important question we can ask. What do we gain by the optimal two-rate base and bonus strategy over the more traditional approach of using a single rate? The single rate we should compare to is the one we get if we assume noise density corresponding to $\epsilon$-weather, the noise density $N_2$. For, this is the only way we can guarantee that we receive the base data in a single-rate strategy. The next section considers the gain of the optimal two-rate strategy over the traditional one-rate strategy.

## IV. Tradition Does Not Pay

The traditional single-rate strategy provides a data rate $x_s$ at least equal to $x_2$, because we put all our power into a single channel. Because the noise density $N_2$ is the one for $\epsilon$-weather, $x_s$ will be given by

$$x_s = B \log_2 \left(1 + \frac{P}{N_2 B}\right) \qquad (15)$$

We obtain $x_s$ with probability $1 - \epsilon$, and no data at all with probability $\epsilon$.

The expected total data returned by the traditional strategy will be $E_s$, which is given by

$$E_s = (1 - \epsilon) x_s \qquad (16)$$

With the optimal dual-rate strategy, the total data return $\bar{E}_{opt}$ is given by

$$\bar{E}_{opt} = \max_{x_1} x_1 P(x_1) + (1 - \epsilon) x_2 \qquad (17)$$

Here $P(x_1)$ is the probability that the weather is good enough to support the bonus rate $x_1$ while the second channel is providing the base rate $x_2$.

We may ask, how large can $S = \bar{E}_{opt}/\bar{E}_s$ be? This of course depends on the weather density $r(N_1)$. But, it will always exceed 1 by definition of "optimum."

It is easy to show that $\bar{E}_{opt}/\bar{E}_s$ can be as large as we please if we pick the right parameters. For this, we should have a substantial probability, say around 1/2, that the noise $N_1$ is much smaller than $N_2$. But what if there is a minimum noise $N_{1,min}$? What is the largest that the data rate gain $S = \bar{E}_{opt}/\bar{E}_s$ can be in this case?

We want to choose a probability distribution on the noise to maximize the data rate gain $S$. From (16) and (17), we want to maximize

$$S = \frac{1}{(1 - \epsilon) x_s} \max_{x_1} x_1 P(x_1) + \frac{x_2}{x_s} \qquad (18)$$

by choosing the distribution of receiver noise. Here $\epsilon$ and $x_2$ are given and

$$pr(N_1 > N_2) = \epsilon,$$

or

$$\int_{N_2^+}^{\infty} r(N)\, dN = \epsilon \qquad (19)$$

Finally, $x_s$ is given according to Shannon by (15).

From (13), we find $x_1$ in terms of $N_1$ and $N_2$. The probability that the bonus data can be received is given by

$$P(x_1) = \int_0^{N_1} r(N) dN$$

Since $x_2$ is also given, the only term varying in (18) is $\max_{x_1} x_1 P(x_1)$.

How can we maximize this maximum on $x_1$ by the choice of the probability distribution on the noise density? Here $x_1$ is considered fixed. So we should maximize $P(x_1)$. Since the probability that base data cannot be received is $\epsilon$, $P(x_1)$ can be arbitrarily close to $1 - \epsilon$. We have to say "arbitrarily close" because we need to keep some probability just to the left of $N_2$ to "lock in" the base rate $x_2$.

The *max* of $x_1 P(x_1)$, which is really a sup, is then $(1 - \epsilon) x_1$. By arranging for a near-delta function of probability almost $1 - \epsilon$ at (just to the left of) $N_1$, we can indeed arrange that this $x_1$ be the optimum for the resulting distribution. This distribution, while somewhat artificial, is not too out of line, for the real weather distribution may tend to have a delta function of reasonable probability around the minimum noise, corresponding to clear dry weather. In this case, it also corresponds to $N_1$, because here $N_1 = N_{1,min}$.

We then find the maximum gain ratio $S$ given $N_1$ and $N_2$ from (18) as

$$S_{max}(N_1, N_2) = \frac{x_1 + x_2}{x_s} \qquad (20)$$

Given the base rate $x_2$, the received power $P$, and the bandwidth $B$, we may attempt to maximize (20) over all possible $N_1$ and $N_2$ with $N_1 < N_2$, using (13) for $x_1$ and (15) for $x_s$. This is not very instructive.

However, for infinite bandwidth we can find the maximum in (20). We could do this easily from scratch without going through the finite-bandwidth case, but we shall take the limit as $B \to \infty$ in the finite-bandwidth case. We define a parameter $\rho$ greater than 1 as

$$\rho = \frac{N_2}{N_1} \qquad (21)$$

Here $\rho$ is the ratio of the noise density in $\epsilon$-weather to that in clear dry weather, about 10 dB for low-noise X-band reception. Also let the parameter $\gamma$ less than 1 be defined as

$$\gamma = \frac{x_2}{x_s} \qquad (22)$$

This is the ratio of the base data rate $x_2$ to the data rate that could be supported in $\epsilon$-weather if we put all our power into a single rate $x_s$.

From (13), for infinite bandwidth,

$$x_1 = \frac{P}{N_1 \ln 2} - x_2 \frac{N_2}{N_1} \tag{23}$$

From (15), we similarly have

$$x_s = \frac{P}{N_2 \ln 2} \tag{24}$$

Equation (20) then becomes

$$S_{max}(N_1, N_2) = \frac{\left(\frac{P}{N_1 \ln 2} - x_2\left(\frac{N_2}{N_1} - 1\right)\right)}{P/N_2 \ln 2}$$

$$S_{max}(N_1, N_2) = \rho - \gamma(\rho - 1) \tag{25}$$

We can check (25) by noting that

$$\text{as } \gamma \to 1, S_{max}(N_1, N_2) \to 1.$$

This means that there can be no savings in using the optimal strategy if the link can barely support the base rate in $\epsilon$-weather. This is as must be. But if, for example, $\gamma$ were 1/2, corresponding to $\epsilon$-weather being able to support *twice* the base rate,

$$S_{max}(N_1, N_2) = \rho - \frac{1}{2}(\rho - 1) = \frac{\rho + 1}{2}$$

If $\rho = 5$ for the given $\epsilon$, as may be typical of X-band for deep-space use, we can save up to a factor of 6/2 = 3 or 4.8 dB if the $\epsilon$-weather link could have supported twice the base rate. Here the probability that the noise is at least 5 times as bad as for clear dry weather is $\epsilon$. Half the power is devoted to base data, and half to bonus. The bonus rate is 5 times the base rate. There is five times as much energy devoted to each base bit as to each bonus bit.

In the next section, we do a more realistic case, one in which the bandwidth is finite and in which the probability distribution approximates the real ones that seem to occur for X-band weather.

## V. A Realistic Case

In this section, we adopt a weather model much like the actual weather statistics for deep-space communication at X-band. However, the results should be used only as guidelines and not for mission planning, because these are not actual weather statistics.

We consider a probability distribution varying from $N_{min} = 1$ as a normalization to $N_{max} = 10$, corresponding to the approximate 10 dB maximum X-band weather loss. We shall allow a $\delta$-function at $N = 1$, corresponding to a positive probability that the weather is perfectly clear and dry. Let the $\delta$-function have probability $1 - \beta$, so that the area of the continuous part $r_1(N)$ is $\beta$. We shall find $\beta$ to qualitatively match X-band weather statistics in this example.

If we agree that a 3-dB loss ($N = 2$) occurs 5% of the time, we can find $\beta$ under the assumption that the continuous part of the distribution is part of a parabola with its minimum at $N = 10$. Let the continuous part be $c(10 - N)^2$. Then

$$c \int_{N=1}^{10} (10 - N)^2 \, dN = \beta$$

and so

$$c = \frac{\beta}{243}$$

Then we have for $r_1(N)$

$$r_1(N) = \frac{\beta}{243}(10 - N)^2$$

The 3-dB loss occurs with probability 0.05. Or

$$\frac{\beta}{243} \int_{N=2}^{10} (10 - N)^2 \, dN = 0.05 = \frac{\beta}{729}(0.8)^3,$$

and

$$\beta = 0.0712, 1 - \beta = 0.9288$$

The overall weather density becomes

$$r(N) = 0.9288\,\delta(N - 1) + 2.930 \times 10^{-4}(10 - N)^2 \tag{26}$$

So the probability of clear weather is about 93%. This is not contrary to experience.

We have yet to choose an $\epsilon$. Rather than setting the criterion directly on $\epsilon$, we let it correspond to a 7-dB (factor of 5) loss, whereby $N_2 = 5$. Or

$$\epsilon = \int_5^{10} r_1(N)\,dN = 2.930 \times 10^{-4} \times \frac{5^3}{3} = 0.0122$$

This is slightly more than 1%. We want a 98.8% probability of getting at least the base data. One percent is lower than what is sometimes stated as a requirement in present deep space designs, but is probably close to the real requirement. The only reason 1% isn't required now is that the data rate penalty over clear weather would be too large (7 dB) with the traditional strategy. We shall see that we can pick up almost 5 dB of the 7 dB with the optimal dual-rate strategy.

We shall assume that the base-data-rate-to-bandwidth ratio $x_2/B$ is 1/8, and that, if we put all the power into a single channel at the $\epsilon$-weather noise density $N_2$ (=5), we could support a data rate of $2x_2 = x_s$ instead of the $x_2$ we will actually get. Or, from the channel capacity formula (15),

$$2x_2 = B \log_2 \left(1 + \frac{P}{N_2 B}\right)$$

$$2^{2x_2/B} - 1 = \frac{P}{N_2 B}, \quad 2^{1/4} - 1 = \frac{P}{N_2 B} \qquad (27)$$

So, $P/N_2B = 0.1892$, and, since $N_2 = 5$, $P/B = 0.9460$. We have determined the power-to-bandwidth ratio $P/B$. If we had kept better track of units, the dimensions would be joules/cycle (or watts/Hz).

From (11), we can now find

$$\bar{\alpha} = \left(1 + \frac{1}{0.1892}\right)(1 - 2^{-1/8}) = 0.522, \alpha = 0.478$$

We put 52.2% of the power into the base ($x_2$) channel, and 47.8% into the bonus ($x_1$) channel. But, we still need to find $x_1$ and $N_1$. From (10), with $P/B = 0.9460$, we have

$$x_1 = B \log_2 \left(1 + \frac{(0.478)(0.9460)}{N_1}\right)$$

$$\frac{x_1}{B} = \log_2 \left(1 + \frac{0.4522}{N_1}\right) \qquad (28)$$

As an aside, we remark that, within $\bar{\alpha} = 0.522$, the power to the base channel is reduced by 0.522. And, from (10), the noise density is increased, due to bonus coding, by

$$\left(\frac{\alpha P + N_2 B}{B}\right) - N_2 = \alpha P/B = 0.4522,$$

the same 0.4522 of Eq. (28). This is a percentage increase of $0.4522/N_2 = 0.4522/5 = 0.905^- = 9.0\%$. The noise density in the base channel increases 9.0% due to the bonus codewords being seen as noise by the base decoder.

The noise density increase fraction is, from (10), (11), and (27), actually

$$\frac{\alpha P}{N_2 B} = \left(1 + \frac{P}{N_2 B}\right) 2^{-x_2/B} - 1 = 2^{x_2/B} - 1$$

It depends only on the assumed $x_2/B$ and $x_s/B$, and not on the probability density $r(N)$ chosen nor on $\epsilon$. The power devoted to the base channel drops by a factor of 0.522 due to bonus coding, and the noise increases by a factor of 1.0905. So, the signal-to-noise drops by a factor of $0.522/1.09 = 0.479^-$. This is, of course, the drop in signal-to-noise ratio which exactly cuts a data-rate-to-bandwidth ratio of $2x_2/B = 1/4$ to one of $x_2/B = 1/8$:

$$\frac{2^{1/8} - 1}{2^{1/4} - 1} = 0.478^+$$

The probability that the bonus rate can be supported if we design it for weather producing noise density $N_1$ is say $P_0(N_1)$, where

$$P_0(N_1) = 1 - \int_{N_1}^{10} r_1(N)\,dN = 1 - \frac{(10 - N_1)^3}{729}$$

We seek to maximize, by choice of $N_1$, the function $x_1 P_0(N_1)$, or, by (28), maximize

$$G(N_1) = (\ln 2) F(N_1) = \left[\ln \left(1 + \frac{0.4522}{N_1}\right)\right]\left(1 - \frac{(10 - N_1)^3}{10240}\right)$$

The derivative $G'(N_1)$ is

$$G'(N_1) = \frac{1}{\left(1 + \frac{0.4522}{N_1}\right)} \left(\frac{-0.4522}{N_1^2}\right) \left(1 - \frac{(10 - N_1)^3}{10240}\right)$$

$$+ \left[\ln\left(1 + \frac{0.4522}{N_1}\right)\right](-3)\ \frac{(10 - N_1)^2}{10240}$$

Both terms are clearly negative in $1 \leqslant N_1 < 10$, so the maximum is at $N_1 = 1$. We play for clear weather. Because the weather is clear with probability 0.9288, this is no surprise.

From (28), we see that

$$\frac{x_1}{B} = \log_2 (1 + 0.4522) = 0.5388,$$

$$\frac{x_1}{B} P_0(N_1) = (0.5388)(0.9288) = 0.5004$$

From (17), then,

$$\frac{\overline{E}_{opt}}{B} = 0.5004 + (1 - \epsilon)(1/8) = 0.5004 + \frac{0.9878}{8},$$

$$\frac{\overline{E}_{opt}}{B} = 0.6239.$$

Compared with the $2x_2/B$ of 1/4, we have a gain of a factor of 2.50 or 4.0 dB in total expected data rate, and a factor of 2.86 or 4.6 dB in power.

We note that the adopted optimum strategy gives a data-rate-to-bandwidth total of $0.5388 + 0.1250 = 0.6638$ with probability $1 - \beta = 0.9288$, a data-rate-to-bandwidth total of 0.1250 with probability $0.9878 - 0.9288 = 0.0590$, and no data with probability 0.0122. The traditional strategy would give 0.2500 with probability 0.9878 and no data with probability 0.0122. We only get half as much data in the $\epsilon = 0.0122$ weather, but the enormous gain in $N_1 = 1$ (clear dry) weather much more than makes up for this on the average. The optimal dual-rate strategy gains a 2.5 factor increase in expected data rate, which is equivalent, when the finite bandwidth is taken into account, to a power gain of 4.6 dB. Tradition does not pay.

## References


1. Cover, Thomas M., and Abbas El Gamal, "Multiple User Information Theory," *Proc. IEEE*, Vol. 68, 1980, pp. 1466-1483, esp. Sec. II C.
2. Butman, Stanley "Transmission Strategies for Atmospheric Entry Probes," *Proc. 1972 National Telemetering Conf.* (NTC 72), IEEE Publication, pp. 28E-1 to 28E-4.

# Appendix A

# Penalty for Lack of Prescience

Here we answer the question raised in Section II. What is the maximum loss in expected data return from not being able to know what the weather will be?

If there is no restriction on $p(x)$, the loss can be arbitrarily large. For example, let only integer rates $j = 2, 3, \ldots$ be supportable. The example can be modified so that the probability distribution of rates has a continuous density and a maximum rate. The conclusion will be the same.

Suppose the probability $p_j$ that the actual weather would support rate $j$ is

$$p_j = \frac{c}{j^2}, \qquad j = 2, 3, \ldots$$

Here $c$ is an irrelevant constant (actually $1/[(\pi^2/6) - 1]$. For this weather distribution

$$\bar{E} = \sum_{j=1}^{\infty} j p_j = \sum_{j=2}^{\infty} \frac{c}{j} = \infty$$

But the expected data return $E_i$ if we chose rate $i$ is

$$i \sum_{j=i}^{\infty} p_j = ci \sum_{j=i}^{\infty} \frac{1}{j^2} < ci \int_{x=i}^{\infty} \frac{1}{(x-1)^2}\, dx$$

$$E_i < \frac{ci}{i-1} \leqslant 2c$$

So the $i$ maximizing $E_i$, which we may call $i_0$, has

$$E(i_0) < 2c$$

But the expected data return $\bar{E}$ with prescience is $\bar{E} = \infty$. There is an infinite penalty for lack of prescience. This is as expected.

Now we ask, suppose $p(x)$ must be nondecreasing on the interval $[0, x_{max}]$. What is the largest penalty for lack of prescience now? We may scale rates so that $x_{max} = 1$. If $p(x) = 1$ on $[0, 1]$, then

$$\bar{E} = \int_0^1 x\, dx = 1/2$$

Also,

$$x_0 P(x_0) = x_0(1 - x_0)$$

has maximum $1/4$ at $x_0 = 1/2$. For this $x_0$, $E(1/2) = 1/4$ and the penalty is

$$\bar{E}/E(1/2) = \left(\frac{1/2}{1/4}\right) = 2 = 3 \text{ dB}$$

We shall show that the 3-dB loss encountered above is the worse case when $p(x)$ is nondecreasing. In fact, we shall show that if $\bar{X}$ is the mean or expected data rate given by

$$\bar{X} = \int_0^1 x p(x)\, dx$$

(which we have called $\bar{E}$ above) then using $\bar{X}$ as the actually chosen rate $x_0$ never loses more than 3 dB, and loses less unless $p(x)$ is the uniform density on $[x_{min}, x_{max}]$. We note that if $x_{min} > 0$, $\bar{X}$ will not be the optimum data rate to choose.

We seek to show

$$\frac{\bar{X}}{\left(\bar{X} \int_{\bar{X}}^1 p(x)\, dx\right)} \leqslant 2$$

when $p(x)$ is nondecreasing on $[0, 1]$, with equality only when $p(x)$ is constant on its interval of support. We want to show

$$\int_{\bar{X}}^1 p(x)\, dx \geqslant \frac{1}{2} \qquad \text{(A-1)}$$

for $p(x)$ nondecreasing. Another way to put this is that the median $X_{1/2}$, the point such that half the probability lies to the right and half to the left, satisfies

$$\bar{X} \leqslant X_{1/2} \qquad \text{(A-2)}$$

if $p(x)$ is nondecreasing, with equality only when $p(x)$ is constant for $x_{min} \leqslant x \leqslant x_{max}$.

Let $p(X_{1/2}) = h$. We *increase* $\bar{X}$ for the same $X_{1/2}$ if we replace the original $p(x)$ by one which is equal to the constant

$h$ for $b \leqslant x \leqslant X_{1/2}$ and equals 0 for $0 \leqslant x \leqslant b$. Here the constant $b$ is given by

$$h(X_{1/2} - b) = 1/2$$

$$b = X_{1/2} - \frac{1}{2h}$$

Note that $b \geqslant 0$. For the *maximum* probability to the left of $X_{1/2}$ is the area of the rectangle with base $X_{1/2}$ and height $h$. This probability is exactly 1/2, so

$$h \cdot X_{1/2} \geqslant \frac{1}{2}$$

$$b = X_{1/2} - \frac{1}{2h} \geqslant 0$$

What happens to the right of $\overline{X}$? We can increase $\overline{X}$ keeping $X_{1/2}$ the same by extending $p(x)$ to the right at the constant value $h$ for a length long enough, to $a$, say, with $a > 1$, to make the resulting rectangle have area 1/2 to the right of $X_{1/2}$. Figure A-1 explains these operations.

The resulting uniform density on $[b, a]$ has a *larger* mean than the original $p(x)$, for probability has been pushed to the right. However, it has the *same* median $X_{1/2}$. But the resulting uniform distribution has mean $\overline{X}'$ equal to $X_{1/2}$, for the mean of a uniform distribution is also its median. The result is the following:

$$\overline{X} \leqslant \overline{X}' = X_{1/2} \tag{A-3}$$

This provides (A-2). Note that if the original $p(x)$ were not uniform on its interval of support, we would have actually increased $\overline{X}$ to $\overline{X}'$, so that $\overline{X} < \overline{X}'$. This shows that the 3-dB loss holds *only* for uniform $p(x)$. So 3 dB is the worst loss possible for a nondecreasing $p(x)$. This loss is attained only for uniform $p(x)$, and we can guarantee that we don't lose more than 3 dB from the prescience value by choosing to transmit at the rate $\overline{X}$, the mean data rate we could get with prescience. Of course, if we choose the $x_0$ maximizing (2), we will be able to cut the prescience loss still further in all cases except where $x_{min} = 0$ and $p(x)$ is uniform. But for the uniform distribution when the prescience loss is 3 dB, $x_0 = x_{max}/2$ is the best $x_0$.

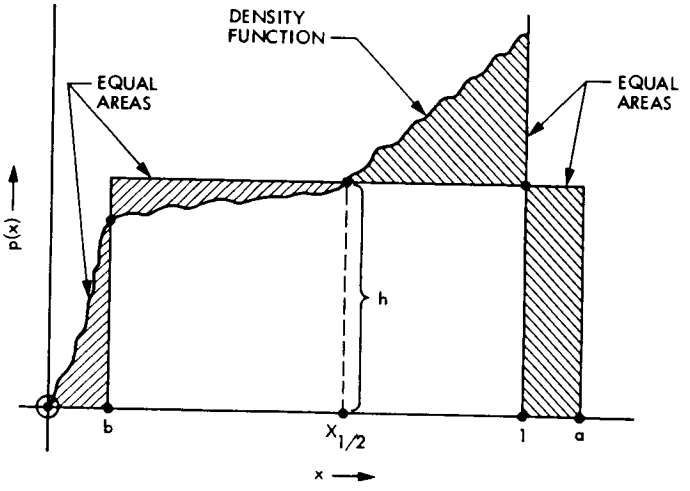



**Fig. A-1. Flattening $p(x)$**

# Appendix B

## Condition for Nondecreasing $p(x_1)$

Here we find a condition on the probability density $r(N_1)$ of the noise $N_1$ that implies that the probability density $p(x_1)$ of the bonus data rate $x_1$ is nondecreasing up to the maximum rate $x_{max}$. We referred to this condition in Section III.

Equation (13) can be written

$$x_1 = B \log_2 \left(1 + \frac{L}{N_1 B}\right) \tag{B-1}$$

Here

$$L = P - (P + N_2 B)(1 - 2^{-x_2/B})$$

is a positive constant, positive because the rate $x_1$ is a positive rate.

We may write

$$pr(X_1 \leqslant x_1) = pr(n_1 \geqslant N_1) = R(N_1) \tag{B-2}$$

Here $X_1$ is the random variable of rates $x_1$, $n_1$ is the random variable of corresponding noise densities $N_1$, and $R(N_1)$ is 1 minus the cumulative distribution of the noise random variable $n_1$ evaluated at the particular $\epsilon$-weather noise $N_2$.

We can differentiate (B-2) with respect to $x_1$, and use (B-1) to differentiate $N_1$ with respect to $x_1$. The result is

$$p(x_1) = \frac{d}{dx_1} R(N_1) = \frac{dN_1}{dx_1} \frac{d}{dN_1} R(N_1) = - \frac{dN_1}{dx_1} r(N_1) \tag{B-3}$$

But, from (B-1),

$$\frac{dN_1}{dx_1} = \frac{1}{dx_1/dN_1} = - \frac{\ln 2}{L} \left(N_1^2 + \frac{L}{B} N_1\right) \tag{B-4}$$

(B-3) becomes

$$p(x_1) = \frac{\ln 2}{L} \left(N_1^2 + \frac{L}{B} N_1\right) r(N_1) \tag{B-5}$$

If $N_1^2\, r(N_1)$ is nonincreasing in $N_1$, then $N_1\, r(N_1)$ is all the more nonincreasing in $N_1$. Since $L/B$ is nonnegative, the sum $N_1^2\, r(N_1) + (L/B) N_1\, r(N_1)$ is nonincreasing in $N_1$. Since increasing noise $N_1$ corresponds to a decreasing bonus rate $x_1$, (B-5) shows that $p(x_1)$ is nondecreasing in $x_1$ if $N_1^2\, r(N_1)$ is nonincreasing in $N_1$. (In fact, the condition on $p(x_1)$ is almost but not quite *equivalent* to the condition on $r(N_1)$.)

N82-11287



# Virtual Center Arraying

L. J. Deutsch, R. G. Lipes, and R. L. Miller
Communication Systems Research Section

*One way to increase the amount of data that can be received from outer planet missions is to array several ground antennas in such a way as to increase the total effective aperture of the receiving system. One such method is virtual center arraying (VCA). In VCA, a combined carrier reference is derived at a point that is, conceptually, the geometric center of the array. This point need not coincide with any of the actual antennas of the array. This report includes a noise analysis of the VCA system and exhibits formulas for the phase jitter as a function of loop bandwidths and the amount of loop damping. If the ratio of the loop bandwidths of the center loop to the vertex loops is greater than 100, then the jitter is very nearly equal to that expected for ideal combined carrier referencing.*

## I. Introduction

Many different antenna arraying systems have been proposed for increasing the data transmission rates from outer planet missions. Two of these schemes, "baseband only combining" and "baseband combining with combined carrier referencing" (or simply "combined carrier referencing"), have been studied in Ref. 1. In ideal combined carrier referencing, the loop signal-to-noise ratio (loop SNR) of the array is equal to that of a single aperture that is the sum of the effective apertures of the individual array elements. In practice, however, it is not possible to achieve ideal combined carrier referencing. Two conceptually different systems have been proposed to achieve performance approaching that of ideal combined carrier referencing.

The approach proposed in Ref. 2 will be referred to as the "master-slave" system. In the master-slave system (see Fig. 1), the carrier power from each element of the array is combined at one of the antenna receivers, called the "master." This receiver derives a combined carrier reference that is used to carrier-aid the other receivers, called "slaves." A short loop in each slave receiver is used to track the frequency and phase differences between the master and local signal. A major benefit of carrier aiding is that the bandwidths of these short loops may be made narrower than would otherwise be possible, since the master receiver provides a good estimate of doppler-induced phase drifts. This bandwidth narrowing produces higher loop SNR's in the slave receivers. The performance of the master-slave system has been studied extensively (Refs. 2, 3, 4).

The approach proposed in Ref. 5 and recently rediscovered by J. W. Layland will be referred to as "virtual center arraying" (VCA). In VCA (see Fig. 2), a combined carrier reference is derived at a point called the center. This point need not coincide with any of the actual array elements. The combined carrier reference is used to carrier-aid the individual receivers, called "vertices." Short loops in the vertices track the frequency and phase differences in the carrier reference. The

vertex short loops, like the slave loops from the master-slave system, may be made narrow because of the carrier aiding. The master-slave system requires one less short loop than VCA. Also, both systems perform like ideal combined carrier referencing in the limit as the short loop bandwidths approach zero.

The VCA system has received less attention than the master-slave system. While a study of the stability of Layland's version of VCA appears in Ref. 6, the performance of VCA has not been previously determined. In Section II, this performance is determined by developing an expression for phase jitter as a function of loop bandwidths and the amount of loop damping. In Section III, loop jitter is given explicitly for expected Voyager 2 Uranus and Neptune encounter conditions for an array of three 34-m antennas with and without a 64-m antenna. These plots demonstrate that when the ratio of the loop bandwidth of the center loop to those in the vertex short loops is greater than 100, the jitter is very nearly equal to that expected for ideal combined carrier referencing.

## II. Jitter Performance of VCA

All signal names in this section correspond to the labeling of Fig. 2. Consider a set of input signals of the form

$$S_k(t) = \sqrt{2}A_k \sin[\omega_0 t + \theta_k(t)] + n_k(t)$$

$$(k = 1, 2, 3, \ldots, N)$$

where $A_k$ is the carrier amplitude, $\omega_0$ is an intermediate frequency, and $\theta_k$ is the phase in the $k^{th}$ vertex. The data part of the input signal is assumed to have been eliminated by filtering. The noises $n_k$ are assumed to be independent white Gaussian processes. $C(t)$ is of the form

$$C(t) = \sqrt{2}K_0 \cos[\omega_1 t + \hat{\theta}_0(t)]$$

where $K_0$ is the center VCO rms output and $\hat{\theta}_0$ is the estimate of the carrier phase in the center. After the first mixer and effective low-pass filtering, the resulting signal is

$$L_k(t) = K_0 A_k \sin([\omega_0 - \omega_1]t + \theta_k(t) - \hat{\theta}_0(t)) + C(t)n_k(t) .$$

The function $W_k$ is defined by

$$W_k(t) = 2K \cos[(\omega_0 - \omega_1)t + \hat{\theta}_k(t)]$$

where $K$ is the VCO rms output in the vertex loops (assumed to be the same for all the short loops) and $\hat{\theta}_k$ is the phase estimate in the $k^{th}$ vertex short loop. After the second mixer (and effective low-pass filtering),

$$X_k(t) = KK_0A_k \sin\phi_k(t) + N_k(t)$$

where

$$\phi_k(t) = \theta_k(t) - \hat{\theta}_k(t) - \hat{\theta}_0(t)$$

and $N_k(t)$ $(k = 1, 2, 3, \ldots, N)$ are independent narrowband Gaussian processes. Each $N_k(t)$ is assumed to have one-sided spectral density $N_0/2$ in the bandwidth of the short loops.

If $f_k(t)$ is the impulse response of the vertex filter $F_k(s)$, then

$$Y_k(t) = \int_{-\infty}^{\infty} X_k(t-u) f_k(u)\, du$$

and, if each vertex has VCO gain equal to $K_{VCO}$,

$$\frac{d\hat{\theta}_k}{dt} = K_{VCO} Y_k(t)$$

$$\approx K_{VCO}KK_0 \int_{-\infty}^{\infty} [A_k\phi_k(t-u) + N_k(t-u)]\, f_k(u)\, du$$

where the last expression holds for small phase errors $\phi_k$. After taking Laplace transforms and solving for $\hat{\theta}_k$,

$$\hat{\theta}_k = H_k(s)\left(\theta_k - \theta_0 + \frac{N_k}{A_k}\right)$$

where

$$H_k(s) = \frac{K_T A_k F_k(s)}{s + K_T A_k F_k(s)}$$

is the closed loop transfer function for the $k^{th}$ vertex short loop and

$$K_T = K_{VCO}KK_0 .$$

The summing junction combines the $X_k$'s with coefficients $\beta_k$. These coefficients are normalized so that

$$\sum_{k=1}^{N} \beta_k^2 = 1 .$$

It is known (see Ref. 1) that the optimal selection for $\beta_k$ is

$$\beta_k = \frac{A_k}{A_T}$$

where

$$A_T^2 = \sum_{j=1}^{N} A_j^2 .$$

The output of the summing junction is

$$Z(t) = \sum_{j=1}^{N} \beta_j X_j(t)$$

whence

$$\frac{d\hat{\theta}_0}{dt} = K_{VCO_0} Y_0(t)$$

$$= K_{VCO_0} \int_{-\infty}^{\infty} Z(t-u) f_0(u)\, du$$

where $f_0(u)$ is the impulse response of the center filter $F_0(s)$. If

$$K_{T_0} = K_{VCO_0} K K_0$$

then, for small phase error $\phi_j$,

$$\frac{d\hat{\theta}_0}{dt} \approx K_{T_0} \sum_{j=1}^{N} \beta_j \int_{-\infty}^{\infty} e(u)\, \{A_j \phi_j(t-u) + N_j(t-u)\} du .$$

It may then be shown that if $H_j = H_1$ for $j = 2, 3, 4, \ldots, N$ (i.e., if all the short closed loop bandwidths are equal), then

$$\hat{\theta}_0 = \frac{1}{A_T} \frac{H_0(s)[1 - H_0(s)]}{1 - H_0(s) H_1(s)} \sum_{j=1}^{N} \beta_k A_k \left(\theta_j + \frac{N_j}{A_j}\right)$$

where

$$H_0(s) = \frac{K_{T0} A_T F_0(s)}{s + K_{T0} A_T F_0(s)} .$$

If each $\theta_j$ is assumed to be slowly varying so that $E(\theta_j) \approx \theta_j$ then

$$\phi_k - E(\phi_k) = [1 - H_1(s)]\, [E(\hat{\theta}_0) - \hat{\theta}_0] - H_1(s) \frac{N_k}{A_k}$$

$$= - \frac{1}{A_T} \frac{H_0(s)\,[1 - H_1(s)]^2}{1 - H_0(s) H_1(s)} \sum_{j=1}^{N} \beta_k N_k - H_1(s) \frac{N_k}{A_k}$$

and so the jitter in the $k^{th}$ vertex is given by

$$\sigma_{\phi_k}^2 = \frac{N_0 (1 - \beta_k^2)}{2 A_T^2} \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} |a_1(s)|^2 ds$$

$$+ \frac{N_0}{2 A_T^2} \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \left| \beta_k a_1(s) + \frac{1}{\beta_k} H_1(s) \right|^2 ds$$

where

$$a_1(s) = \frac{H_0(s)[1 - H_1(s)]^2}{1 - H_0(s) H_1(s)} = a_2(s) - H_1(s)$$

and

$$a_2(s) = \frac{H_0(s)[1 - H_1(s)] + H_1(s)[1 - H_0(s)]}{1 - H_0(s) H_1(s)} .$$

This may be simplified as follows:

$$\sigma_{\phi_k}^2 = \frac{N_0}{2 A_T^2} \left[ \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} |a_2(s)|^2 ds \right.$$

$$\left. + \left( \frac{1 - \beta_k^2}{\beta_k^2} \right) \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} |H_1(s)|^2 ds \right]$$

If all the filters are taken to be of the second-order having the same damping factor $r$, then, using integration techniques described in Ref. 7,

$$\sigma_{\phi_k}^2 = \frac{N_0 B_L}{A_T^2} \left[ \frac{R^2 + \frac{2r}{1+r} R + 1}{R^2 + R} + \left( \frac{1 - \beta_k^2}{\beta_k^2} \right) \frac{1}{R} \right]$$

where $B_L$ is the bandwidth in the center loop and $R$ is the ratio of $B_L$ to the bandwidths in the vertex short loops.

## III. Conclusions

It was shown in Ref. 1 that the loop SNR for ideal combined carrier referencing in the presence of bandpass hardlimiting is given by

$$\rho_{CCR} = \frac{A_T^2}{N_0 B_{L1} \Gamma}$$

where $B_{L1}$ is $B_L$ expanded by the effects of the bandpass hardlimiter and $\Gamma$ is a suppression factor associated with that limiter. Also, the loop jitter for ideal combined carrier referencing is approximately

$$\sigma_{CCR}^2 = \frac{1}{\rho_{CCR}} = \frac{N_0 B_{L1} \Gamma}{A_T^2} \quad .$$

Thus, if bandpass hardlimiting is added to VCA before the first mixers, then the expression for the jitter in the $k^{th}$ vertex becomes

$$\sigma_{\phi_k}^2 = \sigma_{CCR}^2 \left[ \frac{R^2 + \frac{2r}{1+r} R + 1}{R^2 + R} + \left( \frac{1 - \beta_k^2}{\beta_k^2} \right) \frac{1}{R} \right] \quad .$$

The loop SNR of VCA in the $k^{th}$ vertex is then given by

$$\rho_k = \rho_{CCR} \left[ \frac{R^2 + \frac{2r}{1+r} R + 1}{R^2 + R} + \left( \frac{1 - \beta_k^2}{\beta_k^2} \right) \frac{1}{R} \right]^{-1} \quad .$$

Notice that

$$\lim_{R \to \infty} \rho_k = \rho_{CCR} \, .$$

Graphs of $\rho_k$ plotted as a function of $R$ are exhibited in Fig. 3. Arrays consisting of three 34-m antennas with and without a 64-m antenna are considered under typical Voyager 2 Uranus and Neptune encounter conditions.

## References


1. Deutsch, L. J., Miller, R. L., and Butman, S. A., "New Results on Antenna Arraying: Part 1," *TDA Progress Report 42-62, January and February 1981*, Jet Propulsion Laboratory, Pasadena, Calif.
2. Brockman, M. H., "Radio Frequency Carrier Arraying for High-Rate Telemetry Reception," *DSN Progress Report 42-45, March and April 1978*, Jet Propulsion Laboratory, Pasadena, Calif.
3. Brockman, M. H., "Radio Frequency Carrier Arraying for Near Maximum Carrier Signal-to-Noise Ratio Improvement," *DSN Progress Report 42-49, November and December 1978*, Jet Propulsion Laboratory, Pasadena, Calif.
4. Divsalar, D., and Yuen, J., "Carrier Jitter and Radio Loss Analysis of Carrier Arraying," *TDA Progress Report 42-65* (this issue).
5. Schrader, J. H., "Receiving System Design for the Arraying of Independently Steerable Antennas," *IEEE Trans. Space Electronics and Telemetry*, June 1972.
6. Boussalis, H., "Modeling and Stability Analysis of a Multi-Loop Antenna System," *Proceedings of the Twelfth Modeling and Simulation Conference*, Pittsburgh, Penn., 1981 (to be published).
7. Winklestein, R., "Closed Form Evaluation of Symmetric Two-Sided Complex Integrals," *TDA Progress Report 42-65* (this issue).

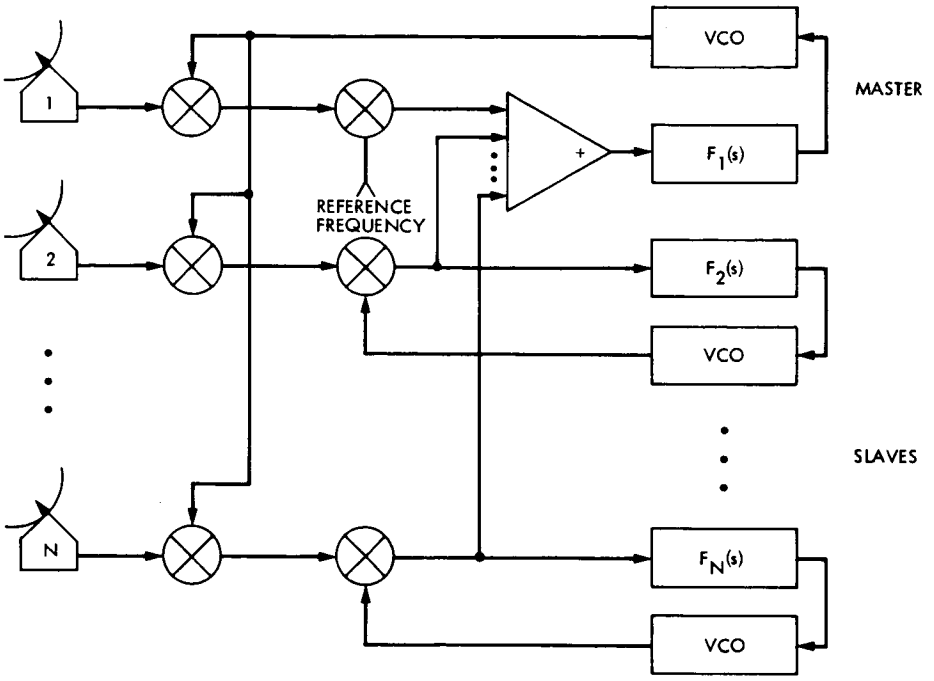



**Fig. 1. Block diagram of master/slave scheme**

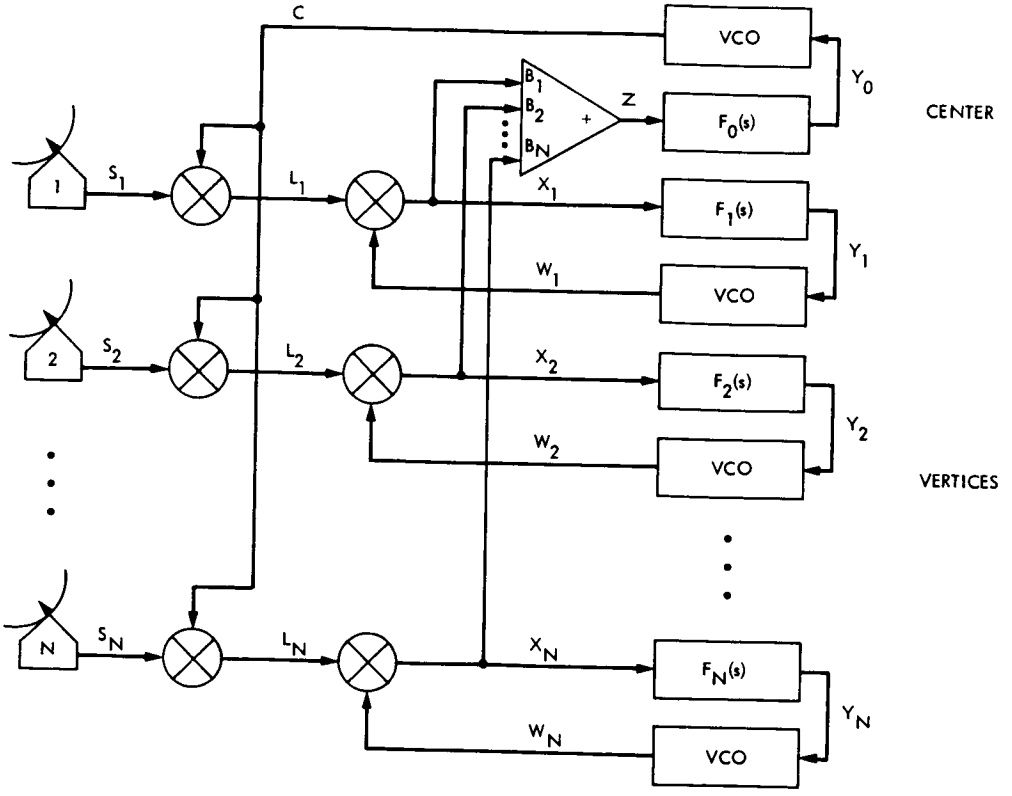



**Fig. 2. Block diagram of virtual center arraying scheme**

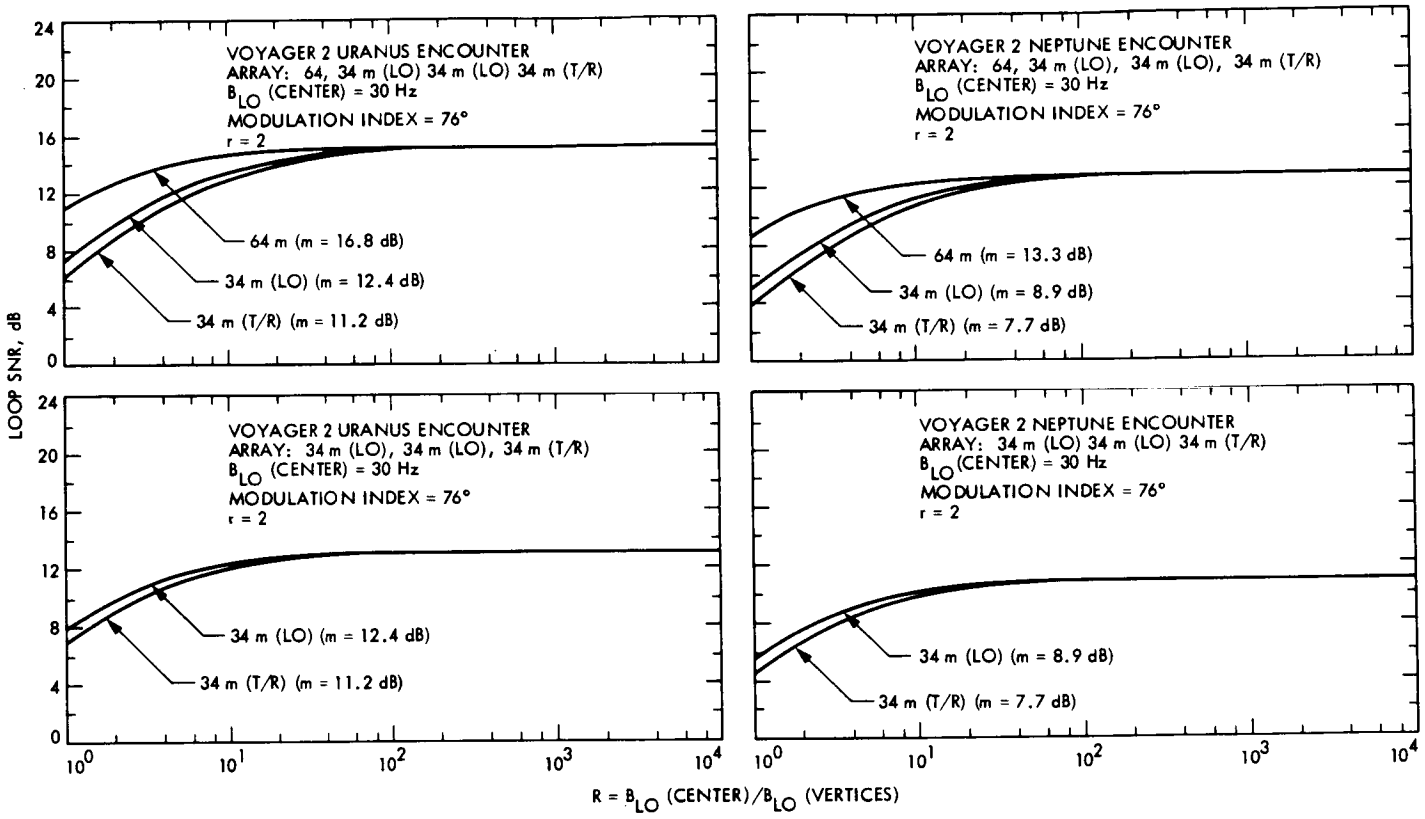



**Fig. 3. Virtual center arraying performance**

N82-11288



# The Deep Space Optical Channel: II. Wave Propagation Effects

J. Katz
Communications Systems Research Section

*Propagation effects of optical waves in the deep space optical channel with a space-based receiver are analyzed. It is found that these effects are negligible for all practical interplanetary systems designs.*

## I. Introduction

The basic considerations involved in the design of deep space optical communication systems are described in Ref. 1. Besides assessing noise effects (Ref. 2), one must also analyze the wave propagation aspects of the optical channel in order to evaluate their impact on the overall system performance.

The purpose of this report is to summarize the wave propagation characteristics which may affect the optical communication link. Since we assume that the receiving station of the link is located outside the earth atmosphere, propagation effects through the atmosphere are not analyzed here. It is found that propagation effects are negligible in all practical situations. If propagation through the atmosphere were to be included, this result would not be correct. This is one of the advantages of locating the receiver outside the earth atmosphere. (Other advantages include the reduction of the total noise in the system.)

## II. Formulas for Calculating the Propagation of Electromagnetic Waves in the Optical Channel

The medium through which the light propagates in a deep space link is basically an inhomogeneous plasma. The general formulas of electromagnetic wave propagation in plasmas are complicated (Ref. 3), but when dealing with optical frequencies (i.e., $3.10^{14} \leqslant f \leqslant 10^{15}$ Hz, corresponding to $0.3 \leqslant \lambda \leqslant 1$ $\mu$m) and with typically encountered plasmas, we can greatly simplify the calculations.

The basic reason for this possible simplification is that the region of frequencies that we are interested in lies far above any other relevant frequency of physical significance. We are mainly referring to the plasma resonance frequency, given by

$$f_p \cong 8.9 \sqrt{N} \text{ (Hz)} \tag{1}$$

(where $N$ is the elctron density in the plasma in $m^{-3}$), to the cyclotron resonance frequency of magnetic fields present in the plasma, and to the rate of collisons of an electron in the plasma with other particles, denoted by $1/\tau_s$. These conditions are discussed in more detail in Ref. 3. For operations close to any of the above frequencies, the interaction between the medium and the field is stronger and hence its consequences are more complicated to describe and evaluate.

Under the above assumptions, the field amplitude of a plane wave propagation in a plasma is typically described by the following exponential:

$$e^{-\frac{1}{2}\alpha Z} e^{i[2\pi ft-(k+\delta k)Z]} \tag{2}$$

where $\alpha$, the intensity (or power) attenuation coefficient, is approximately given by (Refs. 3, 4)

$$\alpha = \frac{1}{c}\left(\frac{f_p}{f}\right)^2 \frac{1}{\tau_s} (m^{-1}) \tag{3}$$

where $f$ is the optical signal frequency, $c$ is the velocity of light, $k$ is the wave-vector of the field, defined as

$$k = 2\pi \frac{f}{c} \ (\mathrm{m}^{-1}) \tag{4}$$

and the phase shift imposed upon the wave is associated with the following change in the wave-vector (Ref. 3)

$$\delta k = -\frac{\pi f}{c}\left(\frac{f_p}{f}\right)^2 \ (\mathrm{m}^{-1}) \tag{5}$$

The basic propagation configuration is shown in Fig. 1. We have an input beam characterized by an input power $P_{in}(W)$, pulsewidth $\tau_{in}$ (sec) and a beam divergence angle $\theta_{in}$ (rad), propagating a distance $L$(m) in a plasma with a density $N(\mathrm{m}^{-3})$, temperature $T$(K) and density variation $\delta N/N$. The emerging signal has output power $P_{out}$ (W), pulsewidth $\tau_{out}$ (sec) and a beam divergence angle $\theta_{out}$ (rad). From this and the previous assumptions, the following propagation effects will be considered:

## A. Attenuation

The intensity attenuation coefficient $\alpha$, defined as

$$\alpha \equiv \frac{1}{L} \ln \frac{P_{in}}{P_{out}} \tag{6}$$

can be calculated using Eq. (3). From Ref. 4 we can obtain the following approximate expression for the frequency of collisions $1/\tau_s$, a parameter which is needed for the calculation of $\alpha$:

$$\frac{1}{\tau_s} = 10^3 N T^{-3/2} (21.4 + 5.44 \ln T - 1.81 \ln N) \tag{7}$$

## B. Temporal Dispersion

The general formula relating the optical signal pulsewidths at the input and output is

$$\tau_{out} = \left[\tau_{in}^2 + \left(\frac{1}{2\pi}\frac{L}{\lambda}\frac{f_p^2}{f^4}\frac{1}{\tau_{in}}\right)^2\right]^{1/2} \ (\mathrm{sec}) \tag{8}$$

The derivation of this equation is similar to the derivation of the dispersion in optical fibers (Ref. 5).

Differentiating Eq. (8) with respect to $\tau_{in}$, we find that the shortest pulse that can be achieved is

$$\tau_{in}\,(\min) = \sqrt{\frac{1}{2\pi}\frac{L}{\lambda}\frac{f_p}{f^2}} \ (\mathrm{sec}) \tag{9}$$

which results

$$\tau_{out}\,(\min) = \sqrt{2}\,\tau_{in}\,(\min) \tag{10}$$

## C. Beam Divergence

The beam divergence angle is given by the following formula:

$$\theta_{out} = [\theta_{in}^2 + \langle(\delta\theta)^2\rangle]^{1/2} \tag{11}$$

$\delta\theta$ is a random variable related to the fluctuations in the index of refraction of the medium, as explained in the following paragraph.

First we denote by $L_c$ the scale of the fluctuations, which is roughly a characteristic distance over which the parameters of the medium change. For the interplanetary medium, $L_c$ is of the order of several hundreds of kilometers (Ref. 6).

When the condition $L_c \gg \sqrt{\lambda L}$ applies, we can use the simple laws of geometrical optics to analyze the effect of the randomness of the medium on the optical wave. Since this condition applies in our case, we obtain (Ref. 7).

$$\langle(\delta\theta)^2\rangle \cong \langle(\delta n)^2\rangle \frac{L}{L_c} \tag{12}$$

The next step is to relate the refractive index fluctuations $\langle(\delta n)^2\rangle$ to the plasma density fluctuations $\langle(\delta N)^2\rangle$. The index of refraction of a plasma at $f > f_p$ is (Ref. 8)

$$\epsilon = n^2 = 1 - \left(\frac{f_p}{f}\right)^2 = 1 - \frac{1}{4\pi^2}\frac{q^2}{\epsilon_0 m f^2} N \tag{13}$$

so

$$\langle(\delta N)^2\rangle = \frac{1}{4}\left(\frac{f_p}{f}\right)^4 \frac{\langle(\delta N)^2\rangle}{N^2} \tag{14}$$

Combining Eqs. (12) and (14), we obtain the desired expression for $\langle \delta\theta^2 \rangle$, which is to be used in Eq. (11):

$$\langle \delta\theta^2 \rangle = \frac{1}{4}\left(\frac{f_p}{f}\right)^4 \left(\frac{L}{L_c}\right)\left(\frac{\langle (\delta N)^2 \rangle}{N^2}\right) \qquad (15)$$

In the next section we will apply the formulas derived above to some envisioned scenarios that occur in deep space missions.

## III. Wave Propagation Effects in the Optical Channel

In this section we will apply the formulas derived in the last section to find how the optical wave is modified during propagation. Three propagation cases, shown in Fig. 2, are considered. They include (1) propagation through the interplanetary medium, (2) propagation through a big planet's (Jupiter) magnetosphere, and (3) propagation through the Sun's corona. The relevant parameters and the results for each case are given in Table 1. Values assumed for the calculations are $f = 4 \cdot 10^{14}$ Hz (corresponding to $\lambda = 0.75$ $\mu$m), $L/L_c \cong 2 \cdot 10^7$ and $\langle (\delta N)^2 \rangle / N^2 \cong 1$. Values of $N$ and $T$ for these examples were taken from Refs. 8 and 9. Most of the values assumed are on the worst-case side. Analyzing the results, we see that the power attenuation is very small, less than 0.5% in the worst-case scenario. The temporal dispersion is less than 0.1 ps, and the addition to the beam spread angle is less than 1 nrad. It is interesting to note that propagation through a planet's magnetosphere has the largest effect. This is not true if we consider propagation paths closer to the Sun than the 0.1 AU distance assumed in this report. However, in this case noise problems due to the radiation of the Sun will become very severe (Ref. 2), rendering the link impractical.

To summarize, in all three cases the beam spread and power attenuation are always negligible, and temporal dispersion is negligible for pulses longer than $10^{-13}$ sec.

Other propagation effects not related to plasma phenomena are light scattering and absorption by the interplanetary dust. It can be shown that these effects are also negligible. For example, only 0.25% of the sunlight is scattered in its propagation to earth (Ref. 9), in spite of the fact that the region near the sun has the highest concentration of interplanetary dust.

## IV. Conclusion

Propagation effects of optical waves in the deep space to near space link have been analyzed. It has been found that for practical wave propagation considerations, the medium can be taken as a free-space medium.

# References


1. Katz, J., and Lesh, J., "Optical Communication for Deep Space Missions Using Integrated Semiconductor Injection Lasers," JPL Publication (in preparation).
2. Katz, J., "The Deep Space Optical Channel: I. Noise Mechanisms," *TDA Progress Report 42-64*, Jet Propulsion Laboratory, Pasadena, Calif., August 15, 1981.
3. Condon, E. U., and Odishaw, H., *Handbook of Physics*, McGraw Hill, N.Y., 1961, Ch. 4-11.
4. Valley, S. L., *Handbook of Geophysics and Space Environment*, McGraw-Hill, N.Y., 1965, pp. 12-25.
5. Yariv, A., *Introduction to Optical Electronics*, 2nd edition, Holt, Rinehart and Winston, New York, 1976, pp. 45-50.
6. Yakovlev, O. I., "Radio Wave Propagation in the Solar System," Kunner (LEO) Associates, Redwood City, California, 1974 (74N32605).
7. Hodara, H., "Laser Wave Propagation Through the Atmosphere," *Proc. IEEE*, 54, pp. 368-375, 1966.
8. Ref. 3, p. 4-189.
9. *The Planet Jupiter*, NASA SP-8069, 1971.
10. Daniels, G. M., "A Night Sky Model for Satellite Search Systems," *Opt. Eng. 16*, pp. 66-71, 1977.

**Table 1. Parameters and results for wave propagation in the deep space link (all values are approximate)**

| Propagation medium | Conditions | | | | Results | | |
|---|---|---|---|---|---|---|---|
| | Distance $L$, m | Plasma frequency $f_p$, Hz | Electron density $N$, $m^{-3}$ | Electron temperature $T$, K | Power attenuation exponent $\alpha L$ | Temporal dispersion $\tau_{out}$(min), sec | Beam spread $\sqrt{\langle(\delta\theta)^2\rangle}$, rad |
| Interplanetary medium | $\leqslant 10^{13}$ | $2 \cdot 10^4$ | $10^5$ | $10^4$ | $4 \cdot 10^{-16}$ | $3 \cdot 10^{-15}$ | $6 \cdot 10^{-17}$ |
| Planet's magnetosphere | $\leqslant 10^9$ | $10^8$ | $3 \cdot 10^{12}$ | $10^5$ | $5 \cdot 10^{-3}$ | $10^{-13}$ | $10^{-9}$ |
| Sun's corona at distance 0.1 AU from the Sun | $\leqslant 10^{10}$ | $10^7$ | $3 \cdot 10^{10}$ | $10^6$ | $3 \cdot 10^{-8}$ | $5 \cdot 10^{-14}$ | $10^{-11}$ |

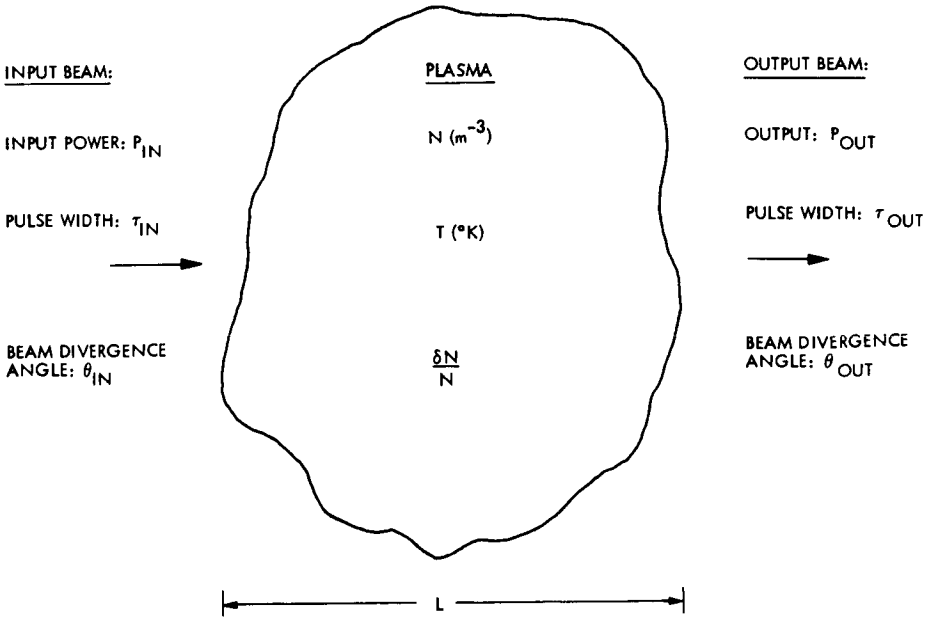



**Fig. 1. Propagation configuration**

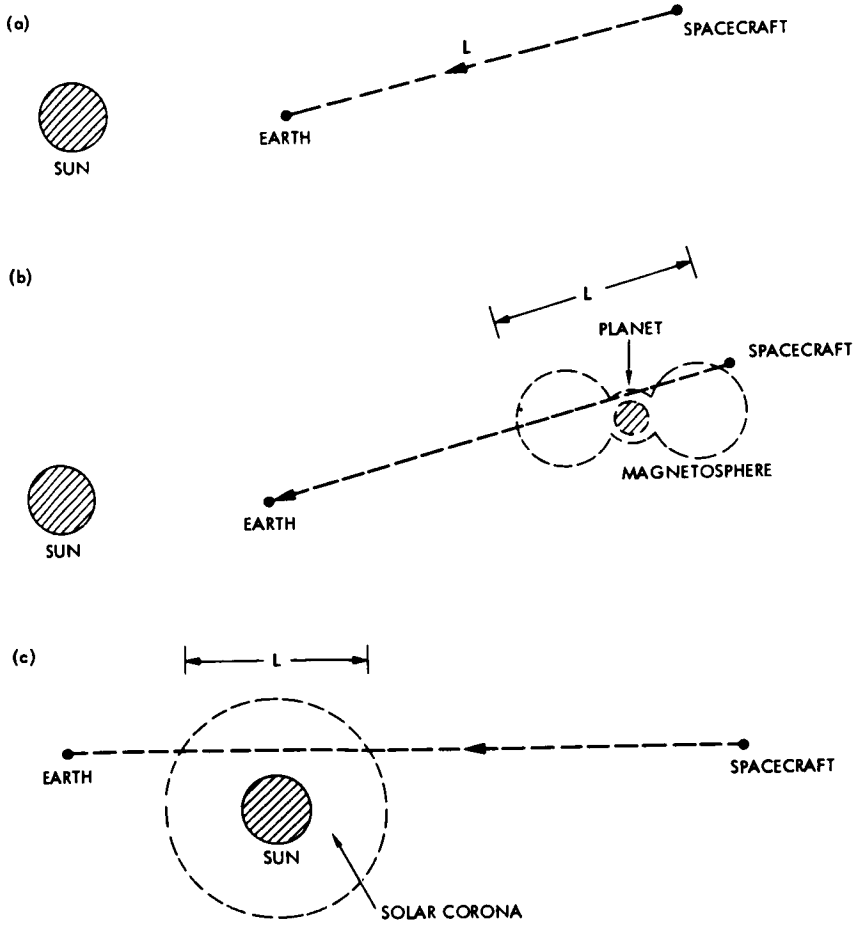



**Fig. 2. Three propagation cases (not to scale)**

N82-11289



# Symbol Error Rate Performance Evaluation of the LM37 Multimegabit Telemetry Modulator-Demodulator Unit

H. Malek
Communications Systems Research Section

*The LM37 multimegabit telemetry modulator-demodulator unit was tested for evaluation of its symbol error rate (SER) performance. Using an automated test setup, the SER tests were carried out at various symbol rates and signal-to-noise ratios (SNR), ranging from +10 to -10 dB. With the aid of a specially designed error detector and a stabilized signal and noise summation unit, measurement of the SER at low SNR was possible. The results of the tests have shown that at symbol rates below 20 megasymbols per second (MS/s) and input SNR above -6 dB, the SER performance of the modem is within the specified 0.65 to 1.5 dB of the theoretical error curve. At symbol rates above 20 MS/s, the specification is met at SNR's down to -2 dB. This report presents the results of the SER tests with the description of the test setup and the measurement procedure.*

## I. Introduction

The current telemetry data handling capability of the Deep Space Network (DSN) is limited to 250 kilosymbols per second (kS/s). Ground telemetry systems for future planetary missions, such as Venus Orbiting Imaging Radar (VOIR) and Mars Rover, require the capability of processing much higher symbol rates than can presently be handled in the DSN.

The Multimegabit Telemetry Development (MTD) project was created for the purpose of developing the needed megasymbol telemetry technology. The developed technology will be used to upgrade and expand the telemetry data handling capability of the DSN. To accomplish the MTD project objectives, a developmental program leading to a feasibility model of a multimegabit telemetry demodulator-detector was initiated at JPL. The program also included acquisition of a commercially developed multimegabit modem for test and evaluation in support of a make or buy decision. To fulfill the latter objective of the developmental program, the Linkabit LM37 multimegabit telemetry modulator-demodulator (MTMD) unit was purchased. This report presents the SER test and evaluation process of the commercial modem. Following are the descriptions of the modem, the test setup, the measurement procedure, the test results, and evaluation of the results with some concluding remarks.

## II. LM37 MTMD Unit

The LM37 MTMD unit is a full duplex, digital biphase, shift-keyed modem (Ref. 1). The unit has the capability of

modulating and demodulating a suppressed carrier at an intermediate frequency (IF) of 55 MHz with data rates ranging from 0.5 to 30 MS/s. Figure 1 shows the block diagram of the modem.

### A. Modulator

The modulator portion of the modem has been provided for the purpose of testing the demodulator. The modulator accepts the telemetry (TX) data and the associated TX clock at a differential emitter coupled logic (ECL) level. The falling edge of the TX clock enters the TX data into a flipflop whose output biphase modulates the 55-MHz carrier signal. The carrier signal is generated by a phase-locked loop whose reference frequency is an external 5-MHz source. To eliminate carrier harmonics, the biphase modulated signal is first passed through a four-pole Butterworth bandpass filter, with a bandwidth of 60 MHz, centered at the carrier frequency of 55 MHz. The bandpass signal is then amplified to a zero dBm output level.

### B. Demodulator

The demodulator portion of the LM37 MTMD unit produces four-bit (soft decision) data, labeled as RX data, and the corresponding symbol clock at a differential ECL level. The front end of the demodulator is an analog quadrature demodulator whose outputs are the lowpassed in-phase (I) and quadrature (Q) baseband signals. The I and Q channels are digitized into four bits of data, at a rate between 32 and 64 MS/s, depending on the input symbol rate. The digitized I and Q data are then used in all the subsequent digitally implemented functions. These functions include carrier acquisition and tracking, symbol acquisition and tracking, symbol and carrier lock detection, automatic gain control (AGC), and data detection.

The demodulator symbol rates, ranging from 0.5 to 30 MS/s, are entered through eight thumbwheel switches on the front panel. The front panel also contains the bit (symbol) and the carrier lock indicator lights. All of the input/output (I/O) signal connectors are located at the rear panel.

## III. Automated Test Setup

Figure 2 shows the block diagram of the automated test setup which was used in the SER test of the LM37 MTMD unit. By configuring test instrumentation and support hardware around the HP-9845C desk top computer, accurate signal measurement and rapid data logging were possible. This test setup also provided a capability for on-line computation and verification of the test parameters. The functions of the various instruments in the test setup are given in the test procedure section. The following is a description of the special support hardware developed for the SER tests.

### A. Noise Generator

The required white noise for the SER tests was generated by the noise generator of Fig. 3. The operation of the unit is based on amplification of the thermal noise contributed by the 50-ohm input terminating resistor and the equivalent input impedance of the preamplifiers. The bandwidth of the noise at the output of the preamplifiers extends from 5 to 500 MHz. Mixing the noise with a local oscillator (LO) accomplishes two objectives. First, the noise is translated closer to dc, and second, by selecting an appropriate LO frequency, noise folding can be used to improve noise spectrum flatness.

To exclude the higher frequency components of the mixed signal, it is passed through a five-pole Butterworth lowpass filter with 150-MHz bandwidth. The lowpass signal is then amplified and becomes the output noise signal with a uniformly distributed power spectrum in the desired frequency range. The regulation of the generated noise was measured to be ±0.2 dB. Figure 4 shows the plot of the power density spectrum of the generated noise signal.

### B. Stabilized Signal and Noise Summation Unit

The output power level of the noise generator is a function of the system noise temperature. By simply adding the generated noise to the signal, the resulting SNR level will vary as the system temperature varies with time. Such a variation in the input SNR will greatly affect the accuracy of the SER tests since the symbol errors are time averaged.

The stabilized signal and noise summation unit, shown in Fig. 5, was designed to maintain the input SNR at constant levels throughout the SER measurements. The SNR stability is achieved through a feedback control loop. Equal portions of signal and noise are compared by the null detector of Fig. 6. The result of the comparison (output of the null detector) is the control signal which is applied to the AGC circuit. The AGC circuit (see Ref. 2) changes gain in response to the above control signal, thus maintaining a constant SNR. The regulation of the feedback control loop was measured to be within ±0.1 dB for ±5-dB changes in the signal or the noise power levels.

### C. Interface Unit

An interface unit was designed to provide the required digital I/O signal levels for the LM37 MTMD unit. Figure 7 shows the schematic of the interface unit. The unit consists of an input and an output translator for the modulator input and the demodulator output. The input section translates the TX data and clock signals from a TTL level into a differential ECL level. The output section converts the symbol data and clock output of the demodulator from a differential ECL level into a TTL level. For the TTL-level translation, the

receivers and drivers are designed for 50-ohm coaxial cables. The receivers and drivers for the ECL-level translators are designed for 120-ohm triaxial cables.

### D. Delay and Error Detector Unit

The delay and error detector unit of Fig. 8 was designed to detect symbol errors at very low SNR levels. Through a shift register, the data from the word generator (TX data) is delayed by a time delay equal to the internal delay of the LM37 MTMD unit. The delay time is set by a thumbwheel switch located on the front panel of the unit. A delay resolution of one-half symbol is required to permit clocking with the LM37 clock. To obtain this resolution, the TX data is shifted at twice the frequency of the TX clock. The shifting clock is divided by 2 and output as the TX clock.

The error detector section compares the delayed TX data with the detected RX symbol data from the demodulator. An error pulse is output when the RX symbol does not agree with the delayed TX symbol. To insure proper timing and also allow sufficient data settling time, the delayed TX data and the RX data are both sampled and stored for comparison at the trailing edge of the RX symbol clock from the demodulator. The two samples are then compared by means of an exclusive OR gate. The return-to-zero (RZ) symbol error pulse is obtained by gating the output of the exclusive OR gate with the delayed RX clock.

Because of the Costas carrier tracking loop in the demodulator, there is a 180° ambiguity in the polarity of the demodulated symbols. When the RX data is in the true or noninverted state, the detected symbol error pulses correspond to the actual symbol errors. However, when the RX data is in the complemented form, the detected symbol errors are also complemented. This ambiguity in the RZ symbol error pulse is resolved in the following way: For binary symbol data, the maximum possible error probability is 0.5. In terms of the SER, the maximum error probability corresponds to one-half of the input symbol rate. When the rate of the detected symbol errors is less than or equal to the above maximum rate, the resulting error pulses correspond to the actual SER. When the rate of the detected errors exceeds the maximum possible rate, the actual SER is obtained by subtracting the detected symbol error rate from the input symbol rate.

## IV. Test Procedure

The SER testing of the LM37 MTMD unit was carried out in two phases. In the initial phase, the power measuring instruments, such as the HP-8568A spectrum analyzer and the HP-436A power meter, were calibrated with a known reference signal. Then the calibrated power meter was used to determine power correction factors for the various symbol rates.

The second phase of the test consisted of a sequence of signal and noise power level settings, followed by the measurement of the input SNR and the corresponding SER. The sequence was repeated for each symbol rate. Each step in a sequence was executed according to directives issued by an HP-9845C computer control and monitor program. The detailed description of a typical sequence is given in Appendix A. Following is the description of the measurement methods and the technical aspects of the test procedure. The description makes reference to a more detailed block diagram (Fig. 9) of the test setup.

### A. Calibration

The premeasurement calibration steps were carried out to insure the accuracy of the measured input SNR. To reduce the effect of the absolute accuracy of the test instruments, both the signal and the noise power levels were measured with the same instrument, namely, the HP-8968A spectrum analyzer.

The absolute accuracy of the measured power, displayed by the maximum value of the marker on the spectrum analyzer, is a function of several parameters. These parameters include uncertainty of the calibration and the reference level, the real fidelity between the reference level and the marker position, and the bandwidth accuracy of the selected filter (see Ref. 3). To eliminate the uncertainty in the calibration, the spectrum analyzer and the power meter were calibrated with the same known reference signal. For maximum accuracy, the coupled functions on the spectrum analyzer were selected such that the power readings on both instruments differed by less than 0.01 dB. The uncertainty in the reference level was removed by measuring the signal and the noise power levels using the reference levels on the spectrum analyzer which were calibrated with the power meter. By maintaining a 1-dB per division scale resolution on the spectrum analyzer, the scale fidelity error was reduced to within 0.01 dB.

### B. Power Correction Factors

The power correction factors at different symbol rates are the ratios of the modulated signal power to the signal power with the modulation removed, as measured with the HP-436A power meter. Table 1 gives the value of the power correction factor as a function of the modulation symbol rate. These factors are required because of the effect of the output bandpass filter of the modulator. At low symbol rates, the difference in the power levels between the unmodulated signal and the modulated signal is small, since most harmonics of the modulated signal are present in the modulator output. The power difference becomes larger as the symbol rate increases, since

more harmonics of the modulated signal fall outside the bandwidth of the output filter.

### C. Signal and Noise Power Measurement

For the evaluation of the input SNR, the power levels in the noise and the signal channels were separately measured with the HP-8568A spectrum analyzer. As shown in Fig. 9, the power measurement at the coupled port of the directional coupler 3 measures the same power ratios as appear at the input to the demodulator. The signal or the noise channel measurements were selected by the proper positions of the SW1 and SW2 switches. The SW3 switch was in the power meter position only during the calibration and the computation of the power correction factors.

In the SER curves, the input SNR level is expressed in terms of the ratio of the energy per symbol ($ST_s$) to the noise power spectral density ($N_0$). $N_0$ can be directly measured with the spectrum analyzer in units of dB/Hz in the selected bandwidth. For greater accuracy the value of $N_0$ was obtained from the average of 100 points, measured in a 25-85 MHz bandwidth. The selected bandwidth was equal to the bandwidth of the input filter in the LM37 demodulator.

It is difficult to directly and accurately measure the total modulated signal power from its power spectrum, especially when the carrier is modulated with pseudorandom (PN) data. For this reason, the modulated signal power level was determined by first measuring the unmodulated carrier signal power with the spectrum analyzer. This was done by removing the PN data (through the SW4 switch) from the modulator input. The total modulated signal power was then obtained by correcting the measured carrier power with the corresponding power correction factors. The following expression was used to compute $ST_s$,

$$ST_s = (P_c \cdot F_{ps})/S$$

where $P_c$ is the carrier power, and $F_{ps}$ is the power correction factor at symbol rate $S$.

### D. Symbol Error Rate Measurement

SER was measured by counting the symbol error pulses with the HP-5370A counter. To reduce the variance of the sampled error pulses, a time average of 100 seconds was used. The value of SER was determined by the rat[illegible] of the averaged symbol error pulses to the modulation rate (symbol rate).

### E. Control and Monitor Program

Under the supervision of a software program, written in BASIC, the HP-9845C computer controlled and monitored the flow of data to and from the test instruments. The interfacing of the test instruments was accomplished with the HP-IB (IEEE 488) interface card.The control program monitored the test process and issued commands to the instruments. The program also displayed the directives for the manual functions, such as keyboard entry of the test parameters. The measured data were manipulated and stored for off-line graphic display of the test results.

## V. Test Results

The results of the SER tests for the various symbol rates are summarized in two types of graph. These graphs, consisting of the SER and the input SNR degradation curves, were prepared with the HP-9845C computer and the HP-9872B plotter. The following are the descriptions of the two types of graph.

### A. SER Curves

The SER curves present the relations between the LM37 demodulator input SNRs and its output probability of symbol errors ($P_e$) at the various symbol rates. These curves were generated by plotting the logarithm of SER versus the $ST_s/N_0$, expressed in units of dB. For comparison, the theoretical error probability curve is plotted along with each of the SER curves. Figure 10 corresponds to the results of the SER tests at extreme and midrange symbol rates.

For the plotting of the theoretical error curve, the error probability for a given SNR was computed from the following expression,

$$P_e = \tfrac{1}{2}\, erfc\,(\sqrt{ST_s/N_0}) = \tfrac{1}{2}\,[1 - erf\,(\sqrt{ST_s/N_0})]\ .$$

The *erf* function was evaluated by using the following approximation,

$$erf(x) = 1 - (a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 t^5)e^{-x^2} + \epsilon(x),$$

where $t = 1/(1 + P)$, $|\epsilon(x)| < 7.5 \times 10^{-8}$, $p$ and $a_i$, $i = 1, 2, \ldots, 5$ are known coefficients (see Ref. 4).

### B. Input SNR Degradation Curves

The input SNR degradation curves represent a quantitative measure of the demodulator SER performance. Figure 11 shows the input SNR degradation curves for 11 different symbol rates.

The input SNR degradation curves were generated from the SER curves by first determining for each measured value of SER ($P_e$) the corresponding SNR values from the measured and the theoretical SER curves. The difference between the two SNR values constituted the degraded input SNR. These different values were then plotted versus the SNR values obtained from the measured SER curves. The SNR values from the theoretical curves were determined by evaluating the following rational approximation of the inverse *erfc* function

$$ST_s/N_0 = [erfc^{-1}(2P_e)]^2$$

$$erfc^{-1}(2P_e) = \frac{1}{\sqrt{2}}\left[t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3} + \epsilon_Q\right]$$

where $t = \sqrt{\ln 1/(P_e^2)}$, $|\epsilon_Q| < 4.5 \times 10^{-4}$, and the $c$ and $d$ coefficients are known constants (see Ref. 4).

## VI. Conclusions and Remarks

The SER performance of the LM37 modulator unit for symbol rates ranging from 0.5 to 20 MS/s and input SNR above 6 dB is within the specified 0.65 to 1.5 dB of the theoretical error probability curve. At symbol rates higher than 20 MS/s, however, the SER specification is not met for input SNRs lower than -2.0 dB. Because of loss of the carrier lock, SER testing at input SNR lower than -6 dB was not possible. Table 3 shows the LM37 modulator SER specifications at $ST_s/N_0 = -4$ dB against the measured values.

The effect of the change in bandwidth of the demodulator input filter on SER performance is observed at some selected symbol rates. For example, contrary to what is expected, the input SNR degradation at 8.0 MS/s is 0.2 dB lower than at 7.99 MS/s. This difference in the SER performance is due to a change of the input filter to a wider bandwidth at 8.0 MS/s than at 7.99 MS/s.

The automated test setup will be used to further test and evaluate the LM37 MTMD unit. The followup testing will include acquisition time and phase jitter measurements. The test setup will also be used to evaluate the feasibility model of the multimegabit modulator-detector system which is being developed at JPL.

## References


1. *Technical Manual For Megasymbol Telemetry Modulator-Demodulator*, Linkabit Corp, Part No. 21340, May 1980.
2. Stevens, G. L., "Input Signal Conditioner for the Multimegasymbol Telemetry System Feasibility Model," *TDA Progress Report 42-58*, May-June 1980, pp. 49-58, Jet Propulsion Laboratory, Pasadena, Calif., August 15, 1980.
3. *Operating and Service Manual for 8568A Spectrum Analyzer*, Volume 1, Hewlett Packard Part No. 08568-90012, October 1978.
4. Abramowitz and Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards, 1968.

**Table 1. Power correction factor**

| Symbol rate, MS/s | Correction factor, dB |
|---|---|
| 0.50 | −0.060 |
| 1.00 | −0.060 |
| 3.99 | −0.060 |
| 4.00 | −0.06 |
| 7.99 | −0.07 |
| 8.00 | −0.07 |
| 15.99 | −0.14 |
| 16.00 | −0.14 |
| 20.00 | −0.18 |
| 29.99 | −0.26 |
| 30.00 | −0.26 |

**Table 2. LM37 internal delay time and required setting on the delay and error detector unit**

| Symbol rate, MS/s | Internal delay | | Delay switch setting |
|---|---|---|---|
| | Time, nsec | Symbols | |
| 0.50 | 8300 | 4.15 | 0-1 |
| 1.00 | 4260 | 4.26 | 0-1 |
| 3.99 | 1227 | 4.89 | 2-3 |
| 4.00 | 1265 | 5.06 | 2-3 |
| 7.99 | 714 | 5.70 | 3-4 |
| 8.00 | 763 | 6.10 | 4-5 |
| 15.99 | 463 | 7.40 | 6-7 |
| 16.00 | 509 | 8.14 | 8-9 |
| 20.00 | 438 | 8.76 | 10 |
| 29.99 | 352 | 10.50 | 12-13 |
| 30.00 | 350 | 10.50 | 12-13 |

**Table 3. LM37 SER performance specifications vs measured values at $ST_S/N_0$ = -4 dB**

| Symbol rate, MS/s | SER deviation from theoretical | | Performance margin, dB |
|---|---|---|---|
| | Specification, dB | Measured, dB | |
| 0.50 | 0.65 | 0.76 | −0.11 |
| 1.00 | 0.65 | 0.69 | −0.04 |
| 3.99 | 0.65 | 0.50 | +0.15 |
| 4.00 | 0.65 | 0.55 | +0.10 |
| 7.99 | 0.65 | 0.71 | −0.06 |
| 8.00 | 0.65 | 0.51 | +0.14 |
| 15.99 | 1.00 | 0.97 | +0.03 |
| 16.00 | 1.00 | 1.00 | 0.00 |
| 20.00 | 1.00 | 1.01 | −0.01 |
| 29.99 | 1.50 | 2.55 | −1.05 |
| 30.00 | 1.50 | 2.44 | −0.94 |

# Appendix A

# Measurement Procedure

The following is a typical sequence of steps which were carried out in the second phase of the test procedure:

1. Set the LM37 unit at the test symbol rate.
2. Set the HP 3225A frequency synthesizer at twice the test symbol rate.
3. Select the delay time corresponding to the test symbol rate (see Table 2).
4. Set SNR approximately to 10 dB.
5. Press the RUN key on the HP 9845C computer keyboard. From this point on the measurement procedures followed the control and monitor program directives.

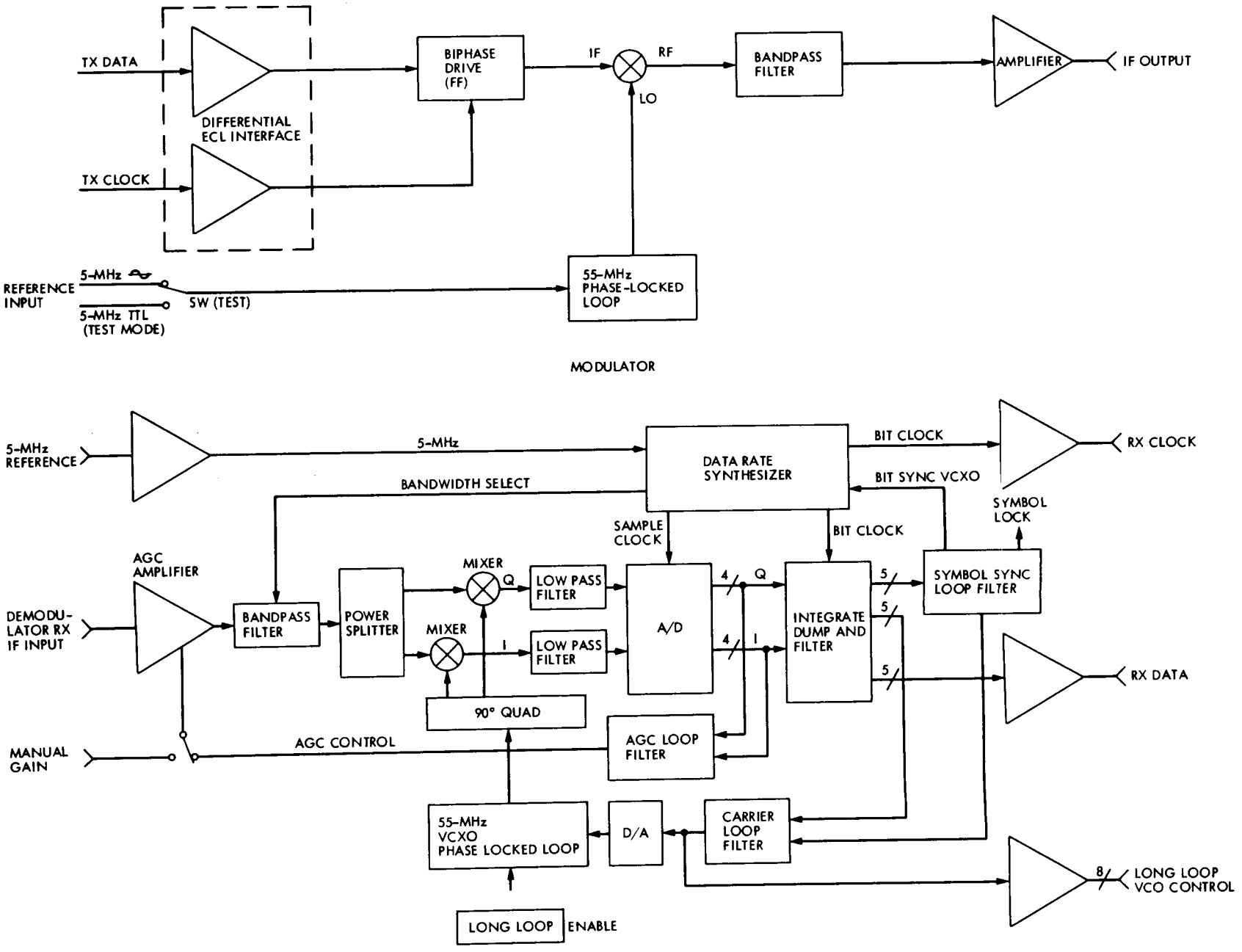



**Fig. 1. LM37 MTMD unit block diagram**

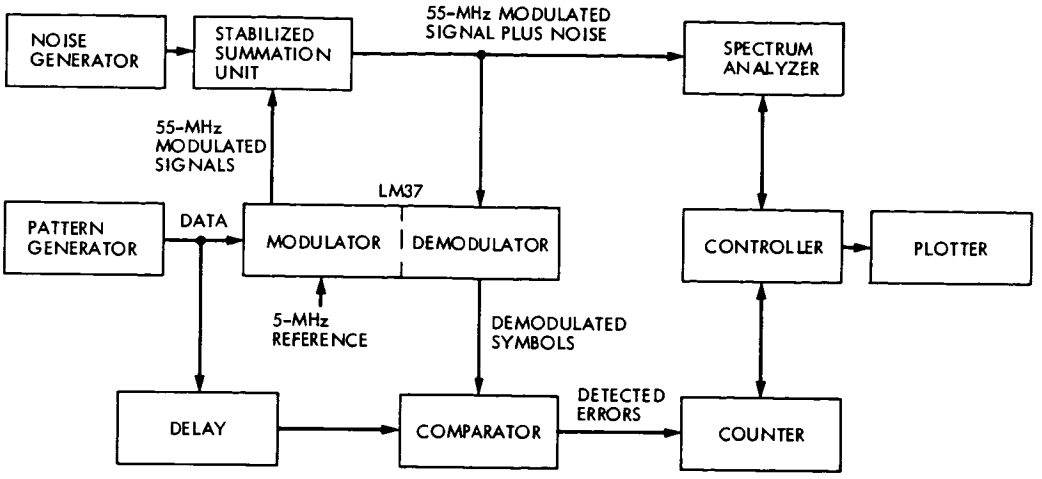
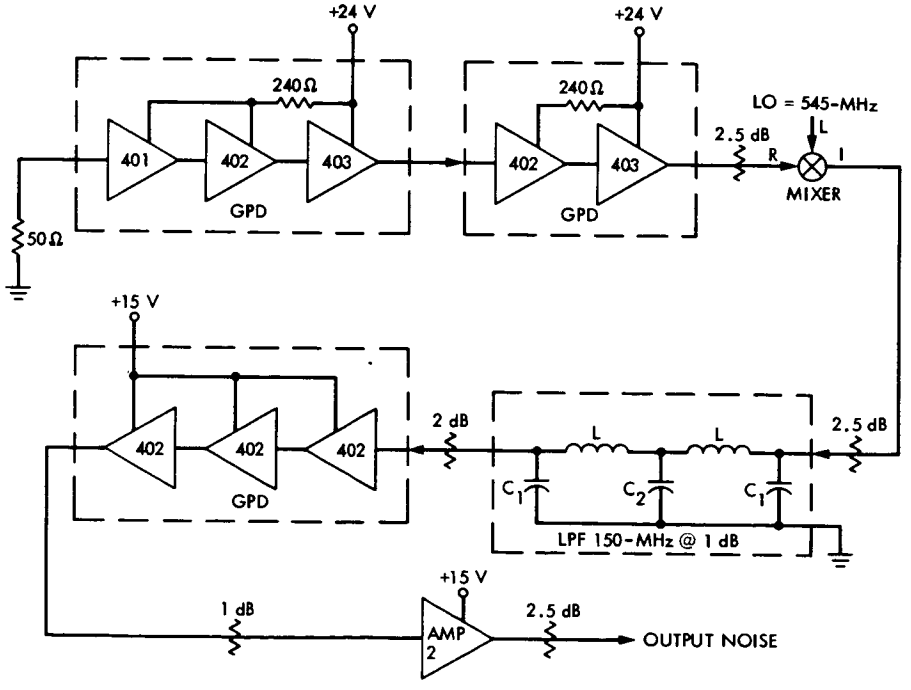



**Fig. 2. Symbol error rate test setup**




GPD ≙ MICROCIRCUIT AMPLIFIER
L = 0.085 μH
$C_1$ = 13.1 pF
$C_2$ = 42 pF
AMP 2 ≙ AVANTEK UTO 501 AND 502 AMPLIFIER

**Fig. 3. Noise generator**

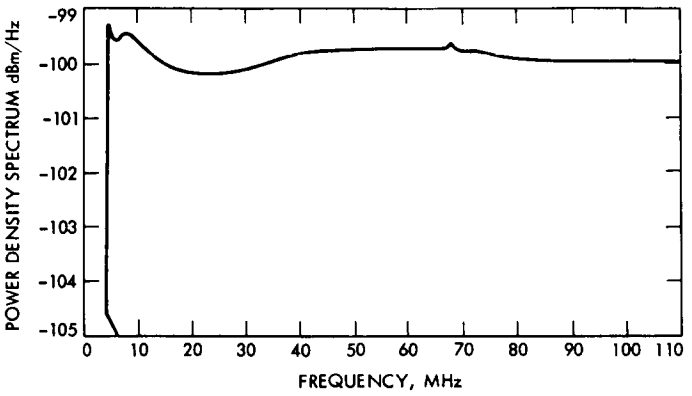



**Fig. 4. Generated noise power density spectrum**

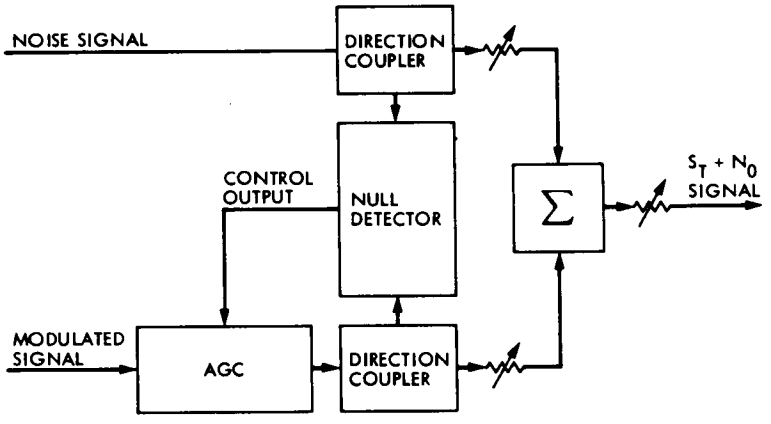



**Fig. 5. Stabilized signal and noise summation unit**

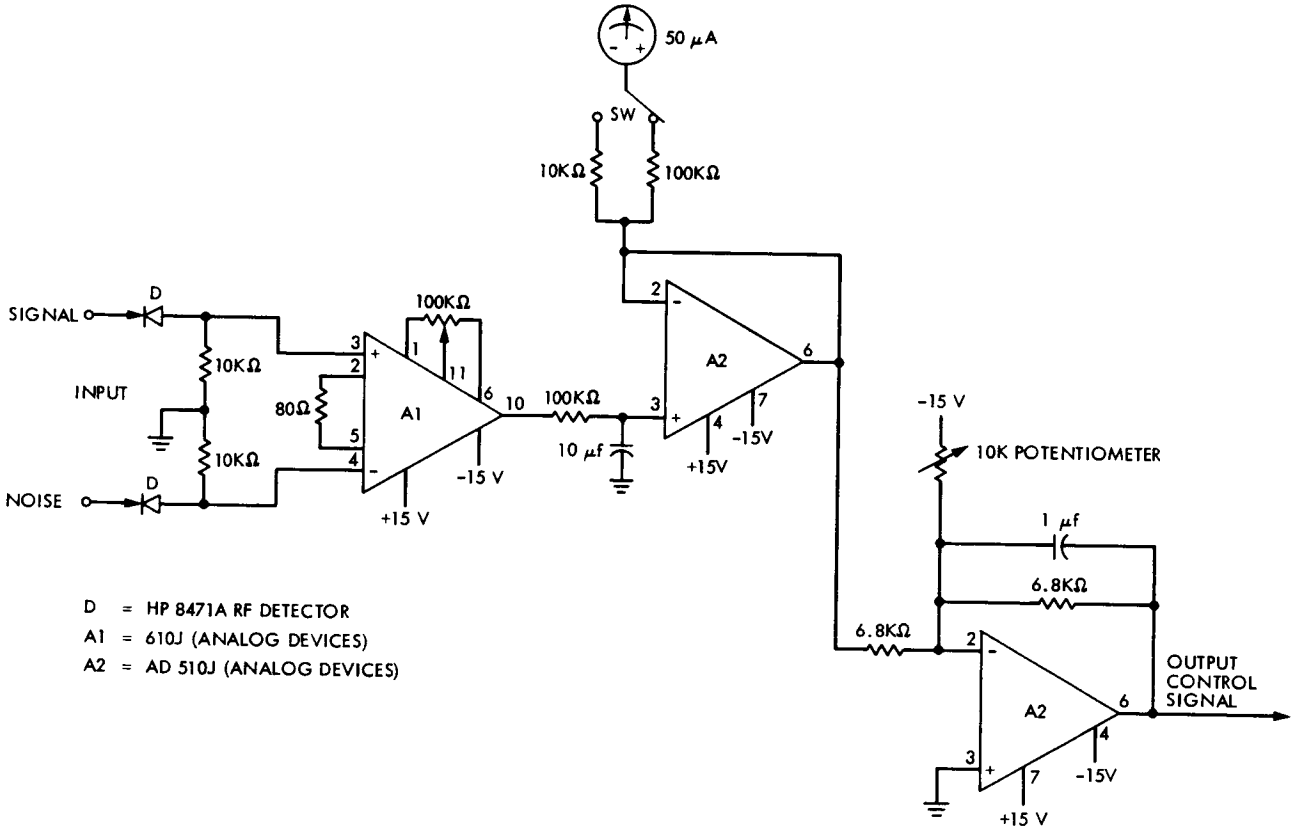
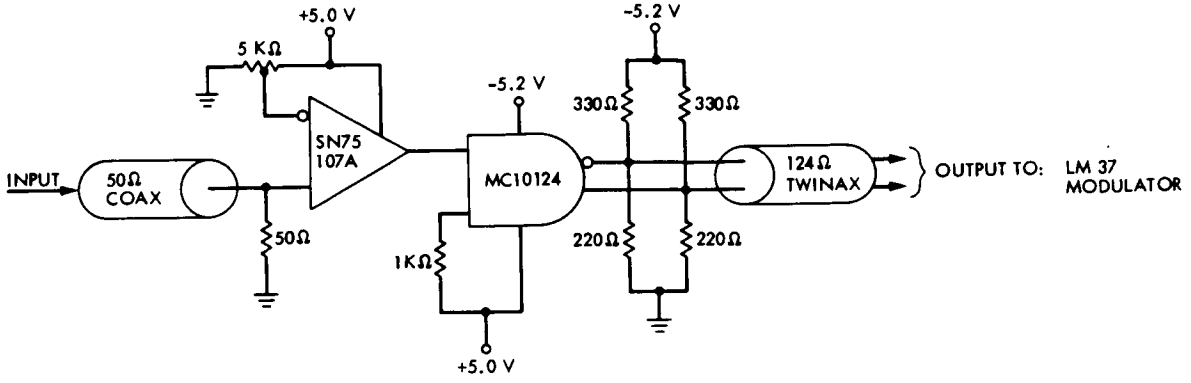
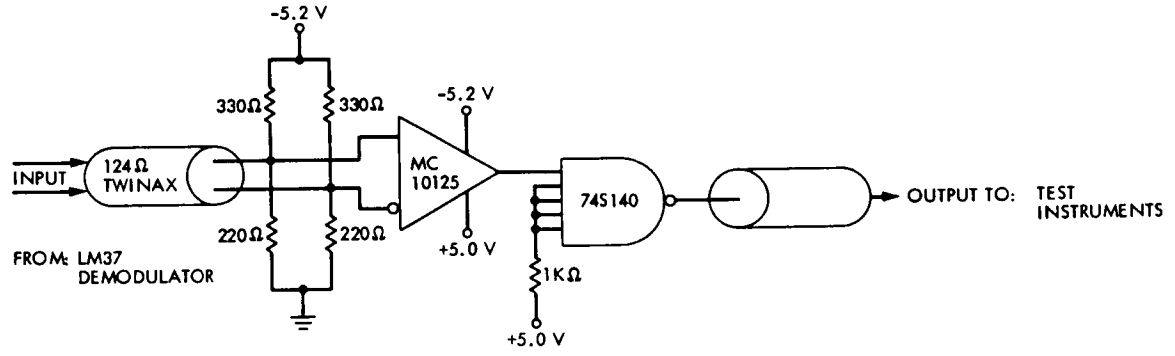



**Fig. 6. Null detector**




A TYPICAL INPUT SECTION




A TYPICAL OUTPUT SECTION

**Fig. 7. Interface unit**

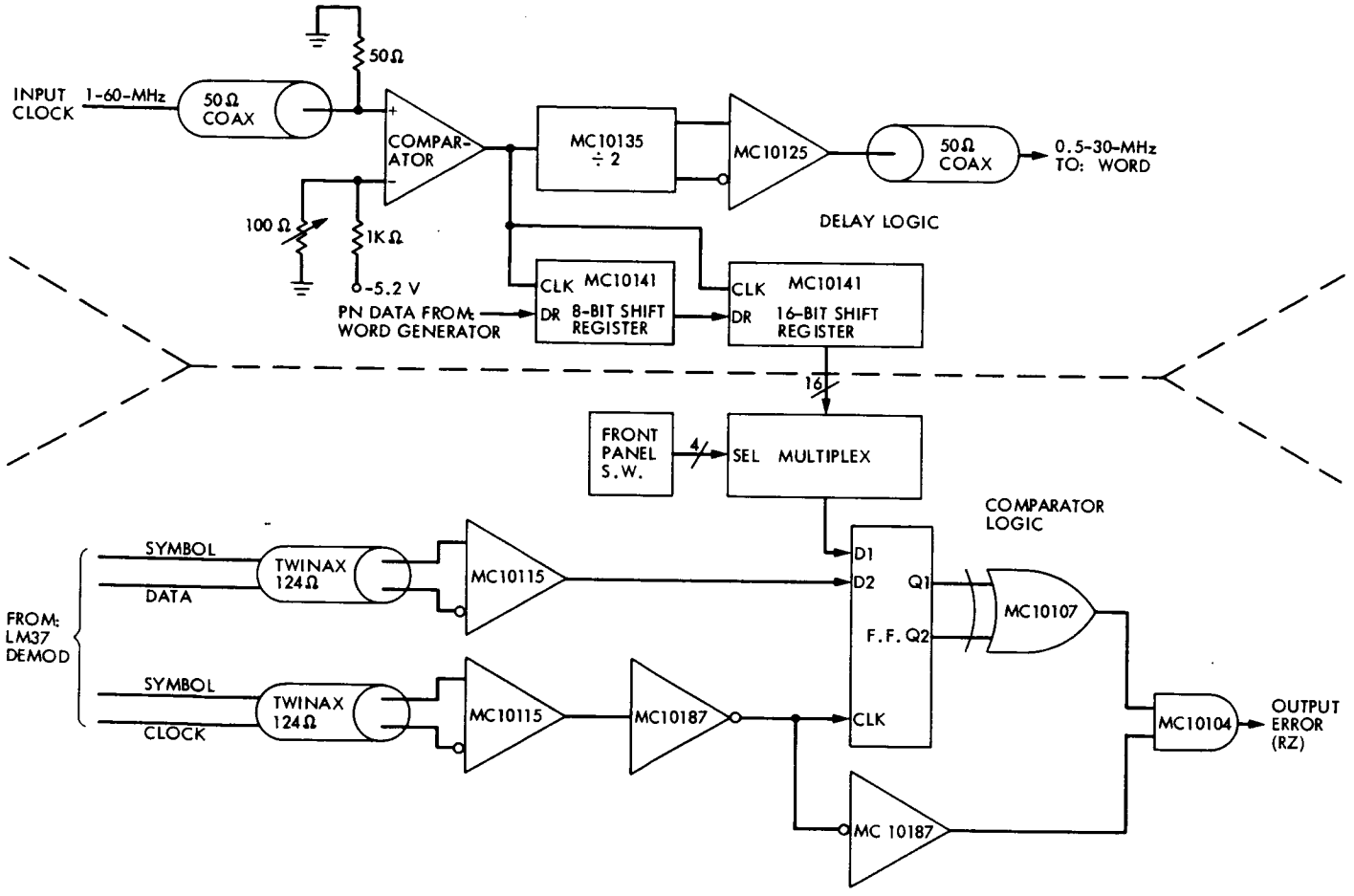
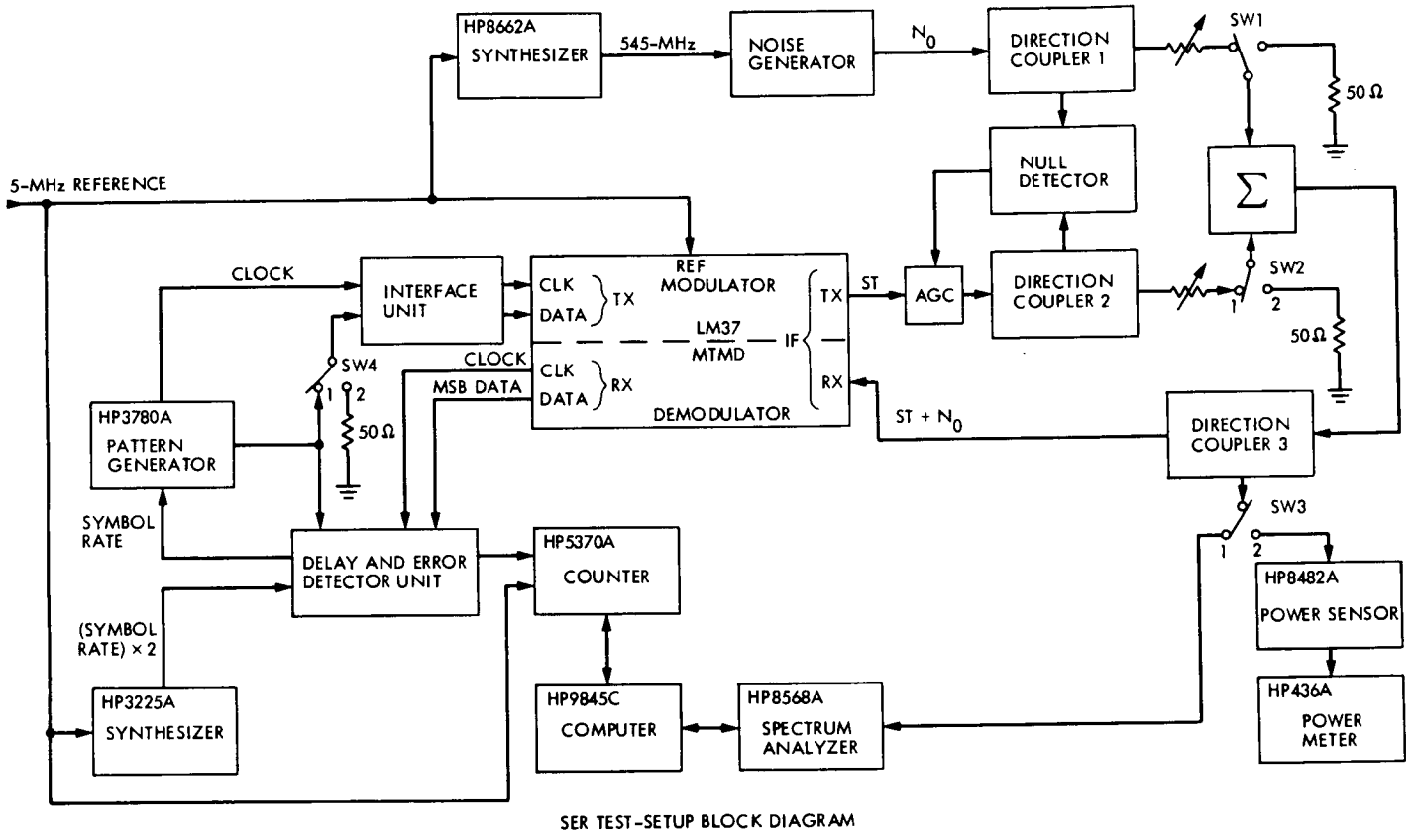



**Fig. 8. Delay and error detector unit**




SER TEST-SETUP BLOCK DIAGRAM

**Fig. 9. SER test setup block diagram**

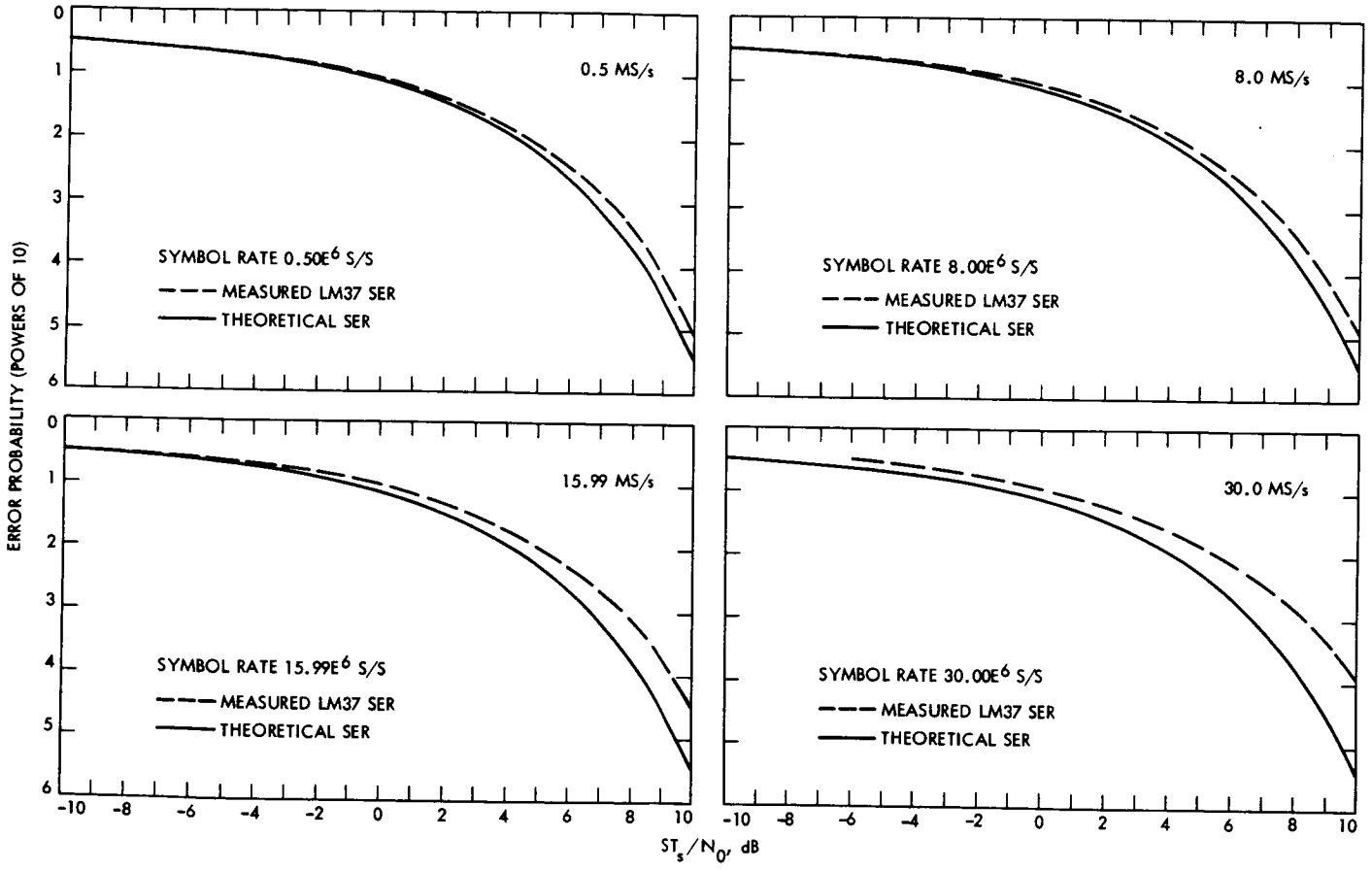



**Fig. 10. SER curves**

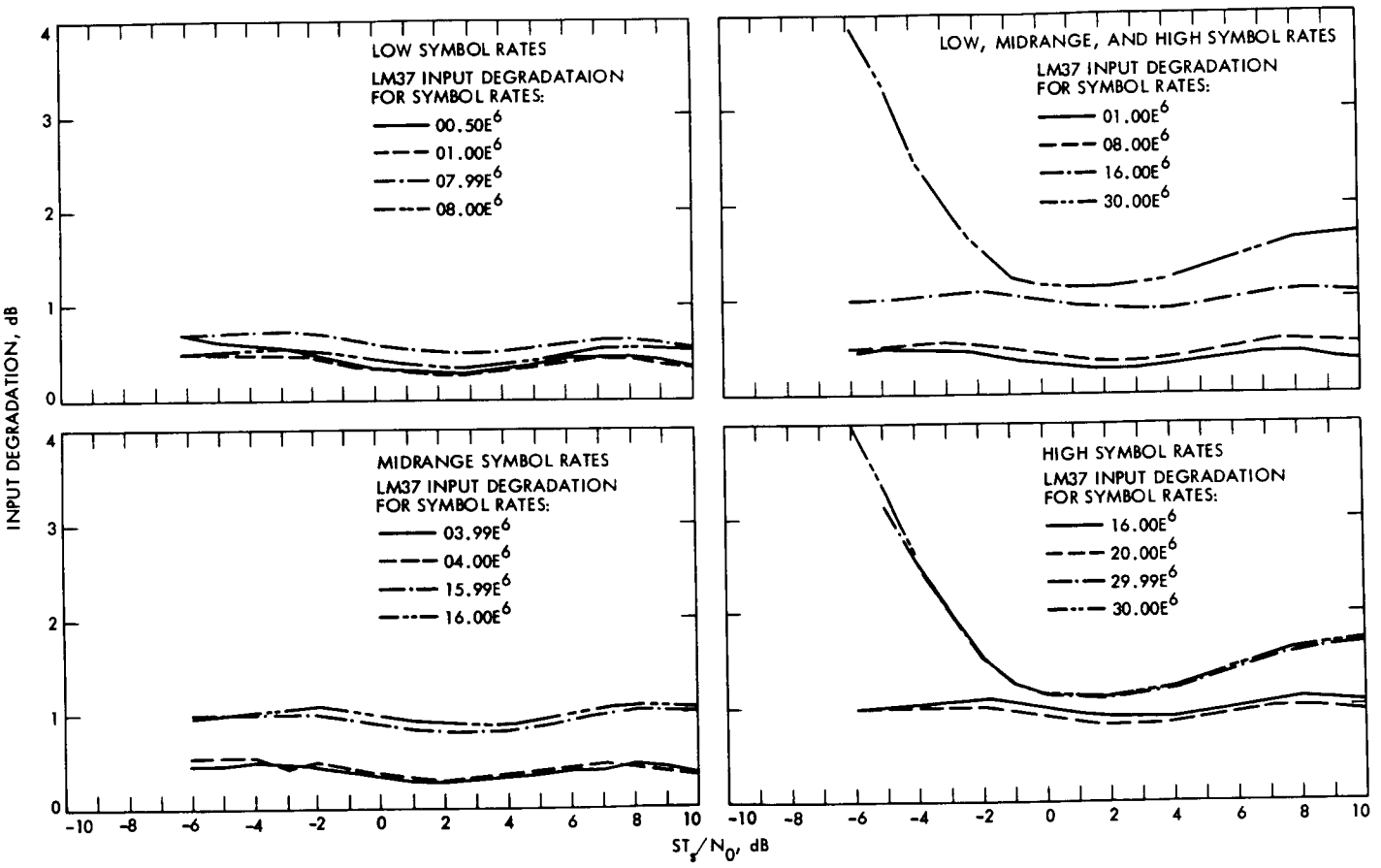



**Fig. 11. Input SNR degradation curves**

N82-11290



# RFI Channels, II

R. J. McEliece
Communications Systems Research Section

*Continuing a previous study, we calculate the cutoff parameters for a class of channel models exhibiting burst noise behavior and evaluate the performance of interleaved coding strategies. We conclude that, provided the channel memory is large enough and is properly exploited, interleaved coding is nearly optimal.*

## I. Introduction

In Ref. 1 we introduced a general class of discrete channels, which we called RFI channels. These channels, which were motivated by an earlier study (Ref. 2) of pulse-position modulation in optical channels, exhibit a simplified kind of burst-noise behavior, and in Ref. 3 we calculated their channel capacities. In this paper, after reviewing our previous results in Section II, we continue our analysis of RFI channels, as follows. In Section III we give formulas for $R_0$, the cutoff parameter for these channels. This parameter is considered by many engineers to be a more meaningful measure of the channel's quality than capacity, and its behavior on RFI channels is quite interesting. In Section IV we begin to deal with the practical problems of coding for RFI channels by considering the merits of interleaved codes. In Section V we give a numerical calculation to illustrate our results and state our conclusions.

## II. Review of Previous Results

We start with a set $\{\zeta_1, \zeta_2, \ldots, \zeta_K\}$ of $K$ discrete memoryless channels, each with the same input alphabet $A$, and output alphabet $B$, and a probability vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_K)$, with $K$ nonzero components. For each positive integer $b$, we define two *mixture channels* $\zeta(b)$ and $\overline{\zeta}(b)$, as follows. When a sequence of letters from $A$ is to be sent over $\zeta(b)$, each block of $b$ consecutive letters (such blocks we call *packets*) is transmitted over one of the auxiliary channels $\zeta_k$. Which channel is selected to transmit a given packet is determined by an external random variable $Z$, which is described statistically by $Pr\{Z = k\} = \alpha_k$. Informally we think of $Z$ as determining the noise *level* and $b$ (the *burst length*) as the length of time that a given noise level persists. Formally, the channel $\zeta(b)$ can be viewed as an orthodox discrete memoryless channel (DMC) with input alphabet $A^b$, output alphabet $B^b$, and transition probabilities

$$p(\underline{y}|\underline{x}) = \sum_{k=1}^{K} \alpha_k \prod_{i=1}^{b} p_k(y_i|x_i)$$

where $\underline{y} = (y_1, y_2, \ldots, y_b)$, $\underline{x} = (x_1, x_2, \ldots, x_n)$, and $p_k(y|x)$ denotes the transition probability function of the channel $\zeta_k$.

The second RFI channel $\overline{\zeta}(b)$, which might be called *the channel with side information*, is identical to $\zeta(b)$ except that the channel provides to the user along with every packet the index of the channel used to transmit that packet. Informally, this side information can be thought as being provided by a noise level detector, perhaps an automatic gain control device.

Formally $\overline{\zeta}(b)$ is a DMC with input alphabet $A^b$, output alphabet $B^b \times \{1, 2, \ldots, K\}$, and transition probabilities

$$\overline{p}(\underline{y}, k|\underline{x}) = \alpha_k \prod_{i=1}^{b} p_k(y_i|x_i) .$$

Let us denote the capacities of $\zeta(b)$ and $\overline{\zeta}(b)$ by $C(b)$ and $\overline{C}(b)$, respectively. The main results of Ref. 1 can be summarized as follows. First, $\overline{C}(b)$ is independent of $b$, and under a mild additional hypothesis* we have the formula

$$\overline{C}(b) = \overline{C} = \sum_{k=1}^{K} \alpha_k C_k ,$$

where $C_k$ denotes the capacity of $\zeta_k$. Second, $C(b)$ is less than $\overline{C}(b)$, in general strictly less, and in fact

$$\overline{C}(b) - \frac{\log K}{n} \leqslant C(b) \leqslant \overline{C}(b)$$

It follows in particular that $\lim_{b\to\infty} C(b) = \overline{C}$.

## III. The Calculation of $R_0$

For a general DMC, the calculation of $R_0$ depends on the function $j(x_1, x_2)$, defined for pairs of input letters:

$$j(x_1, x_2) = \sum_{y \in B} p(y|x_1)^{1/2} \, p(y|x_2)^{1/2} .$$

If then $X$ is any random variable taking values in the input alphabet $A$ we define $J(X) = E[j(X_1, X_2)]$, where $X_1$ and $X_2$ are independent random variables with the same distribution as $X$. Finally,

$$R_0 = \max_X \{-\log J(X)\} .$$

We now consider the channel $\overline{\zeta}(1)$, which is the easiest case. Here the output alphabet is $B \times \{1, 2, \ldots, K\}$. Let us denote the function $j$ for this channel by $\overline{j}$, and the corresponding functions for the auxiliary channels $\{\zeta_k\}$ by $\{j_k\}$.

*This hypothesis is that the *same* input distribution can be used to achieve channel capacity on each of the auxiliary channels. Throughout the paper we will describe this state of affairs by saying the channels $\{\zeta_i\}$ are *compatible*.

*Lemma 1*:

$$\overline{j}(x_1, x_2) = \sum_{k=1}^{K} \alpha_k j_k(x_1, x_2) .$$

*Proof*:

$$\overline{j}(x_1, x_2) = \sum_{y,k} \overline{p}(y, k|x_1)^{1/2} \, \overline{p}(y, k|x_2)^{1/2}$$

$$= \sum_{y} \sum_{k} [\alpha_k p_k(y|x_1)]^{1/2} \, [\alpha_k p_k(y|x_2)]^{1/2}$$

$$= \sum_{k} \alpha_k \sum_{y} p_k(y|x_1)^{1/2} \, p_k(y|x_2)^{1/2}$$

$$= \sum_{k} \alpha_k j_k(x_1, x_2) .$$

QED.

*Corollary 1*:

$$\overline{R}_0(1) \leqslant -\log_2 \sum_{k} \alpha_k 2^{-R_0^{(k)}} ,$$

and equality holds if the auxiliary channels are compatible.

*Proof*: Let $X$ achieve $\overline{R}_0(1)$. Then

$$2^{-\overline{R}_0(1)} = \overline{J}(X)$$

$$= \sum \alpha_k J_k(X) \quad \text{by Lemma 1}$$

$$\geqslant \sum \alpha_k 2^{-R_0^{(k)}}$$

On the other hand, if the channels are compatible, and if $X$ simultaneously achieves $R_0$ for each of the auxiliary channels, we have

$$\overline{J}(X) = \sum_{k} \alpha_k 2^{-R_0^{(k)}}$$

and so

$$\bar{R}_0(1) \geq -\log_2 \sum_k \alpha_k 2^{-R_0^{(k)}}, \text{ as well.}$$

QED.

*Corollary 2*:

$$\bar{R}_0(b) \leq -\frac{1}{b} \log_2 \sum_{k=1}^{K} \alpha_k 2^{-bR_0^{(k)}},$$

with equality if the channels are compatible.

*Proof*: An examination of the proof of Lemma 1 shows that if we denote the $j$-function for $\bar{\zeta}(b)$ by $\bar{j}^{(b)}$, we have

$$\bar{j}^{(b)}(\underline{x}_1, \underline{x}_2) = \sum_k \alpha_k j^{(b)}(\underline{x}_1, \underline{x}_2),$$

here $j^{(b)}$ denotes the $j$-function for $b$ parallel copies of $\zeta_k$. Arguing as in the proof of Corollary 1, we get the desired inequality. It is possible to show (see Ref. 3, p. 150, Eq. 5.6.59) that $R_0$ for $b$ parallel copies of $\zeta_k$ is exactly $b$ times the $R_0$ for $\zeta_k$, and is achieved by an input $(X_1, X_2, \ldots, X_b)$ of independent input random variables, each distributed according to the input that achieves $R_0^{(k)}$. Thus if the channels are compatible and we choose $\underline{X} = (X_1, X_2, \ldots, X_b)$, we get

$$\bar{j}^{(b)}(\underline{X}) = \sum \alpha_k 2^{-bR_0^{(k)}}$$

and so (remembering to divide by $b$),

$$\bar{R}_0(b) \geq -\frac{1}{b} \log_2 \sum_k \alpha_k 2^{-bR_0^{(k)}},$$

which combined with the opposite inequality (which is true in general) yields the desired result.

QED.

*Corollary 3*:

Let

$$R_0^{(min)} = \min_k \{R_0^{(k)} : k = 1, 2, \ldots, K\}.$$

Then

$$\lim_{b \to \infty} \bar{R}_0(b) \leq R_0^{(min)},$$

with equality if the channels are compatible.

*Proof*: This follows immediately from Corollary 2, since the probabilities $\alpha_k$ are all nonzero by assumption.

We close this section with several elementary remarks about $R_0(b)$. We have no simple formula analogous to that of Corollary 2 for $R_0(b)$, but in any given case it is not difficult to compute since, as remarked in Section II, $\bar{\zeta}(b)$ can be viewed as a DMC with alphabets $A^b, B^b$. If $R_0$ indeed measures the channel's quality, the side information present in $\bar{\zeta}(b)$ should not *decrease* $R_0$, and indeed we can prove $R_0(b) \leq \bar{R}_0(b)$. This result follows from Lemma 2, which relates the $j$-functions for $\zeta(b)$ and $\bar{\zeta}(b)$.

*Lemma 2*:

$$\bar{j}^{(b)}(x_1, x_2) \leq j^{(b)}(x_1, x_2).$$

*Proof*:

$$\bar{j}^{(b)}(x_1, x_2) = \sum_{\underline{y}, k} \bar{p}(\underline{y}, k|\underline{x}_1)^{1/2} \bar{p}(\underline{y}, k|\underline{x}_2)^{1/2}$$

$$= \sum_{\underline{y}} \sum_k [\alpha_k p_k(\underline{y}|\underline{x}_1)]^{1/2} \cdot [\alpha_k p_k(\underline{y}|\underline{x}_2)]^{1/2}$$

$$\leq \sum_{\underline{y}} \left[ \sum_k \alpha_k p_k(\underline{y}|\underline{x}_1) \cdot \sum_j \alpha_j p_j(\underline{y}|\underline{x}_2) \right]^{1/2}$$

(by Schwarz inequality)

$$= \sum_{\underline{y}} p(\underline{y}|\underline{x}_1)^{1/2} p(\underline{y}|\underline{x}_2)^{1/2} = j^{(b)}(\underline{x}_1, \underline{x}_2)$$

QED.

*Corollary 4*: $\bar{R}_0(b) \geq R_0(b)$.

*Proof*: Let $X$ achieve $R_0(b)$. Then

$$2^{-R_0(b)} = J(X) \geq \bar{J}(X) \geq 2^{-\bar{R}_0(b)}.$$

QED.

*Corllary 5*: If $R_0^{(min)} = 0$, then

$$\lim_{b\to\infty} R_0(b) = \lim_{b\to\infty} \bar{R}_0(b) = 0 .$$

*Proof*: This follows from Corollary 3 and 4.

When $R_0^{(min)} > 0$, the limits of Corollary 5 are at present unknown to us. Even if the auxiliary channels are compatible, the limit of $R_0(b)$ is as yet unknown, although we conjecture that the two limits are always the same.

## IV. A Study of Interleaving

The values of $C$ and $R_0$ for our channel models only indicate possible ranges of rates for reliable communication. To design a practical system for these channels requires a study of coding. If we choose to view $\zeta(b)$ or $\bar{\zeta}(b)$ as DMCs, the coding alphabet $A^b$ is exponentially large and the prospects of devising practical codes using such a large alphabet are rather poor. On the other hand, motivated by practical experience with real bursty channels, we might try to communicate over $\zeta(b)$ or $\bar{\zeta}(b)$ by *interleaving* codes over the basic alphabet $A$. Of course if this is done, the channels we are really coding for are $\zeta(1)$ and $\bar{\zeta}(1)$, respectively. Now one normally expects such interleaving to *decrease* channel capacity, and in Ref. 1 we showed that indeed $\zeta(b)$ is an increasing function of $b$. However, we also showed there that $\bar{\zeta}(b) = \bar{\zeta}(1)$ for all $b$, so that in the presence of side information apparently no penalty is paid if interleaving is employed. Given the results of Section III, we can now easily describe what happens to $R_0$ when interleaving is used.

Let us assume for purposes of discussion that the $K$ auxiliary channels are compatible. In that case, according to Corollary 2 in Section III, the value of $R_0$ for the channel $\bar{\zeta}(b)$ is given by

$$\bar{R}_0(b) = \frac{1}{b}\log_2 \sum_{k=1}^{K} \alpha_k 2^{-bR_0^{(k)}}$$

From this it follows that $\bar{R}_0(b)$ is a decreasing function of $b$. If we further assume that $R_0^{(min)} = 0$, we have the expression

$$\bar{R}_0(b) \sim \frac{K}{b} ,$$

where $K = -\log_2(\alpha_k)$, $k$ being the index of the auxiliary channel with $R_0 = 0$. Since we already know that under these assumptions the capacity of $\bar{\zeta}(b)$ is a constant independent of $b$, we have the peculiar situation that

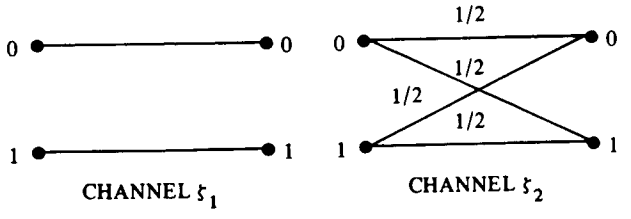
$$\lim_{b\to\infty} \frac{\bar{C}(b)}{\bar{R}_0(b)} = \infty .$$

We say peculiar because, as we mentioned in the introduction, both $C$ and $R_0$ are believed to be measures of the channel's quality, and yet as the burst length $b$ of our RFI channels increases, these measures diverge. For the channels $\zeta(b)$ without side information, the situation is if anything even more puzzling. We omit the details, but what happens in general is that $C(b)$ is a strictly increasing function of $b$, while $R_0(b)$ is a strictly decreasing function of $b$. We will comment on this apparent paradox in Section V.

Before leaving the subject of interleaving, however, we would like to describe a modification which can be used on the channel $\zeta(b)$, and which for large values of $b$ makes this channel almost as good as $\bar{\zeta}(b)$. The idea, which we first suggested in Ref. 1, is to attach to each transmitted packet a fixed "test pattern," and on the basis of the received version of the test pattern to make a statistical decision about which of the $K$ auxiliary channels was used to transmit the packet. If, say, the test pattern is of length $\log b$, then for large values of $b$ one would expect this "noise estimate" to be increasingly reliable, and yet the fraction of the transmitted letters devoted to the test pattern is quite small. What this means is that for large $b$ the channel $\zeta(b)$ behaves as if the side information were available, and so interleaving should not cause a severe degradation in performance. In the next section we will illustrate this point with a specific numerical example.

## V. A Numerical Example and Some Conclusions

We shall now illustrate our results with a specific example, the same example introduced in Ref. 1, which is indeed the "RFI channel" associated with pulse-position modulation in optical channels (Ref. 2). Here $K = 2$; the auxiliary channel $\zeta_1$ is a noiseless binary symmetric channel, and $\zeta_2$ is a "useless" BSC with transition probability 1/2.




Let us assume that $\alpha_1 = 1 - \epsilon, \alpha_2 = \epsilon$ are the probabilities that the "channel selector" chooses $\zeta_1$ and $\zeta_2$, respectively. We

present below a table of the various values of $R_0$ and $C$ for $\epsilon = 0.1$ and $b = 2^m, m = 0, 1, 2, \ldots, 10$. We now present some notes on the calculations:

1. $R_0(b), C(b)$: For purposes of computation the channel $\zeta(b)$ can be viewed as a DMC with input and output alphabets both equal to the set of binary $b$-tuples. The transition probabilities are

$$p(y|x) = (1-\epsilon) + 2^{-b}\epsilon \quad y = x$$

$$= 2^{-b}\epsilon \qquad y \neq x .$$

From this it follows that the $j$-function for $\zeta(b)$ is

$$j(x_1, x_2) = \frac{(2^b - 2)}{2^b}\epsilon + \frac{2}{2^b}\{\epsilon\,[\epsilon + 2^b(1-\epsilon)]\}^{1/2} \quad x_1 \neq x_2$$

$$= 1 \qquad x_1 = x_2 .$$

Since $\zeta(b)$ is symmetric, $R_0$ is achieved for equiprobable inputs, and indeed

$$R_0(b) = -\frac{1}{b}\log_2 (J_0),$$

where

$$J_0 = \frac{2^b - 1}{2^b}\overline{J} + \frac{2}{2^b}, \text{ where}$$

$\overline{J}$ is the value of $\overline{j}(x_1, x_2)$ for unequal $x$'s given above. We computed $C(b)$ in Ref. 3; we repeat the formula here:

$$C(b) = (1-\epsilon_b) - \frac{1}{b}[H_2(\epsilon_b) + \epsilon_b \log_2 (1 - 2^{-b})],$$

where $\epsilon_b = (1 - 2^{-b})\epsilon$, and $H_2$ is the binary entropy function.

2. $\overline{R}_0(b), \overline{C}(b)$: Since both channels $\zeta_1$ and $\zeta_2$ are symmetric, both $R_0$ and $C$ are achieved by equiprobable inputs, and the channels are compatible in the sense of this paper. Since $R_0^{(1)} = 1$, $R_0^{(2)} = 0$, we have by Corollary 2 of Section III,

$$\overline{R}_0(b) = -\frac{1}{b}\log_2 [(1-\epsilon)\,2^{-b} + \epsilon] .$$

Of course from our previous paper

$$\overline{C}(b) = 1 - \epsilon \text{ for all } b .$$

3. $\widetilde{R}_0(b), \widetilde{C}(b)$: This is a new notation and it refers to the channel $\zeta(b)$ when a specific kind of "smart" interleaving of the general kind described in Section IV is implemented. Here we use an all-zeros test pattern of length $t$ in each transmitted packet. If the received test pattern is not all zeros, the entire packet is *erased*; if it is all zeros, the packet is accepted. What this means is that after interleaving the channel $\zeta(b)$ becomes a binary symmetric erasures-and-errors channel with erasure probability $p = \epsilon(2^t - 1)/2^t$ and error probability $q = \epsilon\, 2^{-(t+1)}$. The $R_0$ for this channel is given by

$$R_0(p, q) = 1 - \log_2 [1 + p + 2\sqrt{(1-p-q)\,q}] ,$$

and so the $R_0$ for the channel $\zeta(b)$ when depth $b - t$ interleaving is employed together with this "noise detection" procedure is given by

$$\widetilde{R}_0(b) = \max_{0 \leq t \leq b} \left(1 - \frac{t}{b}\right) R_0(p, q) .$$

The maximization is over all possible test pattern lengths, and the factor $(1 - t/b)$ reflects the rate loss due to the presence of the test pattern.

Similarly the *capacity* of the above erasures-and-errors channel is given by

$$C(p, q) = (1-p)\log_2 \frac{2}{1-p}$$

$$- (1-p-q)\log_2 \frac{1}{1-p-q} - q \log_2 \frac{1}{q},$$

and so

$$\widetilde{C}(b) = \max_t \left(1 - \frac{t}{b}\right) C(p, q) .$$

We now present our table, with $\epsilon = 0.1$.

| $b$ | $R_0$ | $\overline{R}_0$ | $\widetilde{R}_0[t_{opt}]$ | $\widetilde{C}[t_{opt}]$ | $C$ | $\overline{C}$ |
|---|---|---|---|---|---|---|
| 1 | 0.47805 | 0.86250 | 0.4871 [0] | 0.7136 [0] | 0.71360 | 0.9000 |
| 2 | 0.47783 | 0.81074 | 0.4781 [0] | 0.7136 [0] | 0.74841 | 0.9000 |
| 4 | 0.45193 | 0.66952 | 0.4781 [0] | 0.7136 [0] | 0.79622 | 0.9000 |
| 8 | 0.35444 | 0.40901 | 0.4923 [1] | 0.7136 [0] | 0.84199 | 0.9000 |
| 16 | 0.20552 | 0.20761 | 0.5622 [3] | 0.7343 [1] | 0.87069 | 0.9000 |
| 32 | 0.10381 | 0.10381 | 0.6506 [5] | 0.7781 [3] | 0.88534 | 0.9000 |
| 64 | 0.05191 | 0.05191 | 0.7260 [7] | 0.8214 [4] | 0.89267 | 0.9000 |
| 128 | 0.02595 | 0.02595 | 0.7794 [9] | 0.8519 [5] | 0.89634 | 0.9000 |
| 256 | 0.01298 | 0.01298 | 0.8138 [12] | 0.8716 [6] | 0.89817 | 0.9000 |
| 512 | 0.00649 | 0.00649 | 0.8347 [14] | 0.8837 [8] | 0.89908 | 0.9000 |
| 1024 | 0.00324 | 0.00324 | 0.8469 [16] | 0.8909 [9] | 0.89954 | 0.9000 |

The numbers $R_0$, $\bar{R}_0$, $C$, $\bar{C}$ behave as expected, but the behavior of $\tilde{R}_0$, $\tilde{C}$ is rather interesting. For small values of $b$ (up to about $b = 8$), the optimal test pattern length is seen to be $t = 0$; i.e., no test pattern should be used. However, for larger $b$'s the test pattern does help, and indeed as $b \to \infty$, $\tilde{R}_0$ appears to be, and indeed is, approaching the capacity 0.900 of $\bar{\zeta}(b)$. If $R_0$ is in some sense a practical measure of the channel's quality, this indicates that for large $b$, the "smart" interleaving idea makes $\zeta(b)$ a very tractable channel for coding.

To further illustrate our ideas, we next present a table for $b = 128$, $\epsilon = 0.1$ giving the values of $R_0$ and $C$ for 5 different combinations of side information and interleaving.

| Option | $R_0$ | $C$ |
|---|---|---|
| No side information, no interleaving | 0.02595 | 0.89634 |
| No side information, "dumb" interleaving | 0.47805 | 0.71360 |
| No side information, "smart" interleaving | 0.7794 | 0.8519 |
| Side information, no interleaving | 0.02595 | 0.9000 |
| Side information, interleaving | 0.86250 | 0.9000 |

At first we found the fact that interleaving could increase $R_0$, and increase it dramatically, very puzzling. But if we take the view that $R_0$ is an inverse measure of the *delay*, rather than the *complexity*, required to achieve a given performance, the data become comprehensible. Suppose, for example, one can achieve a given bit error probability and rate with delay $D$ on the channel $\bar{\zeta}(1)$. Then exactly the *same* performance can be achieved on $\bar{\zeta}(b)$, with delay $D \cdot b$, by interleaving $b$ copies of the code used on $\zeta(1)$. Thus we would predict $\bar{R}_0(b) \geq 1/b\ \bar{R}_0(1)$, and indeed the data in the above table satisfy this inequality. Indeed, since as we showed above

$$\bar{R}_0(b) \sim \frac{-\log(\epsilon)}{b},$$

we have

$$\frac{\bar{R}_0(1)}{\bar{R}_0(b)} \sim f(\epsilon) \cdot b$$

where

$$f(\epsilon) = \frac{\log\left(\frac{2}{1+\epsilon}\right)}{\log(\epsilon^{-1})},$$

and $0 \leq f(\epsilon) \leq 1/2$. Similar but computationally messier results for $\zeta(b)$ confirm these observations.

On the basis of this numerical example and several others, we tender the following conclusion. On a general RFI channel, i.e., one that exhibits long periodic bursts of poor data quality, the best coding strategy is probably a "smart" interleaving strategy. By this we mean a strategy that uses the received data to estimate the noise severity, and passes along these estimates to the parallel decoders. In a later paper we hope to verify this conjecture by considering the performance of specific coding schemes on specific RFI channels.

## References


1. McEliece, R. J., "RFI Channels," *TDA Progress Report 42-60*, pp. 103-106, Jet Propulsion Laboratory, Pasadena, Calif., Dec. 15, 1980.
2. McEliece, R. J., and Welch, L. R., "Coding for Optical Channels with Photon Counting," *DSN Progress Report 42-52*, pp. 61-66, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1979.
3. Gallager, R. G., *Information Theory and Reliable Communication*, McGraw-Hill, New York, 1968.




C-2

N82-11291 +12



# The $R_0$-Parameter for the Gaussian Channel

R. J. McEliece and E. R. Rodemich
Communications Systems Research Section

*We first define, and then compute, the cutoff parameter $R_0$ for the additive white Gaussian channel. This important channel parameter seems not to have been previously computed for this important channel model, except in the case when the input is restricted to be binary.*

## I. Introduction

The computational cutoff parameter $R_0$ has lately begun to assume an important significance in communication systems. It appears in many situations to measure a given channel's "quality" in a way that is superior, from a practical standpoint, even to the capacity of the channel. It is our object in this paper to compute the $R_0$ parameter for the important additive white Gaussian channel (AWGC), which is, for example, the appropriate channel model for deep-space communication. This parameter is well known, when the channel input is restricted to two levels:

$$R_0 = \log_2 \frac{2}{1 + e^{-E/N_0}} ,$$

where $E/N_0$ is the signal-to-noise ratio (Ref. 7, Eq. 5-56). However, $R_0$ seems not to have computed for the AWGC when there has been no restriction on the number of channel inputs.

In the next section we shall give what we feel is the correct definition of $R_0$ for the AWGC, but also discuss the merits of another candidate, the quantity $R_0^*$ discussed by Shannon (Ref. 6). In Section III we shall prove that the input distribution achieving $R_0$ is always concentrated at a finite number of points. Finally in Section IV we will give some numerical values of $R_0$.

## II. A Definition of $R_0$ for the Gaussian Channel

The additive white Gaussian channel (AWGC) can be described as follows (Ref. 4, Chapter 4): If $\ldots, X_{-1}, X_0, X_1, X_2, \ldots$ denotes the input sequence and $\ldots, Y_{-1}, Y_0, Y_1, Y_2, \ldots$ the output sequence, we have $Y_k = X_k + Z_k$, where $\{Z_k\}$ is a sequence of independent, identically distributed (iid), mean zero, variance $N_0/2$ random variables. The input sequence is constrained in "average energy" by requiring that $\{X_k\}$ be iid, and $E(X_k^2) = A$. It is convenient, and involves no real loss in generality, to use the normalization $N_0/2 = 1$, and we shall do so. Our goal in this section is to give a defensible definition for the cutoff parameter $R_0$ for this channel.

We first assume that the input distribution is given, viz., that $F(x)$ is the cumulative distribution function for each of

the random variables $X_k$. Then according to Eq. (5.53b) in Ref. 7, the cutoff parameter *with respect to* $F$ is given by

$$R_0(F) = -\log_2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x-y)^2/8}\, dF(x)\, dF(y) \tag{1}$$

Since the input distribution must satisfy the energy constraint

$$\int_{-\infty}^{\infty} x^2\, dF(x) = \mathrm{A}\,, \tag{2}$$

it is reasonable to define $R_0$ for the Gaussian channel by

$$R_0 = \sup\,\{R_0(F) : F \text{ satisfies (2)}\} \tag{3}$$

Indeed this is the definition we take, and Sections III and IV describe the solution to this mathematical optimization problem. However, one cannot assert that the quantity so defined is "$R_0$" for the AWGC, without discussing a competing number discussed by Shannon (Ref. 6) in 1959. This we now do.

There are numerous conjectures about the practical significance of the cutoff parameter for a given channel (see for example, Ref. 3 or 5), but there are also two provable theorems about $R_0$. The first theorem is that for any rate $R$ less than $R_0$, then there exists a code of length $n$ for which the error probability is bounded by $P[\mathscr{E}] \leqslant 2^{-n[R_0-R]}$ (see for example, Ref. 2, Chapter 5 for the discrete channel case of this theorem.) It is true, and not hard to prove, that this theorem is true for the AWGC with $R_0$ defined as in (3). However, in the paper cited above (Ref. 6), Shannon proved that this theorem remains true for what turns out to be a larger number; viz.,

$$R_0^* = \frac{\log_2 e}{2}\left[1 + \frac{A}{2} - \sqrt{1 + \frac{A^2}{4}}\right]$$

$$+ \frac{1}{2}\log_2\left[\frac{1}{2}\left(1 + \sqrt{1 + \frac{A^2}{4}}\right)\right] \tag{4}$$

Thus although our definition (3) is perhaps plausible, if one defines $R_0$ to be the largest possible intercept of a line of slope $-1$ which supports the reliability exponent $E(R)$ for the given channel, then the quantity (4) is the correct definition. However, there is another possible definition, which is derived from a theorem with communications significance, which favors our definition. The theorem deals with the expected number of computations needed for a sequential decoding algorithm.

In a celebrated paper on sequential decoding, Berlekamp and Jacobs (Ref. 1) showed that there exists a certain rate, called $R_{\text{comp}}$, which represents the supremum of all rates $R$ such that the average number of computations made by a sequential decoder operating on a code of rate $R$ remains bounded. They showed that $E_0(1) \leqslant R_{\text{comp}} \leqslant \hat{E}_0(1)$, where $E_0(p)$ is a certain function which depends on the channel statistics, and $\hat{E}_0(p)$ is the convex $\cap$ hull of $E_0(p)$. For the Gaussian channel, if the code being used must satisfy the average energy constraint $E(X^2) \leqslant A$, it is easy to show that the parameter $E_0(1)$ is precisely our definition (3) of $R_0$. Now for "ordinary" channels, the function $E_0$ is already convex, and so $R_{\text{comp}} = E_0(1)$. And we conjecture that this holds for the Gaussian channel too, but have not yet been able to prove it. If our conjecture proves to be correct, then our definition (3) will have been proved to be the value of "$R_{\text{comp}}$" for the AWGC, and incidentally will have been shown to be *strictly less* than the "$R_0$" for this channel.

In the next two sections, we will discuss the computation of $R_0$, as defined by (3).

## III. A Characterization of $R_0$

We recall that $R_0$ is defined as follows. If $Q$ is defined as the value of the program:

minimize: $$Q = \iint K(x,y)\, dF(x)\, dF(y),$$

$$\left(K(x,y) = e^{-\frac{1}{8}(x-y)^2}\right)$$

subject to: $$\int dF(x) = 1, \int x^2\, dF(x) = A,$$

where $F$ is a distribution function. Then $R_0 = -\log_2 Q$. In this section we will show that the optimizing distribution $F$ is discrete, i.e., has mass at only a finite number of points.

We shall use the calculus of variations to find necessary conditions that must be satisfied by an external distribution $F$.

If we use Lagrange multipliers $\mu$ and $\lambda$ for the two side conditions, and apply a variation $\delta F$ to $F$, the variation of the Lagrangian function

$$L = \frac{1}{2} Q + \mu \int x^2 \, dF - \lambda \int dF$$

is given by

$$\delta L = \frac{1}{2} \delta Q + \mu \delta \int x^2 \, dF - \lambda \delta \int dF$$

$$= \int \left\{ \int K(x,y) \, dF + \mu x^2 - \lambda \right\} d(\delta F)$$

Since we are looking for a minimum, $\delta L$ must be $\geqslant 0$ for all admissible variations $\delta F$. Because the Lagrange multipliers account for the two integral side conditions, the only restriction on $\delta F$ is that $F + \delta F$ must be an increasing function. Hence $d(\delta F)$ can be concentrated at or near one point, and must be nonnegative if this point is not in support of $dF(x)$. Hence we must have

$$\phi(x) = \int K(x,y) \, dF(y) + \mu x^2 - \lambda \geqslant 0, \text{ all } x, \qquad (5)$$

$$\phi(x) = 0 \qquad (6)$$

at all points of support of $dF$.

We note that $0 \leqslant \int K(x,y) \, dF(y) \leqslant 1$ for all $x$, and this integral approaches zero as $x \to \pm\infty$. Thus if we divide (5) by $x^2$ and let $x \to \infty$, we see that $\mu \geqslant 0$. Also, if $x$ is in the support of $dF$ the integral in (5) is positive, and by (6), $\phi(x) = 0$. Hence $\lambda > 0$. If $\mu = 0$, then as $x \to \infty$, we would have $\phi(x) \to -\lambda$, which contradicts (5). Hence in (5) we must have

$$\mu > 0, \lambda > 0 \qquad (7)$$

Now for any $x$ in the support of $dF$, we see from (6) that

$$\mu x^2 = \lambda - \int K(x,y) \, dF(y) \leqslant \lambda,$$

$$x^2 \leqslant \lambda / \mu$$

Thus the mass of the distribution $F$ all lies in a bounded interval. Thus $\phi(x)$ is analytic, and so can have only a finite number of zeroes on a bonded interval 0. But by (6) this means that $F$ has only a finite number of support points. This is what we set out to prove.

In the next section we shall give some numerical values of the function $R_0$.

## IV. Some Numerical Results

Once it is known that the optimizing distribution $F$ is concentrated at a finite number of points it is possible to program a computer to calculate $R_0$. The number of points needed is an increasing function of the parameter $A$, which we denote by $n(A)$. It turns out, for example, that $n(A) = 2$ for $A \leqslant 2.38586$.

In the table below we list for $k = 2, 3, 4$ the largest value of $A$ for which a $k$-point distribution is optimal:

| $k$ | $A(k)$ |
|---|---|
| 2 | 2.386 |
| 3 | 5.292 |
| 4 | 8.6913 |

In the next table we list the actual value for $R_0$, as a function of $A$. For reference we also tabulate Shannon's function $R_0^*$ as given by (4).

| $A$ | $R_0$ | $R_0^*$ |
|---|---|---|
| 0.0 | 0 | 0 |
| 0.5 | 0.1691 | 0.1692 |
| 1.0 | 0.3161 | 0.3169 |
| 1.5 | 0.4419 | 0.4456 |
| 2.0 | 0.5481 | 0.5583 |
| 2.5 | 0.6367 | 0.6578 |
| 3.0 | 0.7149 | 0.7464 |
| 3.5 | 0.7861 | 0.8260 |
| 4.0 | 0.8512 | 0.8982 |
| 4.5 | 0.9110 | 0.9461 |
| 5.0 | 0.9659 | 1.0247 |
| 5.5 | 1.0165 | 1.0808 |

| $A$ | $R_0$ | $R_0^*$ |
|---|---|---|
| 6.0 | 1.0638 | 1.1330 |
| 6.5 | 1.1082 | 1.1817 |
| 7.0 | 1.1502 | 1.2274 |
| 7.5 | 1.1898 | 1.2704 |
| 8.0 | 1.2274 | 1.3111 |
| 8.5 | 1.2630 | 1.3496 |
| 9.0 | 1.2969 | 1.3861 |
| 9.5 | 1.3292 | 1.4210 |
| 10.0 | 1.3602 | 1.4542 |
| 10.5 | 1.3899 | 1.4860 |

These numbers should be compared to Fig. (5.18) in Ref. 7, where "$R_0$" is computed using a set of equally spaced points, each less in absolute value than $\sqrt{A}$. For the optimal distribution the points are not in general equally spaced, nor are they all less than $\sqrt{A}$.

# References


1. Berlekamp, E. R., and Jacobs, I. M., "A Lower Bound to the Distribution of Computation for Sequential Decoding," *IEEE Trans. Inform. Theory*, IT-13, pp. 167-174, 1967.

2. Gallager, R. G., *Information Theory and Reliable Communication*, McGraw-Hill, New York, 1968.

3. Massey, J. L., "Capacity, Cutoff Rate, and Coding for a Direct Detection Optical Channel," *TDA Progress Report 42-60*, pp. 68-76, Jet Propulsion Laboratory, Pasadena, Calif., Dec. 5, 1980.

4. McEliece, R. J., *The Theory of Information Coding*, Addison-Wesley, Reading, Mass., 1977.

5. McEliece, R. J., Rodemich, E. R., and Rubin, A. L., "The Practical Limits of Photon Communications, *DSN Progress Report 42-55*, pp. 63-67, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1980.

6. Shannon, C. E., "Probability of Error for Optimal Codes in a Guassian Channel," *Bell System Tech., J.*, 38, pp. 611-656, 1959.

7. Wozencraft, J. M., and Jacobs, I. M., *Principles of Communication Engineering*, Wiley, New York, 1965.

N82-11292

TDA Progress Report 42-65

July and August 1981

# Design of Throat Section of Conical Corrugated Horns

A. D. Olver[1]

Queen Mary College, London, England

*A simple graphical method is presented for designing the junction between a conical corrugated horn and a smooth-wall circular waveguide. The method gives both an insight into the field behavior along the horn and an indication of the reflection coefficient at the horn-waveguide junction. It has been applied to the design of the DSN X-S band feed-horn and an improved corrugated throat section.*

## I. Introduction

Conical corrugated horns are usually connected to a section of smooth-wall circular waveguide as the means of transferring power between the radiating horn and a transmission line. The junction, Fig. 1, must be correctly designed if optimum performance is to be achieved. Any mismatch between the modes in the circular waveguide and the modes in the horn will lead to a nonzero reflection coefficient. In addition, any sharp changes in impedance along the throat section of the horn can give rise to mode conversion to a higher order mode. Once excited, this unwanted mode may radiate and cause increased cross-polar power and/or reduced efficiency. The effects can be more complicated if operation over an appreciable frequency band is desired.

This note sets out a graphical method of designing the throat section that gives an estimate of the reflection loss that can be expected from a conical corrugated horn.

[1]Consultant to JPL Sections 333 and 336. Dr. Olver's complete address is: Department of Electrical Engineering, Queen Mary College, Mile End Road, London E1, England.

## II. Simple Model

The standard method of matching into a corrugated horn has been to make the first slot $\lambda/2$ deep at the highest frequency of operation, and then taper over a few slots to $\lambda/4$ depth slots. This procedure is adequate for single-frequency or narrow-band systems, but is too simple for a wideband system, as Fig. 2 indicates. This shows the change in guide wavelength against normalized inner radius ($r_1/\lambda$) for a smooth-wall circular waveguide and a corrugated waveguide with slot depth chosen to match at $r_1/\lambda = 0.5$. The horizontal scale should be interpreted as changing frequency, not changing radius. The variation with frequency is different for the two waveguides, so a good match can be achieved at only one frequency. The reflection coefficient at the junction of the two waveguides is given simply by (Ref. 1)

$$\rho = \left| \frac{\lambda_{g2} - \lambda_{g1}}{\lambda_{g2} + \lambda_{g1}} \right|$$

where $\lambda_{g1}$ and $\lambda_{g2}$ are the guide wavelengths in the two waveguides. Knowledge of these guide wavelengths will give an immediate estimate of the return loss.

The above discussion referred to the junction between two waveguides whereas in reality we have the junction between one waveguide and a conical horn. However, the horn can be approximated as a series of constant diameter waveguides. The guide wavelength for each corrugated waveguide section can be computed and plotted on a composite graph to give an indication of the way in which the field changes as it propagates along the horn. The guide wavelength will be high at the junction and decrease toward the free-space wavelength at the horn aperture. Sharp changes in guide wavelength along the horn are indicative of a change in impedance that can cause a higher order mode to be excited. Thus not only should the guide wavelength on either side of the smooth-wall-to-corrugated waveguide junction be equal, but the change in guide wavelength with distance along the horn (i.e., $\partial\lambda_g/\partial z$) should be as low as possible.

## III. Design Curves

The information needed to design or analyze the throat section of a corrugated horn can be conveniently displayed on a single graph. Curves showing the normalized guide wavelength against normalized slot depth for various normalized inner radii are shown in Fig. 3. These have been computed by solving the $HE_{11}$ mode propagation equation for a corrugated waveguide using a model that includes the effects of space harmonics (Ref. 2). The space harmonic representation enables the exact slot geometry to be studied. In general, this is not strictly necessary for calculation of the propagation coefficient and the simpler surface impedance representation should be adequate (the space harmonic model is needed for the study of attenuation or cross-polar radiation characteristics).

The use of the curves is indicated in Fig. 4. This shows a trajectory for a corrugated horn obtained by plotting the inner radius and slot depth for each corrugation along the horn. The guide wavelength for the smooth-wall waveguide occurs on the slot depth/$\lambda$ = 0.5 line, point A. The first slot (from the junction) in the corrugated section gives point B. The second slot point C, and so on. When the tapering of the slot depths has stopped and the slot depth is constant, point D is reached and from then on the guide wavelength will decrease along the line E.

A good match at the design frequency would have points A and B coincident and a smooth change along the horn. In the example shown in Fig. 4, there would be a mismatch at the junction giving a return loss of about -34 dB. The mode conversion along the horn depends on whether a higher order mode is propagating at the relevant radius. The most troublesome mode is the $HE_{12}$ mode because this is excited by changes in waveguide cross section. The cut-off line for the $HE_{12}$ mode is shown in Fig. 4, and in this case the uneven curve occurs at a point in the horn where the $HE_{12}$ mode cannot propagate.

The guide wavelength curves can be used to design a throat section. An ideal smooth trajectory would be drawn on Fig. 3 between the chosen smooth-wall waveguide value and the chosen final slot-depth value. Points would then be chosen along the curve at which slots could be placed. Care must be taken to ensure that the inner radius changes smoothly with distance along the horn.

The situation becomes more complicated when a band of frequency operation is desired. Then the trajectories at the upper and lower operating frequencies must be placed on the guide wavelength curves. Some compromise will be needed as a perfect match is possible at only one frequency. It is generally better to choose the design frequency for optimum match to be near the middle of the band, rather than at the upper frequency. This is because the guide wavelength changes more rapidly at low frequencies; see Fig. 2. The fact that this makes the first slot depth greater than half a wavelength at the upper frequency is not desirable, but may be a necessary compromise to obtain the lowest return loss over the complete operating band.

A further complication occurs if a tracking mode is present at the junction. It will then be necessary to repeat the analysis for the tracking mode and try to obtain a design that is a compromise between the matching of the signal mode and matching of the tracking mode.

## IV. Application to X-Band Throat Section of DSN X-S-Band Feed Horn

The current design of the throat section has been analyzed using the method described in the previous section. Figure 5 shows the guide wavelength curves at 7.15 GHz, 7.90 GHz, and 8.60 GHz. Clearly the best match occurs at a frequency of about 8.1 GHz with a substantial deterioration at the lower frequencies. This is because the smooth-wall waveguide is operating near to its cut-off frequency where the guide wavelength changes rapidly. The trajectories are smooth and little trouble can be expected from higher order mode excitation.

The values of the reflection coefficient estimated from Fig. 4 are compared with the measured values in Fig. 6. The general agreement is reasonable at the band edge, but not in the middle of the frequency range; the disagreement may be due to the external matching network. However, the relatively simple model gives an indication of the level of performance of a circular waveguide-to-corrugated waveguide junction.

## V. Redesign of Throat Section

It is possible to improve the overall impedance match of the throat section by shifting the effective operating band of the first few slots to a higher frequency. This improves the performance because the guide wavelength changes more rapidly at lower frequencies compared to higher frequencies, so the return loss at the lower frequency can be reduced without a corresponding increase at the upper frequency. The criteria is applied that the theoretical return loss at 8.6 GHz should not be greater than -30 dB; this then fixes the inner radius of the first slot. Next a smooth trajectory is drawn on the guide wavelength curves, Fig. 7. The exact location of the slots along the trajectory is obtained by sketching a smooth geometric profile of the inner radius of the corrugations, Fig. 8. Some juggling of the parameters is needed at this stage to maintain a smooth geometric and smooth electromagnetic profile. The profile shown in Fig. 7 does not fit any analytical function, unlike the original design, which is an arc of a circle. The inner radii and slot depths are given in Table 1.

The redesigned profile is chosen to have the same axial length as the original. The length is not actually necessary, as Fig. 7 shows; most of the transformation takes place in the first seven slots. The major difference between the original design and the new design is the inner radius of the first slot. In the original design this is the same as the radius of the smooth-wall waveguide (17.4 mm). Increasing the radius to 17.8 mm gives the improved match at the junction. This feature is often useful in corrugated horn design to give added flexibility. Sometimes a short section of smooth-wall conical guide can be inserted between the smooth-wall circular waveguide and the first slot as a way of introducing a larger initial radius for the corrugated horn.

## References


1. Dragone, C., "Reflection, Transmission and Mode Conversion in a Corrugated Feed." *Bell Systems Technical Journal*, Vol. 56, July 1977, p. 835-867.

2. Clarricoats, P. J. B., Olver, A. D., and Chong, S. L., "Attenuation in Corrugated Circular Waveguides. Pt I Theory." *Proc IEE* (London), Vol. 122, 1975, p. 1173.

**Table 1. Redesigned Throat Section**

| Slot no. | Inner radius, mm | Slot depth, mm |
|---|---|---|
| 1 | 17.8 | 18.5 |
| 2 | 18.8 | 16.9 |
| 3 | 19.8 | 15.2 |
| 4 | 20.9 | 14.2 |
| 5 | 22.2 | 13.3 |
| 6 | 23.5 | 12.8 |
| 7 | 25.0 | 12.3 |
| 8 | 26.8 | 12.1 |
| 9 | 28.7 | 11.9 |
| 10 | 31.0 | 11.7 |
| 11 | 33.5 | 11.6 |
| 12 | 36.3 | 11.5 |
| 13 | 39.6 | 11.4 |

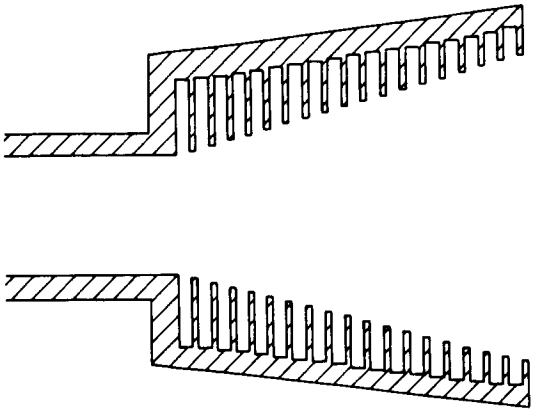


**Fig. 1. The junction of corrugated horn**

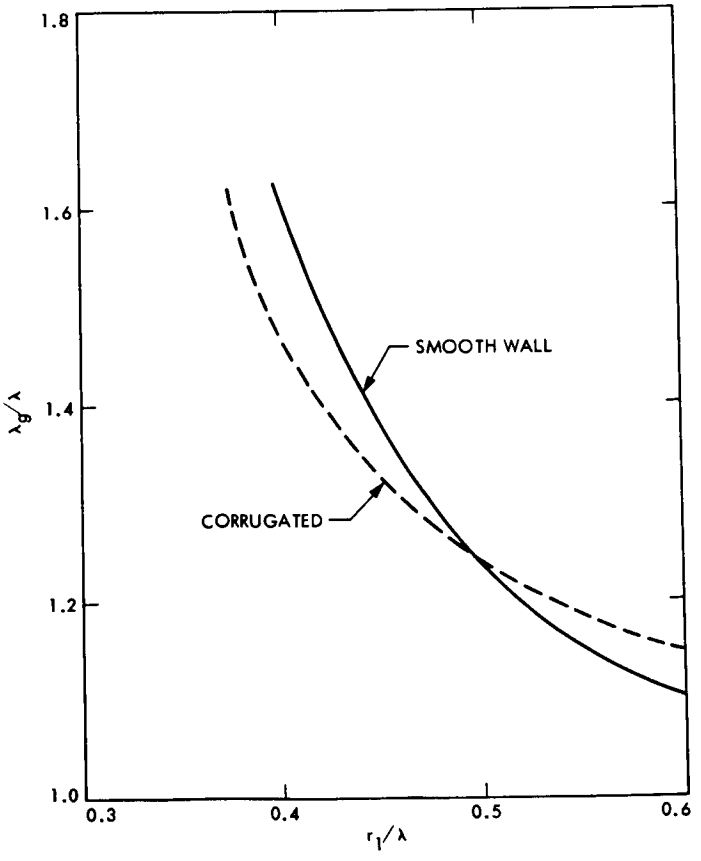



**Fig. 2. Normalized guide wavelength of smooth-wall and corrugated waveguides**

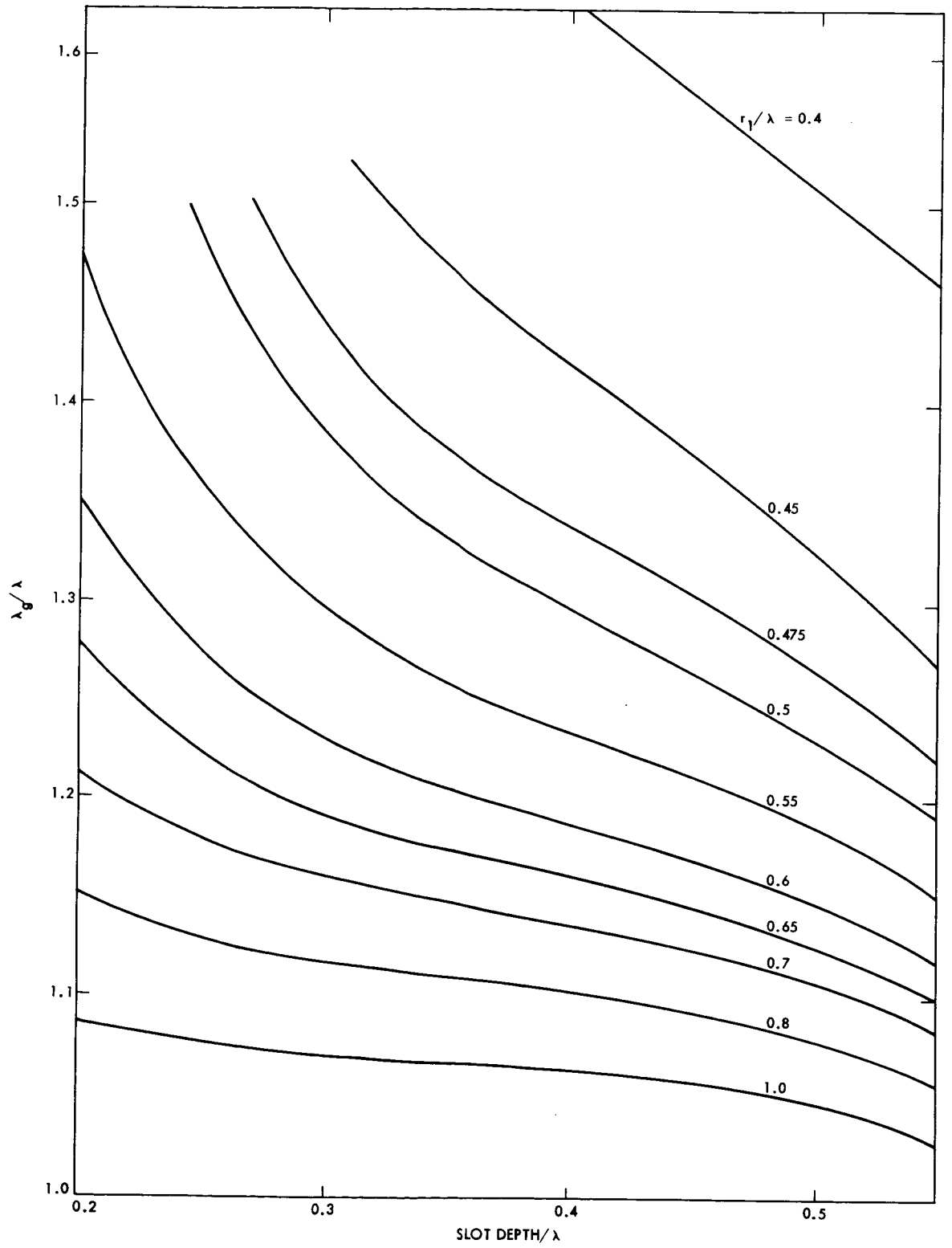



**Fig. 3. Design curves for junction of corrugated horns**

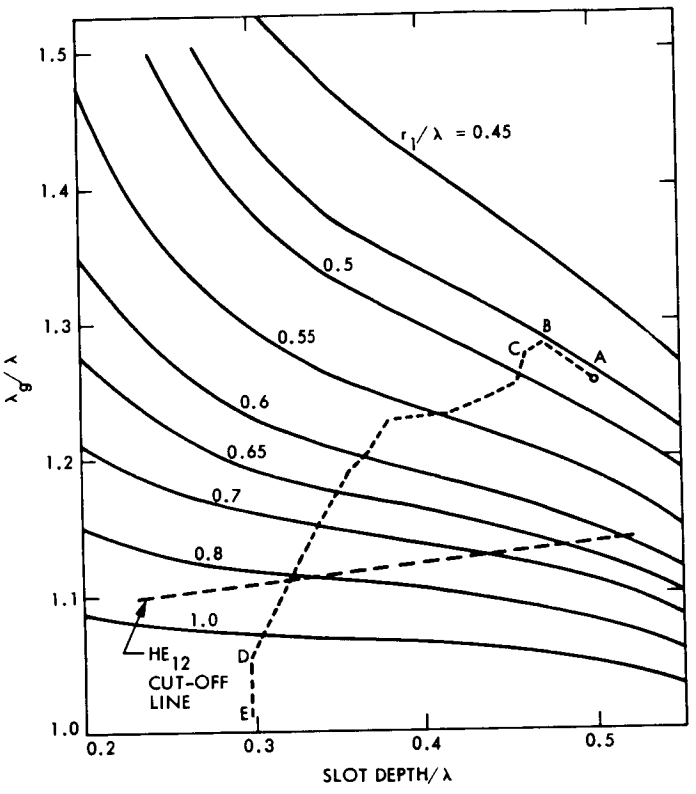



Fig. 4. Trajectory of typical junction

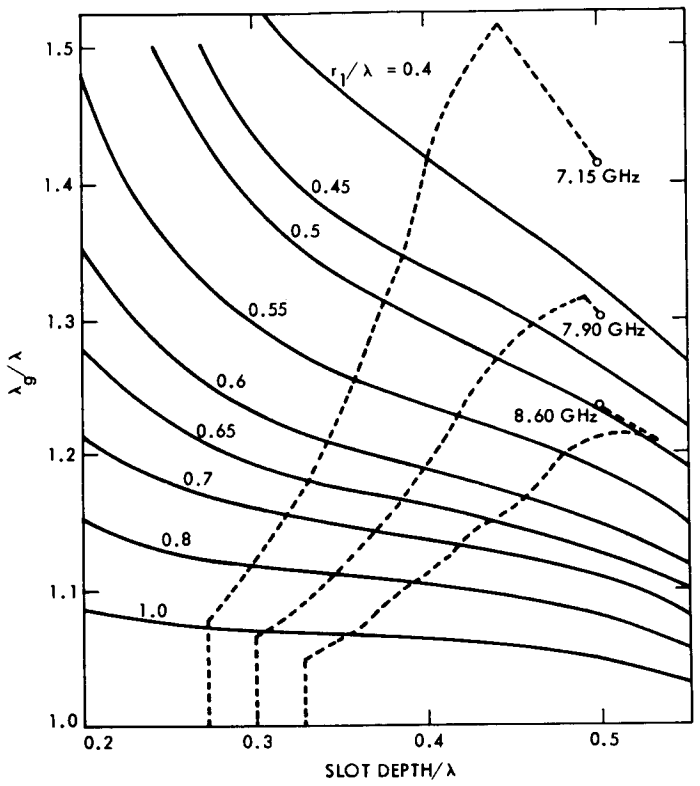



Fig. 5. Guide wavelength curves for DSN X-S-band feedhorn

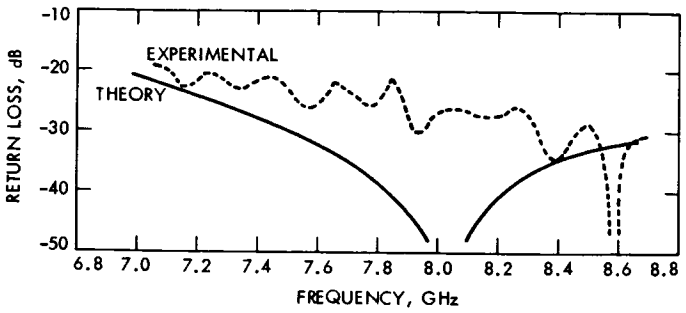



Fig. 6. Predicted and measured return loss of DSN feedhorn

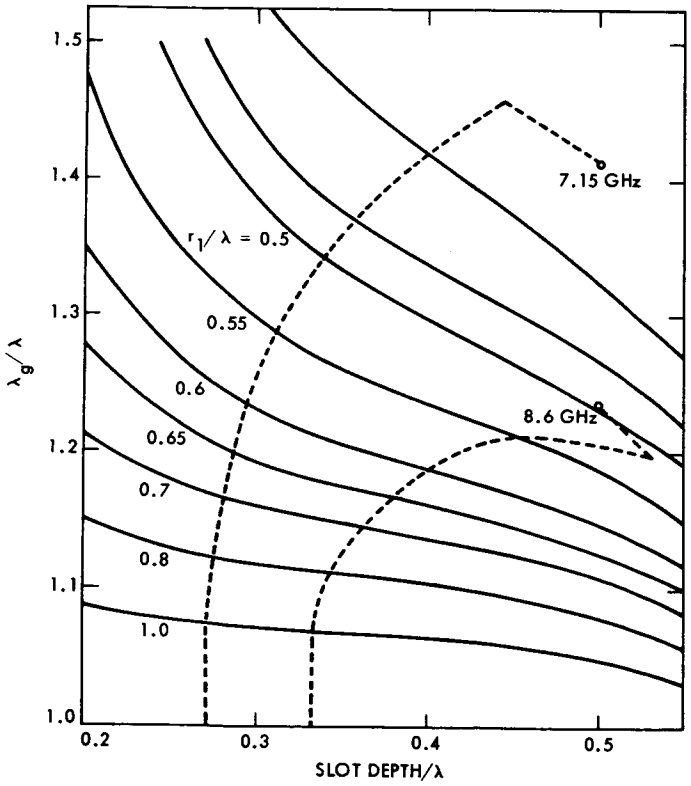



Fig. 7. Redesigned trajectories of DSN X-S-band feedhorn

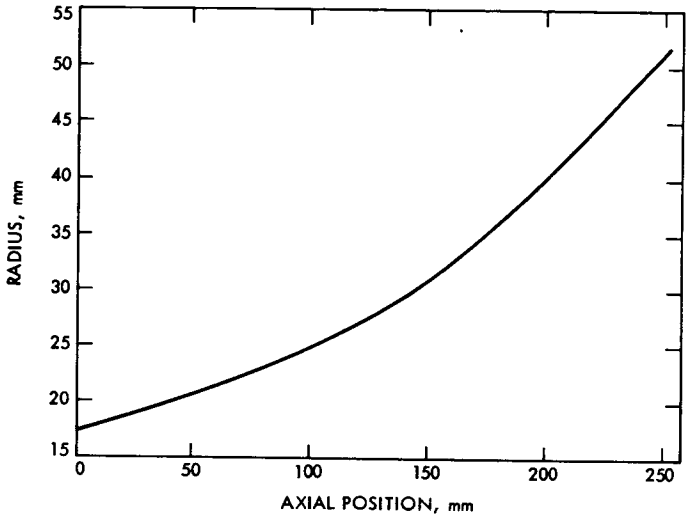



Fig. 8. Geometric profile corresponding to Fig. 7

N82-11293



# Cryogenic Filters for RFI Protection

J. J. Bautista and S. M. Petty
Radio Frequency and Microwave Subsystems Section

*The increased bandwidth and sensitivity of the DSN maser-based receiver systems along with the increase in worldwide microwave spectrum usage have dictated the need for employing additional measures to protect these systems from RFI (radio frequency interference). Both in-band and out-of-band microwave signals at the input of the Deep Space Network (DSN) traveling wave masers (TWM) can adversely affect the maser performance in a variety of ways. Filters fabricated from superconducting materials operating below their superconducting transition temperature ($T_c$) possess the most potential for providing the necessary RFI protection without degrading the system performance.*

## I. Introduction

To meet the increased demands of more challenging spacecraft missions and at the same time provide for more deep space station users such as Very Long Baseline Interferometry (VLBI) and Radio Frequency Interference (RFI) Surveillance, the bandwidth as well as the sensitivity of present masers have been increased over previous designs (Ref. 1). Unfortunately, this increase in performance and versatility has also yielded a greater maser susceptibility to RFI.

The effect of RFI on the maser performance is primarily determined by the level of the RFI, its frequency and how its frequency relates with the frequency of the maser pump source(s). Both in-band and out-of-band RFI (CW and pulse) can result in either gain loss or spurious output signals or both in the maser signal bandpass (Ref. 2).

The long-term goal of the Cryogenic Filter Project at JPL is the development and testing of cryogenically cooled filters which will protect the DSN S-band and X-band masers from in-band and out-of-band RFI without degrading the maser sensitivity, amplitude, phase, or group delay stability, or adding dispersion within the bandpass. The project has been divided into the following six sequential steps:

(1) Measure the effects of CW and pulse RFI on the Block II X-band TWM.

(2) Research the probable levels and frequencies of known sources of RFI.

(3) Develop a fixed, cryogenically cooled filter to protect the Block II X-band TWM from out-of-band RFI.

(4) Measure the effects of CW and pulse RFI on the Block III/IV S-band TWM.

(5) Develop a fixed cryogenically cooled filter to protect the Block III/IV S-band TWM's from out-of-band RFI.

(6) Develop a variable-frequency cryogenic filter which can provide protection from in-band RFI for the above TWM's.

This introductory report will discuss the susceptibility of the Block II TWM (Ref. 3) to out-of-band RFI and the design approaches being investigated for the fixed, cryogenically cooled filter for this TWM.

## II. Block II X-Band Maser Susceptibility to CW RFI

### A. Maser Gain Reduction

In general, from 7 GHz to at least 40 GHz (our upper limit of measurement) levels greater than 0 dBm can cause measurable maser gain reduction (greater than 0.1 dB) in the maser passband. In addition to the above general susceptibility, there are three RFI frequency bands (7.6-9.1, 15.3-15.9 and 34.3-35.5 GHz) where levels below 0 dBm will reduce the maser gain. The minimum RFI power level in each of these frequency bands which results in a 1-dB maser gain reduction in the signal passband (8.4-8.5 GHz) is shown in Fig. 1. RFI at these frequencies induces transitions between the paramagnetic energy states, resulting in a gain loss due to the degradation of the inverted population condition.

### B. Mixing Effects

At the idler frequencies, masers are particularly sensitive to RFI. An idler frequency $f_i$ is defined according to the following equation $f_i = f_p - f_s$, where $f_p$ is a maser pump frequency (or sum of pump frequencies) and $f_s$ is the signal frequency (Ref. 4). The Block II X-band TWM has three idler frequency ranges which are listed in Table 1.

An RFI signal at any one of these idler frequencies can mix with the maser pump energy, resulting in an output in the operating maser passband (8400-8500 MHz). (It should be noted that the 4th, 5th and 17th transmitter harmonics of the DSN S-band transmitters occur in the X-band TWM passband and very near two of the three idler frequency bands (Ref. 5).) Since the pump source is composed of two free-running gunn oscillators that are frequency-modulated at the rate of 100 kHz, the mixed output will also be modulated at the same rate and might appear as broadband noise on a spectrum analyzer. This mixing will occur with a conversion loss of 90 dBm minimum when referenced to the maser input (Ref. 2). For example, a −60 dBm signal at 10.76 GHz mixed with the 19.2 GHz pump source results in a −150 dBm maximum signal at 8.44 GHz. After amplification by the maser (45 dB nominal) the final output is an FM-modulated (100 kHz) signal at 8.44 GHz with signal level of −105 dBm maximum.

Although the mixing discussed above is strictly a function of the electron spin resonances of the maser active material, nonlinear components in the maser such as metal-oxide-metal junctions and YIG isolators can also cause mixing with the maser pump frequencies. The conversion losses for these processes are expected to be much higher than those due to electron spin resonances.

Measurement of these mixing effects in a Block II X-band maser is presently underway.

## III. Filter Design Goals

In view of the variety of ways in which out-of-band RFI can adversely affect maser performance, a filter is needed which precedes the input of the maser amplifier to provide adequate protection from these interferences. Using Fig. 1, the results of Section II, known characteristics of the Block II TWM and practical design considerations, a fixed bandpass preselector is proposed for the X-band Block II TWM. We present the insertion loss and rejection design goals for this filter in Fig. 2. The response limits are contained within the shaded region of this figure.

The proposed filter is designed to adhere to the following guidelines and requirements:

(1) An insertion loss of no greater than 0.1 dB at the bandpass is necessary so that when the filter is cooled down to 4.5 K the filter noise temperature contribution is below 0.2 K.

(2) The rejection characteristics surrounding the maser operating passband must be as steep as possible without adding significant dispersion within the passband.

(3) A 50-dB rejection "floor" from 0-10 GHz gives more than adequate protection in this frequency range ($>>$ 0 dBm RFI) and is easily realizable.

(4) An increased rejection of 100 dB from 10-40 GHz is needed to provide protection from RFI in the idler frequency bands. 100-dB rejection will insure that idler band power levels of less than 0 dBm will result in spurious maser output levels less than the minimum detectable signal level.

## IV. Approach

Because of the low loss characteristics of superconductors operating below $T_c$, their critical temperature, it has been suggested and demonstrated that filters fabricated from these materials would in principle have the ideal characteristics for our applications. In fact, the only known way of obtaining a filter insertion loss of 0.1 dB is by using superconducting

materials because of their extremely low surface impedance when cooled below $T_c$.

The surface impedances of some superconductors have values 5 to 6 orders of magnitude smaller than the best normal conductors at 1-7 Ghz (Refs. 6, 7). In addition, certain A15 compounds like NbTi, $Nb_3$ and $Nb_3Sn$ possess relatively high $T_c$'s (9.5, 9.2, and 18.0 K) and relatively good mechanical properties, allowing repeated thermal cycling without any deleterious effects (Ref. 8). Depending on physical size constraints and engineering difficulties, such filters could be microstrip, air line or cavity construction.

For example, where size is a major constraint and/or frequency ranges dictate smaller dimensions, microstrip filters could be fabricated from Nb or $Nb_3Sn$, utilizing thin film deposition techniques. The thin films would be deposited by either magnetron sputtering or electron-beam evaporation techniques onto a sapphire substrate. This latter process is still under development and Stanford University (in collaboration with JPL) has started a program to evaluate the RF characteristics of such structures.

For masers with less stringent physical constraints and/or higher frequencies, filters could be fabricated from bulk superconductors like NbTi. Currently, we are evaluating the performance of several interdigital X-band filters machined from NbTi. The preliminary results from these tests are very encouraging and will be reported in a future article. This is the approach that has been chosen for the Block II X-band fixed filter.

## V. Plans

Our plans are to develop, test and build a prototype fixed X-band filter for implementation in the X-band Block II-A TWM by the middle of 1982. The prototype fixed S-band filter will be ready for implementation in the Block III and IV S-band masers by early 1983.

In FY 1983 work will begin on tunable cryogenic filters which will protect the maser receiver system from in-band RFI.

## References


1. Clauss, R., Weibe, E., and Quinn, R., "Low Noise Receivers: Microwave Maser Development," in *The Deep Space Network Progress Report*, Technical Report 32-1526, Vol. XI, pp. 71-80, Jet Propulsion Laboratory, Pasadena, Calif.
2. Clauss, R. C., "X- and K- Band Maser Development. Effects of Interfering Signals," in *The Deep Space Network Progress Report 42-42, September and October 1977*, pp.85-87, Jet Propulsion Laboratory, Pasadena, Calif., Dec. 15, 1977.
3. Trowbridge, D. L., "X-Band Traveling Wave Maser Amplifier," in *The Deep Space Network Progress Report 42-28*, California Institute of Technology, Jet Propulsion Laboratory, Pasadena, Calif.
4. Siegman, A. E., *Microwave Solid State Masers*, Chapter 6, McGraw-Hill Book Company, New York, 1964.
5. Hartop, R., and Bathker, D., private communication.
6. Hartwig, W. H., "Superconducting Resonators and Devices," *Proceedings of the IEEE*, Vol. 61. No. 1, Jan. 1973.
7. Kneisel, P., Stoltz, J. and Halbritter, J., "Measurements of Superconducting $Nb_3Sn$ Cavities in the GHz Range," *IEEE Transactions on Magnetics*, Vol. Mag-15, No. 1, January 1979.
8. Beasley, M. R., "Advanced Superconducting Materials for Electronic Applications," *IEEE Transactions on Electron Devices*, Vol. Ed-27, No. 10, Oct. 1980.

**Table 1. Summary of RFI susceptibility: X-band Block II TWM**

| | |
|---|---|
| Signal frequency range | 8.40 – 8.50 GHz |
| Maser net gain | 46 dB max<br>43 dB min |
| Signal power input level for –1 dB gain compression | –84 dBm |
| Pump source frequency range[a] (Low-frequency (3-4) pump; includes FM modulation) | 19.11 – 19.32 GHz |
| Pump source frequency range[a] (High-frequency (1-3) pump; includes FM modulation) | 23.91 – 24.25 GHz |
| Idler frequency ranges | |
| $f_{p_{3\text{-}4}} - f_s$ | 10.61 – 10.92 GHz |
| $f_{p_{1\text{-}3}} - f_s$ | 15.41 – 15.75 GHz |
| $f_{p_{1\text{-}3}} + f_{p_{3\text{-}4}} - f_s$ | 34.53 – 35.18 GHz |
| Mixing possibilities due to nonlinearities in maser structure | |
| $f_{p_{3\text{-}4}} + f_s$ | 27.51 – 27.82 GHz |
| $f_{p_{1\text{-}3}} + f_s$ | 32.31 – 32.75 GHz |
| Maser gain compression due to out-of-band RFI | See Fig. 1 |

[a]These are the minimum and maximum pump frequencies observed on six operating masers.

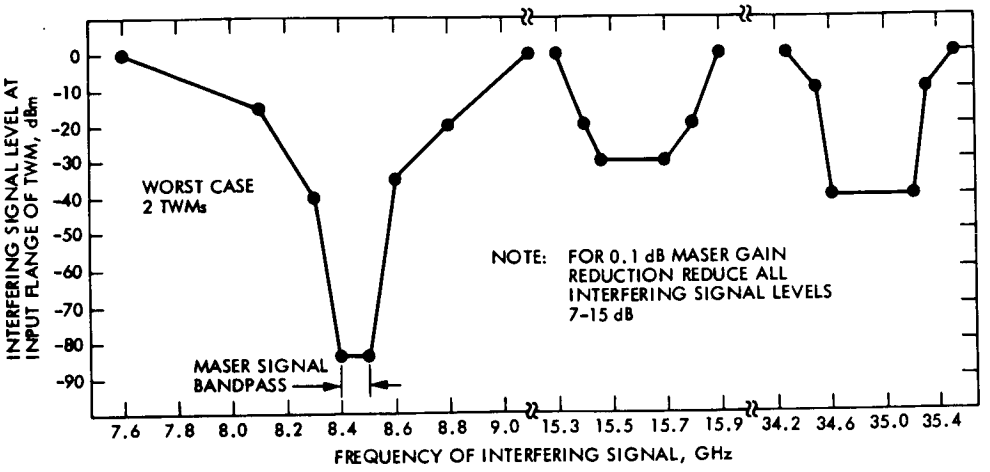



**Fig. 1. Plot of the signal level required to reduce maser gain 1.0 dB as measured at 8.4–8.5 GHz (X-band Block II TWM)**

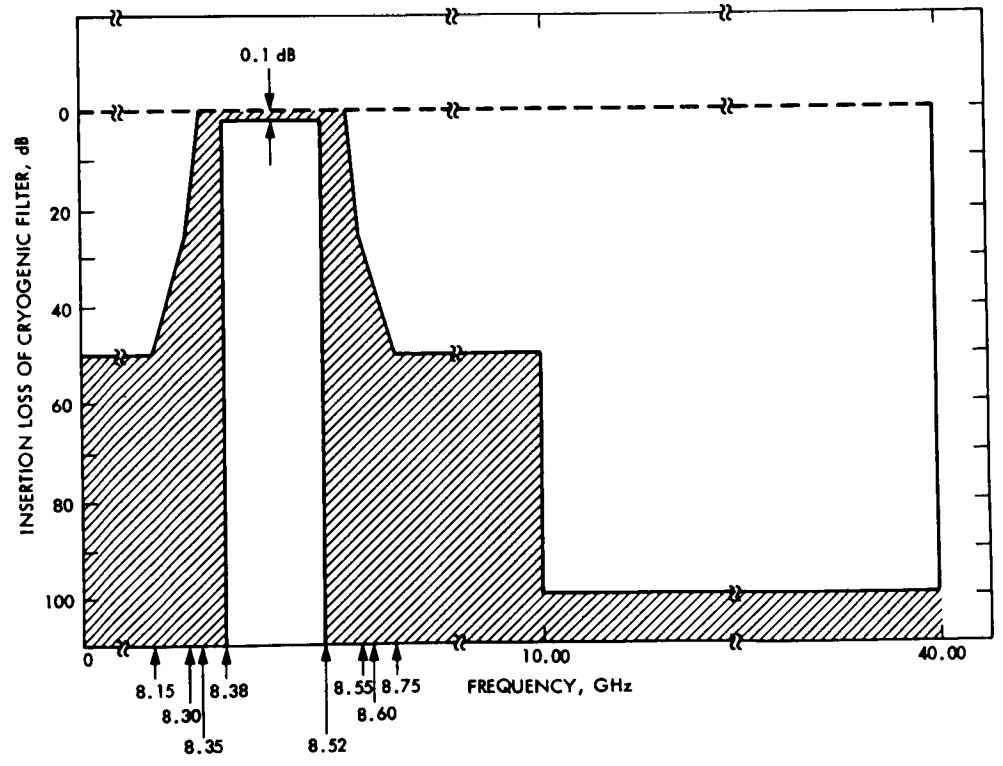



**Fig. 2. Plot of the design goals for a cryogenic low loss input filter (X-band Block II TWM)**

N82-11294



# Traveling-Wave Maser, Closed Cycle Refrigerator Automation

J. A. McNeil* and S. M. Petty
Radio Frequency and Microwave Subsystems Section

*Rationale, functional requirements, implementation guidelines, project organization, high level hardware and software design, and component selection are presented for the maser system automation project.*

## I. Introduction

At the present time all traveling-wave maser and closed-cycle refrigerator system (TWM-CCR) and their power supply and helium compressor assemblies are operated manually. Many man-hours are required of highly trained and experienced operators, resulting in high operator cost. Of greater importance are the problems of unpredictable failure and long downtimes. Since field performance is not monitored, system degradation is detected only when it has progressed to the point of catastrophic failure. Without the aid of performance histories, failure diagnosis is arduous and slow.

Rationale and functional requirements for the TWM-CCR Automation project are the following: (1) to improve reliability and reduce downtime by providing unattended monitor and control, alarm and fault location, failure prediction and self-corrective action; (2) to reduce operator cost by providing automatic tuning and cooldown procedures; and (3) to improve system operations and development by providing the needed engineering data.

This project is being implemented in phases. In Phase I existing technologies are evaluated, functional requirements defined, high-level plans developed, and component elements identified and procured. In Phase II the hardware is integrated, and the monitor and control functions demonstrated. In Phase III algorithms are developed for logging engineering data, performing system diagnostics, providing alarm status and fault location, and performing automated cooldown, tuning, and self-corrective procedures. This report outlines the efforts at the end of Phase I.

## II. System Design and Guidelines for Implementation

Before defining the project organization and guidelines, an extensive review of existing technologies and commercial products was made. To date only one maser has been partially automated, the dual channel S-band maser in operation at the Arecibo Observatory, which was designed and implemented by the Receiver Development Group of Cornell University for the National Astronomy and Ionosphere Center. The system was reviewed in detail, and extensive use has been made of the consulting services of the Resident Support Specialist for that project. Out of this review and the study of presently available commercial products have emerged the following guidelines for implementation.

(1) Components and subsystems are now commercially available so that the design and fabrication of custom maser-controller interfaces are neither necessary, desir-

*On leave from Department of Physics, Loyola Marymount University, Los Angeles, CA 90045.

able, nor cost effective. With the exception of three specialized subsystems (the TWM pumps, RF monitoring assemblies, and the CCR reserve capacity measuring assembly, which are being designed and fabricated at JPL), all sensors and interfaces are selected from readily available commercial products.

(2) Modification to the existing TWM-CCR configuration should be avoided when possible. This is to minimize the effects on the nonautomated operation of the system, to facilitate documentation and maintenance, and to facilitate the automation of other masers in the DSN. To this end, sensors and interfaces are attached to the system as external modules when possible.

(3) Distributed processing should be used when possible to facilitate the communication between the physically separated subsystems and to isolate and modularize functions. Therefore each major subsystem, the TWM-CCR, the system power supplies, and the helium compressor, will have a dedicated microprocessor-based single board providing the analog and digital I/O and the communication between that subsystem and the maser system controller.

(4) Common software and structured programming techniques should be used as much as possible to reduce software development time, maintenance effort, and documentation.

A block diagram of the system hardware is shown in Fig. 1. A discussion of the system hardware and software is given in the next section.

## III. Hardware and Software Implementation

A list of sensors is given in Table 1. These sensors are used to monitor the following parameters in each assembly.

(1) In the TWM-CCR assembly: temperatures of the 4, 15, and 70 kelvin stages; magnetic field; magnet current and charge rate; superconducting switch; refrigerator reserve capacity and drive unit frequency; microwave pump source bias, tune, and modulation voltages, and power; and maser RF gain versus frequency.

(2) In the power supply assembly: system voltages; vacion pump controls, current, and voltage.

(3) In the helium compressor: the pressures at the Joule-Thompson (J-T) supply and return; the pressures at the helium supply, refrigerator supply line and return, and the compressor storage tank and oil separator; the J-T mass flow; the temperatures of the first and second stages of the compressor, of the gas return, and of ambient; the voltages, currents, and power to the three phase motor; and supply voltages. Switch closures will indicate the on/off status of the compressor and refrigerator, as well as alarm conditions in the compressor temperature and pressure and in the airflow fan.

Analog signals control the superconducting magnet charge rate, magnetic field profile currents, microwave source pump tuning and modulation. Digital signals control system power supplies, vacion pump controls, compressor start, CCR start, compressor-CCR stop, and solenoid valves on the helium supply line, refrigerator return, J-T return and vent, storage tank, and the blowdown line. Switches and indicator lamps provide visual alarm indicators as well as a means of communication between field personnel and remote operations. A series of open or closed contacts provide a unique binary code for each maser system so that in the future many maser systems may be remotely identified and monitored. Sufficient extra terminal strip junctions and cabling are provided so that sensors and controls can be conveniently added or reconfigured.

All sensors and control elements connect directly to a commercial microprocessor-based interface, which contains its own power supply, and provides signal conditioning and linearization, excitation signals for the RTD and strain gauges, 13-bit integrating A/D, unit conversion, limit checking and alarm status, high common mode voltage isolation (1000 V peak) and high common mode rejection (160 dB). This interface, manufactured by Analog Devices (uMAC4000), is capable of storing 48 analog and 32 digital inputs in RAM at the rate of 15 channels/sec and can provide up to 32 digital outputs. Communication between this interface and the maser system controller is by RS-232C lines in ASCII format at speeds up to 9600 baud and is controlled by high level commands.

The maser system controller is a Multibus-based system in a configuration being standardized for controllers in the network consolidation project (NCP). Boards used include: CPU – BLC 80/204; disk controller – BLC 8201; 4-channel serial interface – iSBC 534; 64k RAM – BLC 064; and the JPL Quad Standard Interface.

Common software for these controllers includes:

(1) A disk-based operating system, which enables programs to be compiled, debugged, placed in ROM, and executed within the target controller.

(2) A real-time, priority-based, multitasking, multiprogramming executive (EMX) which occupies 1.4K of memory.

(3) A maintenance terminal handler (MTH), which uses a menu-prompting protocol to display system status,

modify the database, and perform subsystem operations.

(4) A driver for a Racah-Vadic modem to provide for remote operating and maintenance.

All coding is in PL/M compatible with both 8080 and 8086 CPU families.

## IV. Plans

Maser system controllers are now being implemented at JPL using existing R/D compressor and X-band TWM-CCR assemblies. Demonstration of all proposed monitor and control functions is planned for next year, and the development of automated procedures and demonstration of the entire system at DDS 13 in the year to follow.

## Acknowledgments

This work was initiated under the supervision of R. C. Clauss, who continues to support the effort with his expertise. The resources of several groups have been drawn upon in this effort. In particular, the contributions of C. Foster, H. Cooper, J. Estrada, B. Barber, F. McCrea, D. Neff, and N. Meyer are gratefully acknowledged. The services of L. Baker of the National Astronomy & Ionosphere Center, Cornell University have been invaluable.

**Table 1. Sensors used to monitor maser system performance**

| Measurement | Sensor | Range |
|---|---|---|
| Pressure | Viatran, series 218 | 0-25, 0-500, 0-3000 psig |
| Mass flow | Hasting, series AFSC | 0-10k, 0-50k SCCM |
| Temperature | RTD | 200-800 K |
| | Analog Devices, AD590 | 220-420 K |
| | Lake Shore Cryotronics, silicon diodes DT-500-DRC | 1-300 K |
| AC voltage | Ohio Semitronics, series VT | 0-150, 0-575 VAC |
| AC current | Ohio Semitronics, series CT | 0-20 A |
| AC power | Ohio Semitronics, PC5-18 | 0-14.4 kW |
| Magnetic field | Siemens, magnetoresistors, FP30L100J | 0-12 kG at 4.5 K |
| RF power | Hewlett-Packard, PDR | |

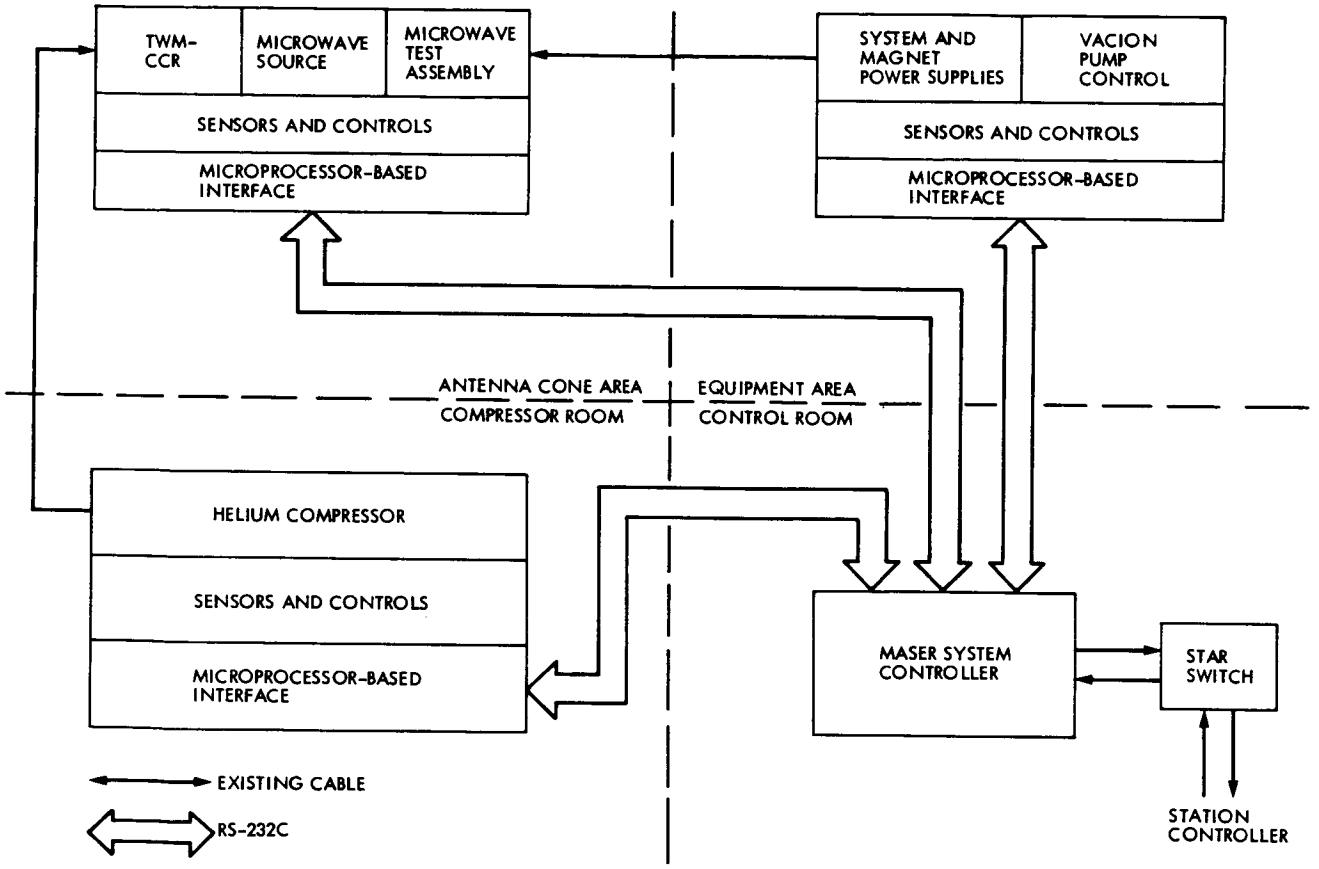



**Fig. 1. Simplified block diagram of maser system automation hardware. Sensors, controls, and interfaces are externally appended to existing assemblies with minimal effect on the manual mode of operation of the system**

D16 N82-11295



# Digital SAR Processing Using a Fast Polynomial Transform

S. Butman, R. Lipes, A. Rubin, and T. K. Truong
Communications Systems Research Section

*In this paper, a new digital processing algorithm based on the fast polynomial transform is developed for producing images from Synthetic Aperture Radar data. This algorithm enables the computation of the two-dimensional cyclic correlation of the raw echo data with the impulse response of a point target, thereby reducing distortions inherent in one-dimensional transforms. This SAR processing technique was evaluated on a general-purpose computer and an actual Seasat SAR image was produced. However, regular production runs will require a dedicated facility. It is expected that such a new SAR processing algorithm could provide the basis for a real-time SAR correlator implementation in the Deep Space Network.*

## I. Introduction

The Venus Orbiting Imaging Radar (VOIR) mission planned for the late 1980's will require handling high rate (> 2-4 msps) telemetry containing synthetic aperture radar (SAR) data. In the planning for upgrading the current DSN system to handle the increased data rates, consideration must be given to how SAR data will affect telemetry validation and monitoring functions. Generally, signal-to-noise and error rate estimates provide sufficient inputs for telemetry monitoring. However, these estimates by themselves are not always good indicators of image quality; direct monitoring of images is a more effective method of detecting telemetry degradation. For missions in which images are readily reconstructed from telemetry data such as in optical imaging, this kind of monitoring is accomplished by real-time image processing at JPL, which takes only a few minutes' delay. However, even in the "quick-look" mode for VOIR SAR data processing, the delay between telemetry reception and image production will be greater than three hours – an unacceptable delay considering that a station can receive for only about eight hours a day. A better method to reduce the time between SAR data reception and image production to an acceptable value must be developed in order to validate the Deep Space Network (DSN) performance. One technique would be to locate a low-cost quick-look SAR processor at each complex or one at the Network Operations and Control Center (NOCC).

Section II explains the background and the concept of a new technique for digital SAR processing. In Section III, it is explained how the range offset data were conveniently translated to baseband for processing. In Section IV, the formulation of the two-dimensional impulse response of a point target is developed. In Section V, some of the data management considerations for using the fast polynomial transform (FPT) for two-dimensional correlation are treated. Finally, Section VI explains how multiple looks are registered and combined and presents results of SEASAT SAR imagery. Section VII presents the conclusions of this report.

## II. Background

Earlier work (Ref. 1) has shown that a VOIR processor handling data from the high-resolution mode must compensate for range migration. When range migration occurs, the two-dimensional impulse response from a point target does not factor into the product of two one-dimensional impulse responses. Currently, SEASAT SAR data accommodates the range migration feature by the digital hybrid correlation algorithm of Wu (Ref. 2). In brief, the azimuth SAR filter is approximated by several linear segments. Then the SAR data corresponding to each segment is correlated separately. Finally the results are coherently summed to produce the full resolution image. The advantage of Wu's hybrid algorithm is that the Fourier transform of the azimuth matched filter is one-dimensional. Hence, one-dimensional FFT can be used to compute azimuth correlation. However, when the two-dimensional impulse response does not change too rapidly with range, full two-dimensional correlation of SAR data with the proper impulse response would eliminate the need for the approximating, linear segments of the hybrid algorithm. Nussbaumer, Quandalle, Arambepola, and Rayner (Refs. 3, 4), and more recently, Truong, Reed, Lipes, and Wu (Ref. 5) showed that a radix-2 FPT could be used to efficiently compute a two-dimensional correlation. It was shown (Ref. 6) that a combination of the FPT and the CRT could be used to very efficiently compute a two-dimensional cycle convolution of $d_1 \times d_2$ array of complex numbers, where $d_2 = 2^m$ and $d_1 = 2^{m-r+1}$ for a $1 \leqslant r \leqslant m$. This FPT convolutional algorithm requires considerably fewer multiplications and about the same number of additions that the usual FFT method needs for the two-dimensional case. It is shown (Ref. 6) that this new algorithm can be implemented readily on a digital computer. It is demonstrated also that the speed of this algorithm is approximately 25% faster than the conventional FFT method for computing two-dimensional convolutions.

To investigate the FPT approach to correlate SAR data with significant range migration, SEASAT SAR data having the characteristics listed in Appendix A was processed on a general-purpose computer. An approach capable of compensating for the magnitude of range migration of the SEASAT data can easily handle VOIR high-resolution data range migration. The main purpose of this effort was to uncover features of the FPT approach that might prevent a real-time implementation and to determine qualitatively whether the range-dependent variation in impulse response would affect image quality. The important real-time processing considerations of autofocusing and Doppler centroid determination have not been treated and will be deferred to later work. The values of Doppler center frequency and Doppler rate used for the impulse response have been supplied to us by C. Wu. A flow diagram of the digital processing system is shown in Fig. 1. Raw SEASAT, range offset data representing 5120 azimuth cells by 3072 range cells were input.

## III. Translation of Offset Spectrum of the Raw Range Data to Baseband

In general, the basic geometry of SEASAT SAR is shown in Fig. 2. The antenna of the uncertainties in the SAR attitude flies at a height $h$ above the moving Earth's surface with the center of beam (boresight) making an incident angle; i.e., $\beta = 20^\circ$ with nadir. The direction along which the spacecraft moves is called the azimuth or $x$ direction and the distance in the direction to a point scatterer as measured from the antenna is called the range or $r$ direction.

The SEASAT SAR transmits a pulse waveform of the form $S(t) = a(t) \cos(2\pi f_0 t + \pi bt^2)$ where $a(t)$ is a rectangular pulse waveform with pulse length $\tau$, $f_0$ is the frequency of the coherent carrier and $\pi bt^2$ is the term needed for linear frequency modulation. Let $T_1$ be the pulse repetition time. In Fig. 2, one observes that if the SEASAT SAR transmits the $n$th pulse, the distance from the SEASAT SAR to a point scatterer location at $(t_0, r(0))$ is $r(nT_1)$, where $r(0)$ is the minimum distance from a point scatterer to the antenna. The returned signal from the $n$th pulse is of form

$$S\left(t - nT_1 - 2\frac{r(nT_1)}{c}\right) = \sigma G\left(\frac{x_0 - nT_1 v}{r(0)}\right) a\left(t - nT_1 - \frac{2r(nT_1)}{c}\right) \cos\left[2\pi f_0 \left(t - nT_1 - \frac{2r(nT_1)}{c}\right) + \pi b \left(t - nT_1 - \frac{2r(nT_1)}{c}\right)^2\right] \quad (1)$$

where $\sigma$ is the radar cross section of a point scatterer location at $(x_0, r(0))$, $f_0$ is the transmitter frequency, $c$ is the speed of light, $G(\theta)$ is the physical antenna pattern in azimuth direction $\theta$, and the pulse is

$$a\left(t - nT_1 - \frac{2r(nT_1)}{c}\right) = \begin{cases} 1 & \text{if } 0 \leqslant t - nT_1 - \dfrac{2r(nT_1)}{c} \leqslant \tau \\ 0 & \text{otherwise} \end{cases}$$

Let $t_1 = nT_1$ and $t_2 = t - nT_1$ for that integer $n$ such that $0 \leqslant t_2 \leqslant T_1$. Then (1) becomes

$$S\left(t_2 - \frac{2r(t_1)}{c}\right) = \sigma G\left(\frac{x_0 - vt_1}{r(0)}\right) a\left(t_2 - \frac{2r(t_1)}{c}\right)$$

$$\cos\left[2\pi f_0\left(t_2 - \frac{2r(t_1)}{c}\right) + \pi b\left(t_2 - \frac{2r(t_1)}{c}\right)^2\right] \quad (2)$$

where $r(t_1) = r(nT_1)$ and

$$a\left(t_2 - \frac{2r(t_1)}{c}\right) = \begin{cases} 1, & \text{if } 0 \leqslant t_2 - \dfrac{2r(t_1)}{c} \leqslant \tau \\ 0, & \text{otherwise} \end{cases}$$

In SEASAT-A SAR, the range data constitutes a block of 4096 samples of raw data, each of 4-bits. In order to convert this real sampled data, the range data, into complex sampled data, advantage is taken of the fact that spectrum is offset from baseband by frequency $f_s/4$, where $f_s = 45.53$ MHz is the sampling rate of the real raw radar data. A method of translating the spectrum to baseband is now described.

If the radar footprint is regarded as a collection of distributed point scatterers, the returned signal has the form

$$x(t_2) = \sum_j \sigma G\left(\frac{x_j - vt_1}{r_j(0)}\right) a\left(t_2 - \frac{2r_j(t_1)}{c}\right) \cos\left[2\pi \frac{f_s}{4}\left(t_2 - \frac{2r_j(t_1)}{c}\right) + \pi b\left(t_2 - \frac{2r_j(t_1)}{c}\right)^2\right] \quad (3)$$

where $G(\theta)$ is the gain of the antenna and $r_j(0)$ is the minimum of the distance from the $j$th point scatterer to the antenna, and $r_j(t_1)$ is the distance from the antenna to the $j$th point scatterer within the footprint. Sampling $x(t_2)$ in (3) at intervals $T_2 = 1/f_s$, yields

$$x(kT_2) = \sum_j \sigma G\left(\frac{x_j - vt_1}{r_j(0)}\right) a\left(kT_2 - \frac{2r_j(t_1)}{c}\right) \cos\left[2\pi \frac{f_s}{4}\left(kT_2 - \frac{2r_j(t_1)}{c}\right) + \phi_j(kT_2)\right], \quad 0 < k < 4095 \quad (4)$$

where

$$\phi_j(kT_2) = \pi b\left(kT_2 - \frac{2r_j(t_1)}{c}\right)$$

is the sampled data sequence. Since $T_2 = 1/f_s$ for even and odd $k$, this formula for $x(kT_2)$ in (4) can be expressed as follows:

$$x(2nT_2) = (-1)^n \sum_j \sigma G\left(\frac{x_j - vt_1}{r(0)}\right) a\left(2nT_2 - \frac{2r_j(t_1)}{c}\right) \cos\left[\frac{2\pi f_s}{4}\frac{2n}{f_s} + \phi_j(2nT_2)\right]$$

$$= (-1)^n \sum_j \sigma_j G\left(\frac{x_j - vt_1}{r_j(0)}\right) a\left(2nT_2 - \frac{2r_j(t_1)}{c}\right) \cos\phi_j(2nT_2) \quad \text{for } n = 0, 1, 2, \ldots\ldots \quad (5a)$$

Similarly,

$$x[(2n+1)T_2] = (-1)^{2n+1} \sum_j \sigma G\left(\frac{x_j - vt_1}{r_j(0)}\right) a\left((2n+1)T_2 - \frac{2r_j(t_1)}{c}\right) \sin\phi_j[(2n+1)T_2] \quad \text{for } n = 0, 1, 2, \ldots \quad (5b)$$

In (5a), one observes that $x(2nT_2)$ is a cosine sequence for $n = 0, 1, 2, \ldots$. In order to convert $x(2nT_2)$ given in (5a) into complex sequence $Z(2nT_2)$, the imaginary part of $Z(2nT_2)$, i.e., a sine sequence, is required to compute for summing to cosine sequence given in (5a). This sequence denoted by $x'(2nT_2)$ is obtained by interpolating the samples of sine sequence $x[(2n+1)T_2]$ given in (5b) for $n = 0, 1, 2, \ldots$. Evidently, from the sum of these sine and cose sequences, the appropriate complex data stream is obtained by forming

$$Z(kT_2) = -[x(2kT_2) + jx'(2kT_2)]$$

$$= \sum_j \sigma G\left(\frac{x_j - vt_1}{r_j(0)}\right) a(2kT_2) e^{-i\phi_j(2kT_2)}$$

$$\text{for } 0 \leqslant k \leqslant 1023 \tag{6}$$

Thus, each range sweep of data has been converted to a stream of 2048 complex data points.

## IV. The Design of the Two-Dimensional SAR Impulse Response of a Point Target

The returned signal of a point scatterer located at $(x_0, r(0))$ is given in Eq. (2). If the moving antenna radiates the successive pulses to a point scatterer on Earth location at $(x_0, r(0))$, then $t_1$ in $r(t_1)$ given in (2) lies between $-T/2$ and $T/2$, where $T$ is the total amount of time a point scatterer is in the mainbeam. Consider the effects of a point scatterer on moving Earth and the effects of the uncertainties in the SAR attitude. Then, using the Taylor series expansion, $r(t_1)$ in (2) about $t_1 = 0$ is given by

$$r(t_1) = r(0) + \dot{r}(0)t_1 + \frac{\ddot{r}(0)}{2} t_1^2 \quad , \quad |t_1| \leqslant T/2 \tag{7}$$

$r(t_1)$ in (7) can be rewritten as

$$r(t_1) - r(0) = \frac{\lambda}{2}\left[f_d t_1 + \frac{1}{2}\dot{f}_d t_1^2\right] \quad , \quad |t_1| \leqslant T/2 \tag{8}$$

where $f_d = 2\dot{r}(0)/\lambda$ and $\dot{f}_0 = 2\ddot{r}(0)/\lambda$ are called the doppler center frequency and the doppler center frequency rate, respectively.

In SEASAT SAR, the values of $f_d$ and $f_d$ for the Los Angeles/Burbank area are $-1415$ Hz and $-515.1$ Hz/sec, respectively. Using the parameters given in Appendix A, one obtains the azimuth integration time to be $T = 25$ sec and $\lambda = 0.235$ (see Ref. 1). Thus, (8) becomes

$$r(t_1) - r(0) = 0.1175\ [-1425\, t_1 - 257.55\, t_1^2], |t_1| \leqslant 1.25 \text{ sec} \tag{9}$$

Assume that the returned signals from the successive pulses are recorded side-by-side. The result forms a two-dimensional impulse response of a point scatterer location at $(x_0, r(0))$ with range $r = c \cdot t_2/2$ along one axis and azimuth $x = vt_1$ along the other axis.

To define a 2-D point scatterer region of a matched filter, i.e., $R$, consider a fixed azimuth location $t_1 = nT_1$; then the range time corresponding to $t_1$ is $t_2 = 2r(t_1)/c$. From (2), the returned signal from the $n$th pulse, $t_1 = nT_1$, only lies in the range time between $t_2 = 2r(t_1)/c$ and $t_2 = \tau + 2r(t_1)/c$. Assume that the moving antenna with uncertainties in the SAR attitude radiates the successive pulses to a point scatterer on moving Earth. From (9) one observes that the azimuth time is located between $|t_1| \leqslant T/2 = 1.25$ sec, where $T = 25$ sec is the total amount of time a point scatterer is in the mainbeam. Thus, $R$ is determined by the pulse width and the pattern of the antenna beam. A 2-D impulse response of a point scatterer only lies in the location between $|t_1| \leqslant 1.25$ sec and $2r(t_1)/c \leqslant t_2 \leqslant 2r(t_1)/c + \tau$. In fact, a two-dimensional matched filter excluding the effect of the physical antenna is

$$h(t_1, t_2) = S\left(t_2 - \frac{2r(t_1)}{2}\right)$$

$$= \begin{cases} e^{-i4\pi r(t_1)/\lambda} \cdot e^{i\pi b(t_2 - 2r(t_1)/c)^2}, & (t_1, t_2) \in R \\ 0, & \text{otherwise} \end{cases} \tag{10a}$$

where $R$ is the 2-D point scatterer region, $\exp\{-i4\pi r(t_1)/\lambda\}$ and $\exp\{i\pi b(t_2 - 2r(t_1)/c)^2\}$ are the azimuth and range matched filters, respectively, and $\lambda = f_0/c$, $t_1$ and $t_2$ are in azimuth and range axis, respectively.

Consider the effect of range migration for a matched filter in SEASAT SAR. For the case $t_1 = 1.25$ sec, (9) becomes

$$r(-1.25) - r(0) = 0.1175\ [1425\,(-1.25) - 257.55\,(-1.25)^2]$$

$$= -162.0 \text{ m}$$

In a similar fashion, $r(1.25) - r(0) = 256.25$ m. Hence the range migration is

$$\Delta r = r(1.25) - r(-1.25) = 256.25 - (-162.0) = 418.25 \text{ m}$$

Since the range migration is greater than the range resolution given in Appendix A, i.e., $\Delta r \gg 6.6$ m, the impulse response of a point target at location $(x_0 = 0, r(0))$ must be two-dimensional. With the parameters given in Appendix A, the size of this two-dimensional matched filter can be computed to be $4096 \times 832$ (see Ref. 1). In SEASAT SAR, $b = 0.562 \times 10^{12}$ cycles/sec$^2$. Thus, substituting $b$ and

$r(t_1) - r(0)$ given in (9) into (10a), a two-dimensional matched filter is of form

$$h(t_1,t_2) = \begin{cases} e^{i2\pi(f_d t_1 + \dot{f}_d t_1^2/2) \cdot e^{ib\pi[t_2 - a(f_d + \dot{f}_d t_1^2/2)]^2}}, & \\ & (t_1,t_2) \in R \quad (10b) \\ 0 & \text{otherwise} \end{cases}$$

where $f_d$ = 1425 Hz, $\dot{f}_d$ = 515.1 Hz/sec, $b = 0.562 \times 10^{12}$ cycles/sec$^2$, $a$ = 0.079 sec, and the 2-D filter region $R$ is located between $|t_1| \leq 1.25$ sec and $2r(t_1)/c \leq t_2 \leq 2r(t_1)/c +$ 33.8 $\mu$sec and is shown in Fig. 3. Since the number of looks for SEASAT SAR is four, the size of the filter per look is a 1024 X 832 array of complex samples (see Fig. 3).

## V. Fast Polynomial Transform for Computing a Two-Dimensional Correlation

In the previous section, one observed that the two-dimensional filter per look is a 1024 X 832 array of complex samples. If the 1024 X 832 filter is extended with zeroes to the size of a 2048 X 2048 array of complex data, then it can be correlated with a 2048 X 2048 array of raw data to obtain a 256 X 1216 image for each look with 6.6- X 25-m resolution in range and azimuth. This two-dimensional cyclic correlation of a 2048 X 2048 array of complex data can be computed by a radix-2 FPT. The radix-2 FPT algorithm is given in Refs. 5 and 6.

## VI. Registration and the Combining of Looks

In order to produce an image, one needs to correlate the raw data with the matched filter given in Eq. (10b). This procedure can be accomplished by using a 4-look overlap-save FPT procedure. The procedure currently being used for this 4-look FPT procedure is the following.

A 6144 X 2048 array of complex raw data is taken. It is divided into 1024 X 2048 pieces denoted by A, B, C, D, E, F. The matched filter is a 4096 X 832 array of complex data. It is divided into 1024 X 832 pieces denoted by looks 1, 2, 3, 4 (see Fig. 3). Each look is extended to a 2048 X 2048 with zeroes. The desired effect is to correlate the complete raw data with the complete matched filter, producing a 2048 X 1216 picture. Since we only have the capability of performing a 2048 X 2048 cyclic correlation, AB is correlated with look 1 to obtain a 1024 X 1216 array of real imagery data that has an azimuth resolution of 25 m for one look. Similarly, look 2 is correlated with BC, look 3 is correlated with CD, and look 4 is correlated with DE. Then these resulting looks are summed to obtain the first 1024 X 1216 image for all looks with 25-m azimuth resolution. To obtain another 1024 X 1216 image, looks 1, 2, 3, and 4 are correlated with BC, CD, DE, and EF, respectively. These resulting looks are summed to obtain the second 1024 X 1216 image for all looks with 25-m azimuth resolution. Hence, a 2048 X 1216 array of real imagery data with 6.6- X 25-m range and azimuth resolution is obtained. Since a 4-look overlap-save FPT is used in this procedure, the actual picture is (2048/4) · 1216 = 512 X 1216 array of real data with 6.6- X 25-m resolution image for all looks.

Using the same procedure described above, a 4-look SEASAT SAR 512 X 1216 array of real imagery data of the Los Angeles/Burbank area with 6.6- X 25-m resolution is shown in Fig. 4.

## VII. Conclusions

An FPT has been used for digital SAR processing and an SAR image has been generated. Moreover, the FPT was shown to have an architecture suitable for hardware implementation. These advantages make this FPT algorithm a good candidate for developing a real-time SAR processor.


## Acknowledgment

The authors express thanks to Dr. I. S. Reed for many helpful suggestions on the subject of SAR and to Dr. C. Wu for supplying them with the raw SEASAT data and corresponding reference functions.

# References


1. Truong, T. K., and Lipes, R., "Concerning the Feasibility of a Real-Time SAR Digital Processor for VOIR Low Resolution Imaging Modes," *TDA Progress Report 42-50, May and June 1980*, Jet Propulsion Laboratory, Pasadena, California, pp. 72-81.

2. Wu, C., "A Digital System to Produce Imagery from SAR" Systems Design Driven by Sensors, AIAA conference, Pasadena, California, October 18-20, 1976.

3. Nussbaumer, H. J., and Quandalle, P., "Computation of Convolutions and Discrete Fourier Transforms by Polynomial Transforms," *IBMJ. Res. Develop*, Vol. 22, No. 2, March 1978, pp 134-144.

4. Arambepola, B., and Rayner, P. J. W., "Efficient Transforms for Multidimensional Convolutions," *Electronics Letters*, March 15, 1979, Vol. 15, No. 6, pp. 189-190.

5. Truong, T. K., Reed, I. S., Lipes, R., and Wu, C., "On the Application of a Fast Polynomial Transform and the Chinese Remainder Theorem to Compute a Two-dimensional Convolution," *IEEE Trans. Acoustics Speech and Signal Processing*, February 1981, ASSP-29, pp. 91-97.

6. Reed, I. S., Shao, H. M., and Truong, T. K., "The Fast Polynomial Transform and Its Implementation on a Computer," *IEE Proc.* Vol. 128, Pt. E, No. 1, March 1981, pp. 50-60.

# Appendix A

# Seasat SAR Sensor Characteristics

SAR System Parameters (SEASAT-A SAR)

| | |
|---|---|
| SAR orbit | SEASAT-A Orbit |
| Nominal altitude | 794 km |
| Nominal speed | 7450 m/sec |
| Transmitter frequency | 1275 MHz |
| Pulse repetition frequency | 1463, 1537, 1645 Hz |
| Pulse width | 33.8 μsec |
| Pulse bandwidth | 19 MHz |
| A/D rate for range offset signals | 45.53 MHz |
| A/D window | 288 μsec |
| Antenna dimension | 2 m X 10.5 m |
| Antenna look angle | 20° cone |
| Attitude (roll, pitch, yaw) accuracy | ±0.5° |
| Image dynamic range | 50 dB |
| Image resolution (range and azimuth) | 6.6 m/25 m |
| Number of looks | 4 |

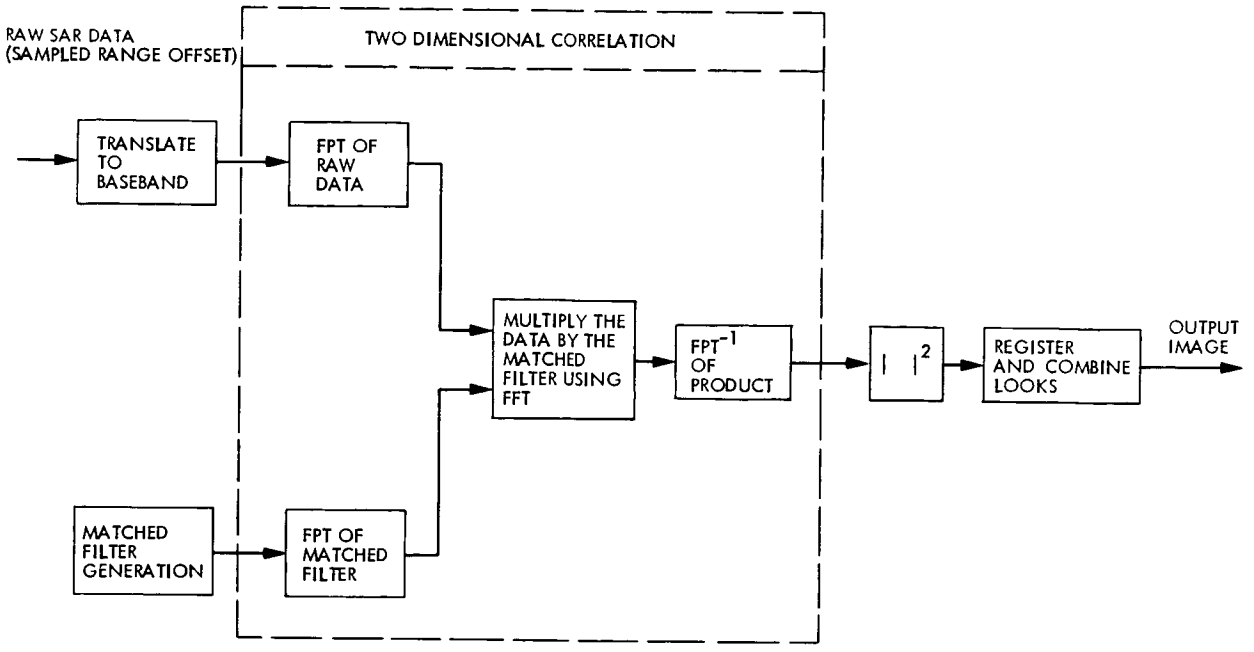



**Fig. 1. Flow diagram of digital SAR processing system using the fast polynomial transform**

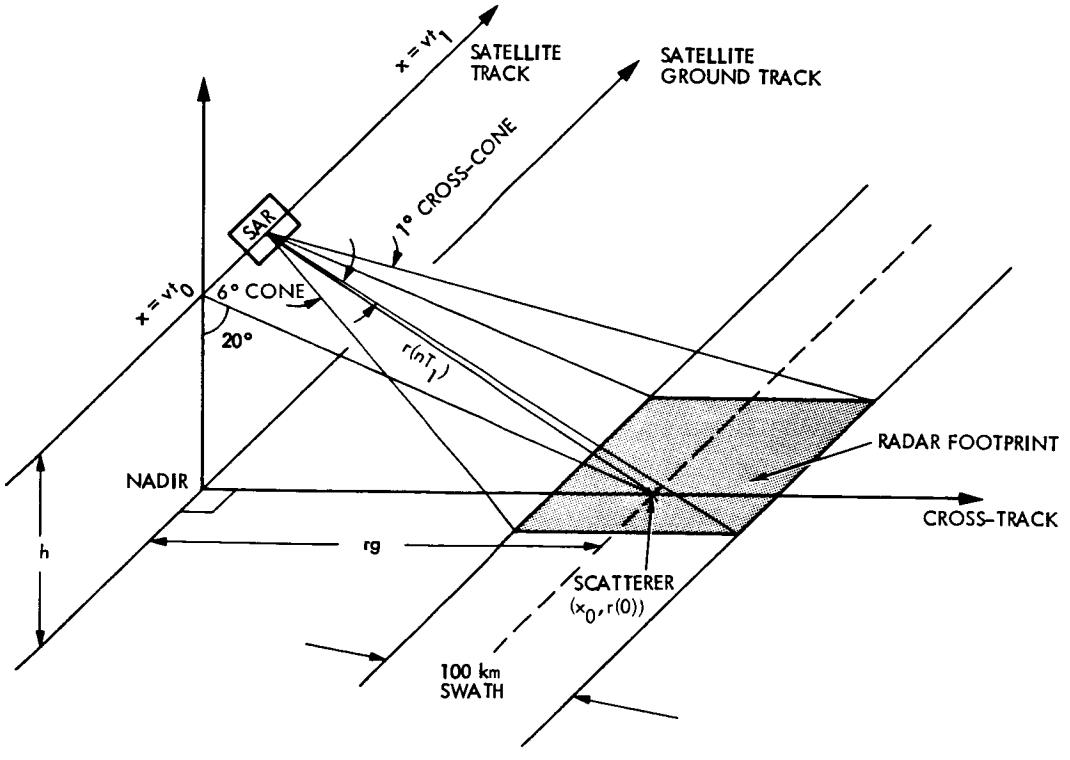



**Fig. 2. SEASAT SAR imaging geometry from the *n*th pulse**

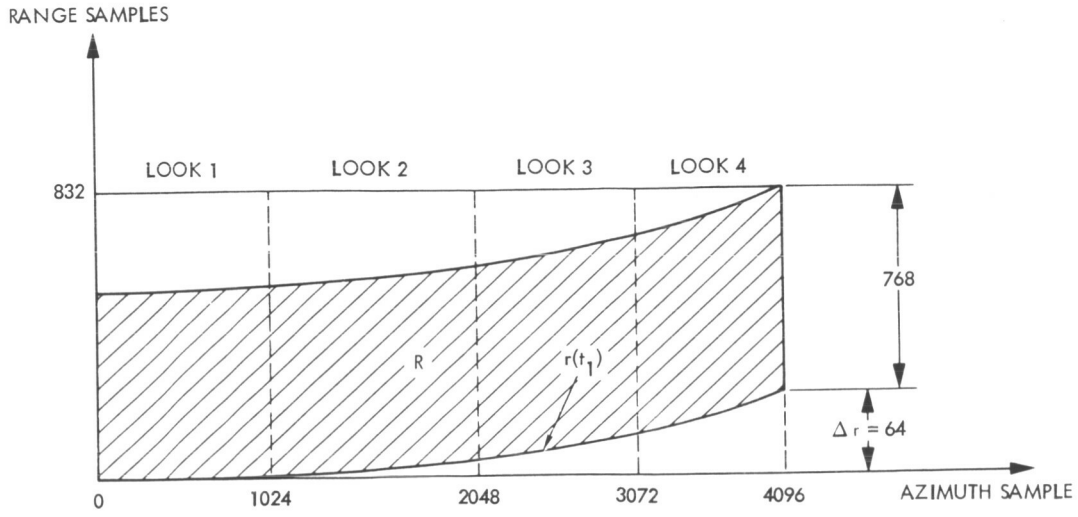



**Fig. 3. A range migration and the 2-D region *R* of a matched filter for SEASAT SAR processing**

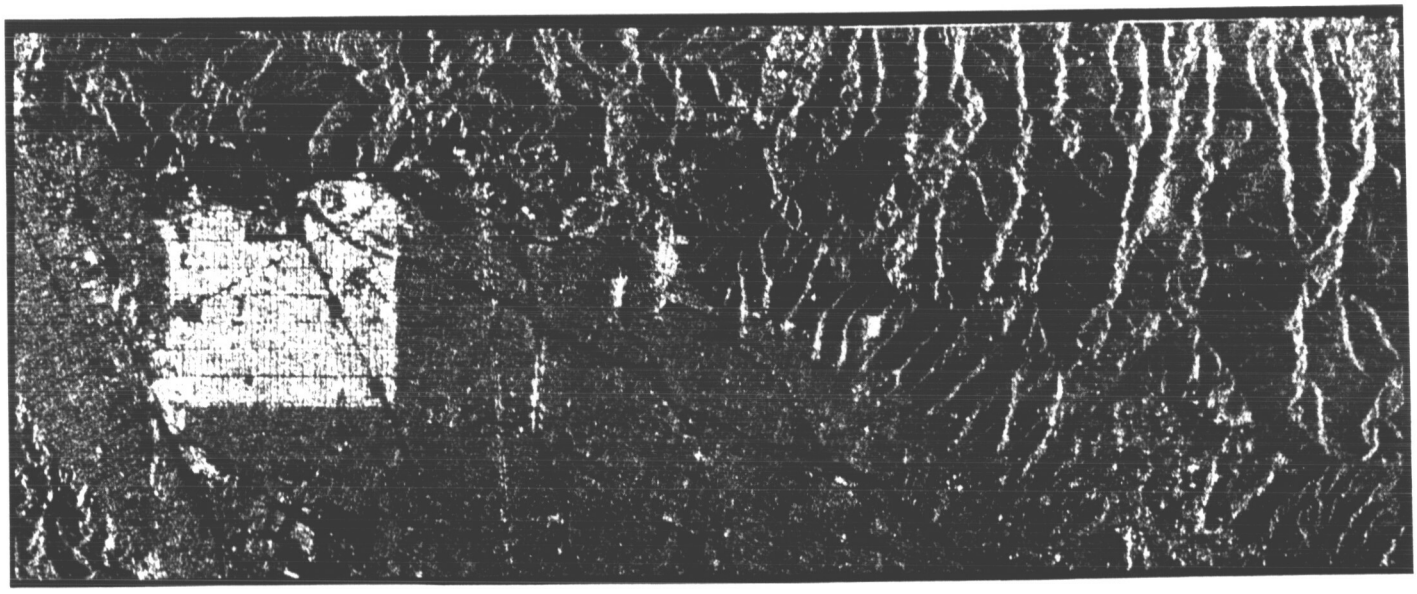


**Fig. 4. A 512 × 1216 array of real imagery data of the Los Angeles/Burbank area with 4-look SEASAT SAR 6.6- × 25-m resolution**

ORIGINAL PAGE IS
OF POOR QUALITY

N82-11296 417



# Minimizing the Time to Troubleshoot a Failed System

E. J. Subelman
Communications Systems Research Section
and
University of California, Los Angeles

*This report presents a procedure to determine the order in which to inspect the components of a failed series system so as to minimize the expected time until the failed component is isolated. Our model includes the case in which inspection of a component may falsely indicate the component is functioning.*

## I. Introduction

In many of the systems and subsystems of the Deep Space Network, one is faced with the problem of troubleshooting a complex system that has failed, that is, with isolating the failed components in order to repair or replace them.

The system's down time can be divided into two phases: troubleshooting and repair. Usually one has little control on the duration of the second phase, except by finding better repair procedures, which may be difficult to do. Furthermore, in the case of modern electronic equipment, repair time consists of the replacement of a circuit board so that the only time to consider is the procurement time for that board.

On the other hand, one usually has a greater degree of control on the troubleshooting procedures, and reducing the total time to detect the fault provides us with a simple way of reducing total system downtime.

Thus the motivation for this paper: to determine the procedures to be followed when troubleshooting a system so as to minimize the total time spent in this phase.

## II. The Model

We will consider a series system of $n$ statistically independent components, that is, a system in which the failure of any one component causes the failure of the system. Suppose the system has just failed, so that we know that one component has failed. In order to repair the system, one would like to determine which component has caused the failure, and this usually requires inspection of the components.

Assume that inspection of component $i$ requires an amount of time $T_i$. Furthermore assume that, if component $i$ is indeed failed and we inspect it, then with probability $\alpha_i$ we will detect the failure, and with probability $\beta_i = 1 - \alpha_i$ we will overlook the failure. Clearly perfect inspection corresponds to $\alpha_i = 1$, and when $\alpha_i < 1$, more than one inspection of component $i$ may be necessary to uncover the fault. We will assume that $T_i > 0$ for all $i$. Our goal is to find an inspection plan that minimizes the total expected time until the faulty component is detected.

Let $F_i(t) = P$ (life of component $i \leqslant t$); that is, $F_i$ is the distribution of the life of component $i$. Let $f_i(t) = (d/dt)\, F_i(t)$

and $\bar{F}_i(t) = 1 - F_i(t)$. We also denote by $r_i(t)$ the failure rate

$$r_i(t) = \frac{f_i(t)}{\bar{F}_i(t)}$$

Although the system will initially have all new components, after several repairs have been completed, the ages of the components will be different. We will, however, assume that a repaired component has the same life distribution $F_i$. This assumption is not necessary but makes our notation simpler.

Let $A_i$ be the age of the component $i$ at the time of the last repair of the system. Let $x$ be the age of the system at the time of failure. Then it can be shown (see Appendix A) that the probability that component $i$ caused the system failure is

$$p_i = \frac{r_i(x + A_i)}{\sum_{j=1}^{n} r_j(x + A_j)} \tag{1}$$

We remark that for the subsequent development, all we require is the value of $p_i$, which could be modified to incorporate any additional information available, such as might be obtained by observation of the symptoms of the system's failure, or even subjective estimates. We will also make the assumption that all components must be inspected; that is, even if $p_i = 1$ we must still inspect component $i$, presumably to pinpoint the cause of failure in order to carry out repairs.

## III. The Optimal Policy

With this information we can proceed to determine an optimal inspection plan, which is simply a list of components $i_1, i_2, i_3, \cdots$ with the interpretation: inspect component $i_1$. If you do not find it failed, proceed to component $i_2$, etc. stopping when the failed component is found. The list may contain repetitions in the case of imperfect inspection.

If we inspect component $j$ and do not find it failed, then this information can be incorporated by modifying the values of $p_i$ $i = 1, 2, \cdots, n$ by means of Bayes' theorem, to obtain

$$p_i' = \begin{cases} \dfrac{p_j}{1 - \alpha_j p_j} & \text{for all } i \neq j \\ \dfrac{(1 - \alpha_j)p_j}{1 - \alpha_j p_j} & \text{for } i = j \end{cases} \tag{2}$$

We can think of the inspection procedure in terms of the set of $\{p_i\}_{i=1}^n$. Given the initial set, we select the first component to inspect. If we do not find it failed, we use (2) to modify the probabilities and incorporate this additional information, and then select a new component to inspect; the process continues until we find the failed component.

It is clear that if $p_i > 0$ then the optimal search policy will eventually examine component $i$ since otherwise it would have an infinite expected duration.

The following result characterizes the optimal policy:

*Theorem*. The optimal policy always inspects the component yielding

$$\min\left\{\frac{T_i}{p_i\alpha_i}\right\}.$$

*Proof*. Suppose, without loss of generality, that

$$\frac{T_1}{p_1\alpha_1} = \min\left\{\frac{T_i}{p_i\alpha_i}\right\},$$

and suppose component 1 is not inspected first. Assume component 2 is the last component inspected before component 1, and that this is the $(a + 1)$st inspection of component 2.

The inspection sequence is then $j_1, j_2 \cdots j_k, 2, 1, \cdots$ with component 2 appearing $a$ times in the first $k$ entries. We will prove that the sequence $j_1, j_2 \cdots j_k, 1, 2, \cdots$ gives a smaller expected inspection time.

For the first sequence, the expected inspection time is, conditioning on which is the failed component,

$$V_1 = p_1[M_1 + T_2 + T_1 + N_1]$$

$$+ p_2[M_2 + \beta_2^a T_2 + \beta_2^{a+1} T_1 + N_2]$$

$$+ (1 - p_1 - p_2)N_3$$

where $M_1$ and $M_2$ depend only on $j_1 \cdots j_k$ and $N_1, N_2$ and $N_3$ depend on the tail of the sequence. These terms are not affected by the interchange of 1 and 2 and so their form is not important for our purposes.

For the second sequence the expected time is

$$V_2 = p_1[M_1 + T_1 + \beta_1 T_2 + N_1]$$

$$+ p_2[M_2 + \beta_2^a T_1 + \beta_2^a T_2 + N_2]$$

$$+ (1 - p_1 - p_2)N_3$$

Clearly $V_1 \geqslant V_2$ if and only if

$$p_1\alpha_1 T_2 \geqslant \beta_2^a p_2 \alpha_2 T_1$$

Thus, if

$$\frac{T_1}{p_1\alpha_1} < \frac{T_2}{p_2\alpha_2}$$

then for any $a \geqslant 0$

$$\beta_2^a p_2 \alpha_2 T_1 < p_1\alpha_1 T_2$$

and therefore

$$V_2 < V_1 .$$

It follows that component 1 should be inspected first.

The procedure for inspection of the system can be summarized as follows:

(1) Initially $A_1 = A_2 = \cdots A_n = 0$ (all new components).

(2) Measure the system uptime until failure, and call it $x$.

(3) For each $i = 1, 2, \cdots n$ compute $p_i$ using (1) or its equivalent (3) in Appendix A.

(4) Inspect the component giving

$$\min\left\{\frac{T_i}{p_i\alpha_i}\right\},$$

say component $j$. If it is failed go to step 6.

(5) For each $i$ replace $p_i$ by $p_i'$ obtained by using (2), and go to step 4.

(6) Component $j$ failed. Repair or replace it. Set $A_j = 0$ and $A_i = A_i + x$ for all $i \neq j$. Let the system operate and return to step 2.

In the special case in which inspection of the components yields perfect information on their state ($\alpha_1 = \alpha_2 \cdots = \alpha_n = 1$) the optimal strategy has an even simpler form:

*Corollary* If $\alpha_i = 1$ for all $i$ then the optimal policy inspects the components in the order of increasing values of $T_i/p_i$.

*Proof* By the theorem above, the first component to be inspected is that with the smallest value of $T_i/p_i$, say component 1. If it is not found failed, then the $p_i$ will be modified using (2) to

$$p_i' = p_i/(1 - p_1) \quad \text{for } i = 2, 3, \cdots n$$

$$p_1' = 0$$

Thus the new ratios will be

$$\infty, T_2(1 - p_1)/p_2, T_3(1 - p_1)/p_3, \cdots T_n(1 - p_1)/p_n$$

Since the new ratios are (excepting that for component 1) the original ratios scaled by $(1 - p_1)$, their ordering remains the same, so that the next component to be inspected should be the one with the second smallest value of $T_i/p_i$.

## IV. An Example

We consider a system consisting of three components, all of them having a Weibull life distribution

$$F(t) = 1 - e^{-(\lambda t)^a}$$

The values of the parameters are

| Component | $\lambda$ | $a$ | Inspection time |
|---|---|---|---|
| 1 | 0.02 | 1/2 | 60 min |
| 2 | 0.01 | 1 | 50 min |
| 3 | 0.008862 | 2 | 30 min |

It can easily be verified that all three components have a mean lifetime of 100 hours. Component 1 has a decreasing failure rate, component 2 a constant failure rate and component 3 an increasing failure rate.

For the Weibull distribution, the failure rate is

$$r(t) = a\lambda t)^{a-1}$$

Suppose all three components are new initially ($A_1 = A_2 = A_3 = 0$) and the system fails after 40 hours of operation. We can compute the probability that each component was the cause of failure by using (1)

$$r_1(40) = 0.5 \times 0.02 \times (0.02 \times 40)^{0.5-1} = 0.01118$$

$$r_2(40) = \qquad = 0.01$$

$$r_3(40) = \qquad = 0.006283$$

so that

$$p_1 = \frac{0.01118}{0.01118 + 0.01 + 0.006283} = 0.4071$$

$$p_2 = \qquad = 0.3641$$

$$p_3 = \qquad = 0.2288$$

Suppose that inspecting any of the three components always determines whether the component is functioning or failed ($\alpha_1 = \alpha_2 = \alpha_3 = 1$).

We can now see that component 1 is the most likely to be failed, but requires the longest inspection time. Using the corollary to find the order of inspection, we compute

$$\frac{T_1}{\alpha_1 p_1} = \frac{60}{1 \times 0.4071} = 147.38$$

$$\frac{T_2}{\alpha_2 p_2} = \qquad = 137.32$$

$$\frac{T_3}{\alpha_3 p_3} = \qquad = 131.11$$

Thus, according to the corollary, we should inspect the components in the order 3, 2, 1: if component 3 is not found failed we will next inspect component 2 and if it is functioning we will examine component 1.

As a further illustration, we will exhibit the computations necessary for this system assuming the corollary does not apply. If inspection of component 3 finds it functioning, we use (2) to recompute the values of $p_1, p_2, p_3$

$$p_1 = \frac{0.4071}{1 - 1 \times 0.2288} = 0.5279$$

$$p_2 = \qquad = 0.4721$$

$$p_3 = \frac{(1 - 1) \times 0.2288}{1 - 1 \times 0.2288} = 0$$

These probabilities reflect the fact that we now know component 3 is functioning.

We now find which component to inspect next.

$$\frac{T_1}{\alpha_1 p_1} = \frac{60}{1 \times 0.5279} = 113.65$$

$$\frac{T_2}{\alpha_2 p_2} = \frac{50}{1 \times 0.4721} = 105.90$$

$$\frac{T_3}{\alpha_3 p_3} = \frac{30}{1 \times 0} = \infty$$

Thus we next test component 2, and, if found functioning, we proceed to component 1.

## V. Conclusions and Suggestions for Further Work

We have presented a procedure for the determination of the sequence in which to test the components of a failed series system in minimal expected time. The procedure requires an estimate of the lifetime distributions of the different components and of their testing times, and requires a minimal amount of computation to determine the optimal order. Additional information required is the age of each component as well as the age of the system since it was last repaired.

This study suggests the following questions that remain to be analyzed: Under what conditions would it be preferable to replace a component by a new one without inspecting it for failure, and under what conditions would it be preferable to replace a component only after it has been tested and found defective? Furthermore, in what order should the replacements or tests of different components be carried out?

# Appendix A
# Probability Derivation

We now present a derivation of the probability that component 1 caused the system's failure.

If $A_i$ was the age of component $i$ when the last system repair was completed, then the probability it will function for an additional $t$ units of time is

$$\overline{G}_i(t) = \frac{\overline{F}_i(t + A_i)}{\overline{F}_i(A_i)}$$

Let $G_i(t) = 1 - \overline{G}_i(t)$ and $g_i(t) = (d/dt)\, G_i(t)$.

Assume $X_i$ is the remaining life of component $i$ after the last system repair, and let $X = \min\ \{X_i\}$ be the time between the last repair and the next failure of the system. Clearly

$$P[X_i > t] = \overline{G}_i(t) \quad i = 1, 2, \cdots n, \text{ and}$$

$$P[X > t] = \prod_{i=1}^{n} \overline{G}_i(t)$$

If it is known that the system failed between times $x$ and $x + \delta$, then the probability that component 1 caused the system's failure is

$$p_1(\delta) = P[X = X_1 \,|\, x < X \leqslant x + \delta]$$

$$= \frac{P[x < X_1 \leqslant x + \delta, X_2 > X_1, X_3 > X_1 \cdots X_n > X_1]}{P[x < X \leqslant x + \delta]}$$

The numerator can be computed by conditioning on the value of $X_1$, yielding

$$p_1(\delta) = \frac{\displaystyle\int_x^{x+\delta} g_1(t)\, \overline{G}_2(t) \cdots \overline{G}_n(t)\, dt}{\overline{G}_1(x)\, \overline{G}_2(x) \cdots \overline{G}_n(x) - \overline{G}_1(x + \delta)\, \overline{G}_2(x + \delta) \cdots \overline{G}_n(x + \delta)}$$

Dividing numerator and denominator by $\delta$ and letting $\delta$ go to 0 we obtain

$$p_1 = p[X = X_1 \,|\, X = x] = \frac{g_1(x)\, \overline{G}_2(x) \cdots \overline{G}_n(x)}{\displaystyle\sum_{k=1}^{n} g_k(x) \prod_{j \neq k} \overline{G}_j(x)}$$

and using the definition of $\overline{G}_i$ and $g_i$ this reduces to

$$p_1 = \frac{f_1(x + A_1) \displaystyle\prod_{j=2}^{n} \overline{F}_j(x + A_j)}{\displaystyle\sum_{k=1}^{n} f_k(x + A_k) \prod_{j \neq k} \overline{F}_j(x + A_j)} \qquad (3)$$

We now recall that $r_i(t) = f_i(t)/\overline{F}_i(t)$ to obtain equation (1):

$$p_i = \frac{r_i(x + A_i)}{\displaystyle\sum_{j=1}^{n} r_j(x + A_j)}$$

N82-11297



# An Error Budget for the JPL Hydrogen Maser Receiver

C. A. Greenhall
Communications Systems Research Section

*This report estimates the frequency instability and phase noise of the hydrogen maser receiver. The errors at each output are given as functions of the errors of the component modules of the receiver. The results are compared to the measured errors of the frequency standard (maser plus receiver).*

## I. Introduction

The hydrogen maser receiver is a synthesizer that converts the maser signal, at about 1420 MHz, to a set of output signals at 100, 20, 10, 5, 1, and 0.1 MHz. This report estimates the contribution of the receiver and its component modules to the frequency instability and phase noise of each of its outputs, and compares these results to published measurements of the 100-MHz output of the frequency standard, which consists of a maser plus its receiver. One can then assess how much the receiver degrades the performance of the frequency standard.

The block diagram of the receiver is shown in Fig. 1, which is extracted from Ref. 1 and modified slightly. It is a double-heterodyne phase-locked loop that provides its own mixing signals to the first and second mixers. The frequency conversion ratio of the receiver is tuned by setting the conversion ratio of the synthesizer that feeds one side of the phase detector. Let us believe that the maser is oscillating at $1420 + f_s$ MHz. If the synthesizer's front panel dials are set to read $f_s$ MHZ, then the loop forces the receiver outputs to have their advertised frequencies of 0.1 up to 100 MHz, relative to the $1420 + f_s$ MHZ that we believe the maser to have. The frequency $f_s$ can be set between 0.4 and 0.51 MHz in steps of $10^{-8}$ Hz.

In the following analysis, we presume to traverse the territory covered by R. Meyer in Ref. 1. There are a few differences. First, many of the component performances in Ref. 1 are merely requirements. Some of these we have replaced by the measured performances reported in a number of other references, to be cited later. Second, we have given a more thorough treatment of the effect of the distribution amplifier chain on the loop and on the receiver outputs. Finally, we have added one more (somewhat redundant) performance measure, the Allan variance for $\tau = 1$ second, because it is easy to compare it with measurements on the frequency standard.

## II. Phase Noise Breakdown

The aim here is to give the phase noises at the six outputs, (13) to (18) in Fig. 1, in terms of the phase noises contributed by the component modules. The sinewave signal at point (i) in the figure has phase noise $\phi_i$ at the indicated frequency (MHz). An exception is the baseband signal at (5). There are many sources of noise. If a module (call it $Z$) has one output, then the additive phase noise contributed by $Z$ to its output is called $n_Z$. If $Z$ has several outputs, at (i), (j), $\cdots$, then there are several noises $n_{Zi}$, $n_{Zj}$, $\cdots$ blamed on $Z$. All of the outputs of each distribution amplifier (DA) are electrically

separate, even though the figure shows the noises $n_{I20}$, $n_{I21}$, $n_{I22}$ as coming from the same point. Each of the modules $H$ through $L$ is a divider, cleanup filter, and DA in tandem.

The signal at ⑤, the phase detector output, is

$$K_D\,[A\sin(\phi_4 - \phi_{10} + n_D) + n_t]$$

where $K_D$ is the phase detector gain, and $A$ is the rms amplitude of the signal at ④. The noise $n_D$ is an exception to the above rules in that it is referred to the input of $D$ instead of the output. The term $n_t$ is the baseband equivalent of the thermal noise of the receiver.

The linearized equation of motion of the phase noise $\phi_6$ can now be written as

$$[1 + 14.204\,G(s)]\,\phi_6 = G(s)n + n_E$$

where the synthesizer frequency $f_s$ has been set to 0.4 MHz, and

$$G(s) = AK_DK_{VCO}F(s)/s\,,$$

$$n = n_t/A + n_A + n_B + n_C + n_D - n_M - n_{H11} - n_G$$

$$- 14n_{F7} - 0.2\,n_{F8} - 0.004\,n_{F9}\,. \qquad (1)$$

From now on, let us drop the term $0.004n_{F9}$ and replace 14.204 by 14.2. The closed-loop transfer function is

$$L(s) = \frac{14.2\,G(s)}{1 + 14.2\,G(s)}\,,$$

whose two-sided noise bandwidth is 100 Hz (Ref. 2). The phase noise $\phi_6$ can be written

$$\phi_6 = L(s)(n/14.2) + [1 - L(s)]\,n_E\,. \qquad (2)$$

Thus, $n/14.2$ appears inside the loop passband, and $n_E$ appears outside.

The output phase noises, presented as if they were multiplied up to 100 MHz, are

$$\begin{aligned}
&100\text{ MHz:} && \phi_{13} = \phi_6 + n_{F13}\\
&20\text{ MHz:} && 5\phi_{14} = \phi_6 + n_{F8} + 5n_{H14}\\
&10\text{ MHz:} && 10\phi_{15} = \phi_6 + n_{F19} + 10n_{I15}\\
&5\text{ MHz:} && 20\phi_{16} = \phi_6 + n_{F19} + 10n_{I20} + 20n_J\\
&1\text{ MHz:} && 100\phi_{17} = \phi_6 + n_{F19} + 10n_{I21} + 100n_K\\
&0.1\text{ MHz:} && 1000\phi_{18} = \phi_6 + n_{F19} + 10n_{I22} + 1000n_L
\end{aligned} \qquad (3)$$

## III. Phase Noise Data Types

We can now discuss Table 1, which contains most of the results of this report. The first data type is called $\Delta f/f$. It is the maximum change in relative frequency that occurs after the receiver environment undergoes a sudden 5°C change in temperature. The *output* column gives the estimated $\Delta f/f$ for each of the six receiver outputs, 100 MHz down to 0.1 MHz. The *standalone* column gives the $\Delta f/f$ for each component by itself. The *receiver* column gives the modular contributions that must be summed according to the arrows to give $\Delta f/f$ for the six outputs. Since some of the $\Delta f/f$ contributions may have different signs, this is a worst-case result.

To arrive at the receiver $\Delta f/f$ numbers, one has to account not only for the coefficients in Eqs. (1) – (3), but also for the frequencies at different places in the loop. For example, $\phi_6$ is at 100 MHz, $n_A$ is at about 1420 MHz, and $n_B$ is at 20.4 MHz. From Eqs. (1) and (2), since $L(0) = 1$,

$$\phi_6 = \frac{1}{14.2}\,(n_A + n_B) + \text{others},$$

$$\frac{\phi_6}{100} = \frac{n_A}{1420} + \frac{20.4}{1420}\frac{n_B}{20.4} + \text{others}\,.$$

This shows that

$$\text{receiver } \Delta f/f = \text{standalone } \Delta f/f \qquad \text{(module A)}$$

$$\text{receiver } \Delta f/f = (20.4/1420)\,(\text{standalone } \Delta f/f) \qquad \text{(module B)}\,.$$

The second data type is one-sided spectral density of phase, $S_\phi(f)$, evaluated at $f = 10$ Hz. The units are dB relative to 1 rad$^2$/Hz. This gives an idea of the relative sizes of the short-term phase noises. *The receiver and output $S_\phi$ values are referred to 100 MHz*, and the receiver values are summed according to the arrows. For the above example,

$$S_{\phi_6}(f) = \left(\frac{1}{14.2}\right)^2\left[S_{n_A}(f) + S_{n_B}(f)\right] + \text{others}\,.$$

This shows that the receiver values are $(1/14.2)^2$ (-23 dB) times the standalone values for those two modules. Note that the figure -100 dB for the 0.1-MHz output is the spectral density of $1000\phi_{18}$, not $\phi_{18}$, whose spectral density is -160 dB.

The third data type is Allan deviation (square root of Allan variance) at $\tau$ = 1 second, computed from the phase noise models used to get $S_\phi(f)$. Each output value is the rss of the receiver values indicated by the arrows.

## IV. The Bottom Line

The hydrogen maser is viewed only through the window of its receiver. The last line of Table 1 gives results of measurements made on the 100-MHz output of the combined system (Refs. 3, 4). These can be compared directly to our estimates of the instability of the receiver alone at its 100-MHz output. The thermal $\Delta f/f$ of the receiver (a worst-case estimate) is 0.52 times the measured system value, the receiver $S_\phi(10)$ is 6 dB below the measured system value, and the receiver $\sigma_y(1)$ is 0.43 times the measured system value.

Figure 2 extends the $\sigma_y(\tau)$ comparison over a wide range of $\tau$. The two upper curves are measurements of Allan deviation (Refs. 3, 5) of hydrogen maser frequency standards at 100 MHz. The lower curve is our estimate of the contribution of the receiver. It is composed mainly of thermal noise from the front end of the receiver and flicker phase modulation from the 100-MHz distribution amplifier. For $\tau < 30$ s the receiver noise is 3 dB to 7 dB below the total system noise. To put it another way, the receiver appears to degrade the short-term stability of the maser by 1 to 3 dB. For larger $\tau$, the degradation from random receiver noise is negligible.


## Acknowledgment

I would like to thank Roger Meyer for his patient assistance and the loan of his hydrogen maser bibliography.

**Table 1. Frequency instability and phase noise of JPL hydrogen maser receiver**

| Module | Standalone | $\Delta f/f$ for 5°C step Receiver | Output | $S_\phi$ (10 Hz) Standalone | MHz | Referred to 100 MHz Receiver | Output | $\sigma_y$ (1 s) Receiver | Output |
|---|---|---|---|---|---|---|---|---|---|
| Thermal noise | | | | −84 dB | 1420 | −107 dB | | 9E-14 | |
| | | | | | | ↓ | | ↓ | |
| A. Front end | 1E-14 | 1E-14 | | −124 dB | 1420 | −147 dB | | 1E-15 | |
| | | ↓ | | | | ↓ | | ↓ | |
| B. First IF | 7E-15 | 1E-16 | | −124 dB | 20.4 | −147 dB | | 1E-15 | |
| | | ↓ | | | | ↓ | | ↓ | |
| C. Second IF | 3.5E-15 | 1E-16 | | −124 dB | 0.406 | −147 dB | | 1E-15 | |
| | | ↓ | | | | ↓ | | ↓ | |
| D. Phase detector | 3.5E-13 | 1E-16 | | −124 dB | 0.406 | −147 dB | | 1E-15 | |
| | | ↓ | | | | ↓ | | ↓ | |
| M. ×14 | 1.8E-13 | 1.8E-13 | | −110 dB | 1400 | −133 dB | | 5.4E-15 | |
| | | ↓ | | | | ↓ | | ↓ | |
| G. Dana synthesizer | 3.4E-15 | 1E-18 | | −128 dB | 0.406 | −151 dB | | 6.6E-16 | |
| | | ↓ | | | | ↓ | | ↓ | |
| E. 100 MHz VCO | 2E-12 | <7E-17 | | −104 dB | 100 | <−147 dB | | 6.6E-15 | |
| | | ↓ | | | | ↓ | | ↓ | |
| F. 100 MHz DA | 3.7E-14 | 7.4E-14 | 2.6E-13 | −111 dB | 100 | −108 dB | −104 dB | 1E-13 | 1.3E-13 |
| | | ↓ | | | | ↓ | | ↓ | |
| H. ÷5, 20 MHz DA | 2.4E-13 | 3.3E-15 → | 2.4E-13 → 5E-13 | −121 dB | 20 | −144 dB → | −107 dB → −103 dB | 1.5E-15 → | 1.1E-13 → 1.7E-13 |
| | | ↓ | | | | ↓ | | ↓ | |
| I. ÷10, 10 MHz DA | 2.4E-13 | 2.4E-13 → | 7.5E-13 | −127 dB | 10 | −107 dB → | −101 dB | 1.1E-13 → | 2.1E-13 |
| J. ÷2, 5 MHz DA | 2.4E-13 | → | 2.4E-13 → 1E-12 | −133 dB | 5 | → | −107 dB → −100 dB | → | 1.1E-13 → 2.3E-13 |
| K. ÷10, 1 MHz DA | 2.4E-13 | → | 2.4E-13 → 1E-12 | −147 dB | 1 | → | −107 dB → −100 dB | → | 1.1E-13 → 2.3E-13 |
| L. ÷100, 0.1 MHz DA | 2.4E-13 | → | 2.4E-13 → 1E-12 | −167 dB | 0.1 | → | −107 dB → −100 dB | → | 1.1E-13 → 2.3E-13 |
| H maser + receiver, measured at 100 MHz | | | 5E-13 | | | | −98 dB | | 3E-13 |

**Table 2. Breakdown of modules H-L**

| | $\Delta f/f$ 5°C step $\tau$ = 2000 s | $S_\phi$(10 Hz) $f_0$ = 1000 MHz | $\sigma_y$(1 s) |
|---|---|---|---|
| Frequency divider | 1.9E-13 | -114 dB | 4.8E-14 |
| Cleanup filter | 9.3E-15 | -111 dB | 6.8E-14 |
| Distribution amplifier | 3.7E-14 | -111 dB | 6.8E-14 |
| Total | 2.4E-13 sum | -107 dB sum | 1.1E-13 rss |

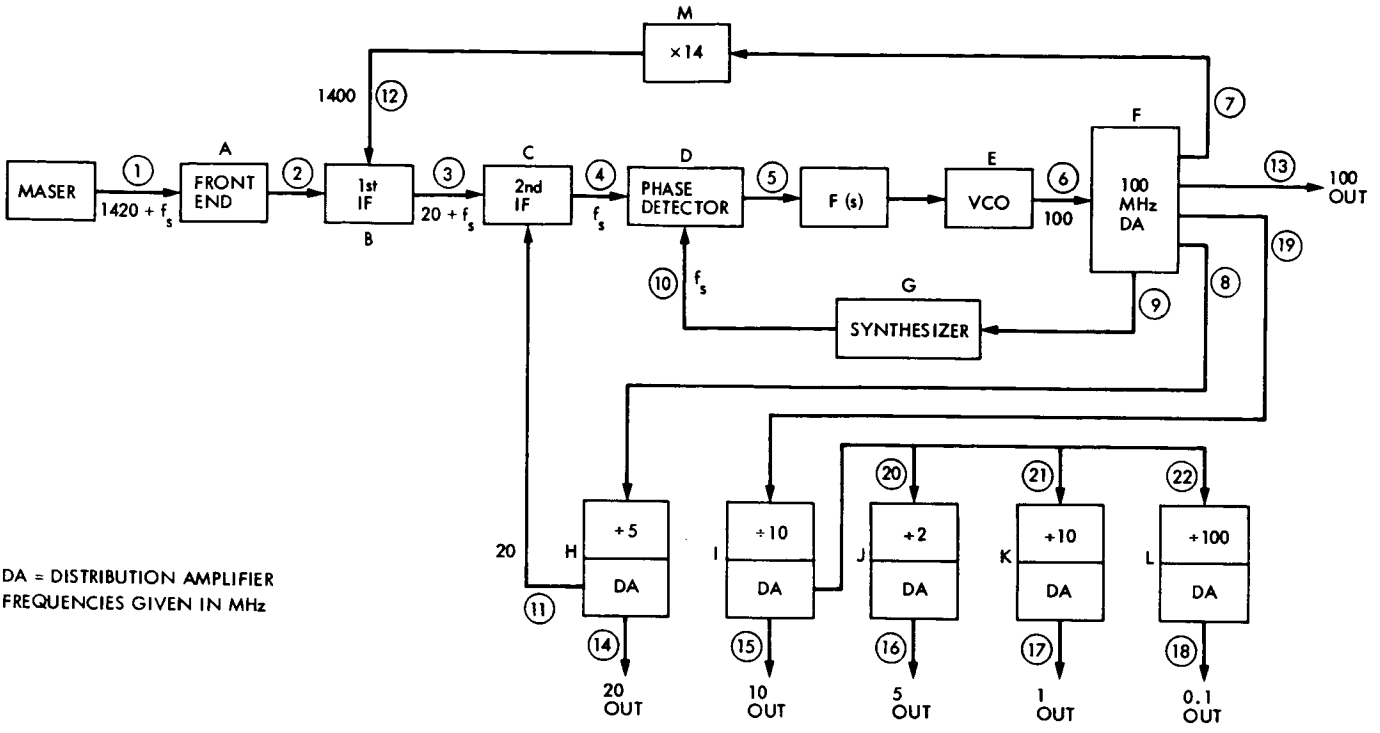



**Fig. 1. Hydrogen maser receiver block diagram**

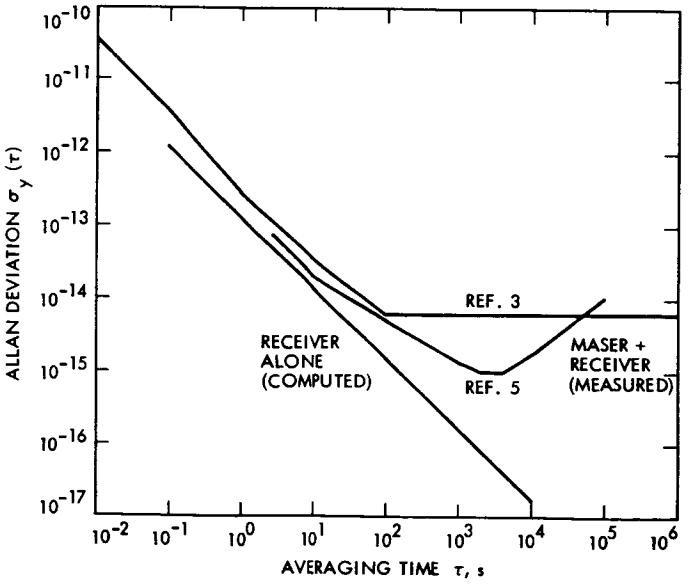



**Fig. 2. Frequency stability of hydrogen maser receiver and total system at 100 MHz**

# Appendix A

# Discussion of Receiver Modules

*Thermal noise.* (This has nothing to do with the $\Delta f/f$ caused by a thermla step.) According to Ref. 2, the receiver loop has SNR $\rho$ = 64 dB in $2b_L$ = 100 Hz. If $N_0$ is the one-sided spectral density of the thermal noise $n_t$, then $N_0/A^2 = 2b_L/\rho$ = -84 dB, which is the standalone value of $S_\phi(0)$ ($=S_\phi(10)$). In effect, this pertains to the 1420-MHz signal. According to Eq. (2), the receiver value of $S_\phi(0)$, which pertains to a 100-MHz signal, is $(14.2)^2$ = 23 dB below the standalone value. Allan deviation can be obtained (Ref. 6) by

$$\sigma_y(\tau) = \frac{\sqrt{3S_\phi(0)b_L}}{2\pi f_0 \tau} \qquad \text{(A-1)}$$

$$= \frac{8.7 \times 10^{-14}}{\tau} \ .$$

*Front end* (module A). The noise $n_A$ is amplifier phase jitter, as opposed to additive thermal noise. The figures for $\Delta f/f$ and $S_\phi(10)$ are merely requirements from Ref. 1. Amplifiers are observed to exhibit flicker phase modulation (Ref. 7), which means that $S_\phi(f)$ = const/$f$ for $0 < f < f_h$, where $f_h$ is a cutoff frequency, set to 100 Hz here. The Allan deviation is then equal to

$$\sigma_y(\tau) = \frac{2.42\ \sqrt{S_\phi(10)}}{f_0 \tau}\ \sqrt{1 + 0.13 \ln \tau} \qquad \text{(A-2)}$$

(Ref. 6, corrected in a memorandum of D. W. Allan.)

*First IF* (module B), *second IF* (module C), *phase detector* (module D). Same remarks as for module A. The $\Delta f/f$ and $S_\phi(10)$ values are requirements from Ref. 1, and flicker PM is assumed.

*X14 multiplier* (module M). The $\Delta f/f$ value is from Ref. 8. Reference 9 gives $S_\phi(f) = 10^{-10}/f$ for $1 \leq f \leq 100$ Hz. The receiver value is 23 dB lower. Equation (A-2) gives $\sigma_y(\tau)$.

*Dana synthesizer* (module G). The $\Delta f/f$ is scaled from the value given in Ref. 10 for a 25°C step. To get the receiver value, multiply by 0.4/1420. The $S_\phi(10)$ value is scaled from the value given in Ref. 11 for a Dana synthesizer at 40 MHz instead of 0.4 MHz. Since this is also flicker PM. Eq. (A-2) gives $\sigma_y(\tau)$.

*VCO* (module E). The VCO phase noise appears outside the one-sided loop passband (50 Hz). We assume it to be pure flicker frequency modulation, $S_\phi(f) = c/f^3$, such that for all $\tau$ of interest, $\sigma_y(\tau) = 2 \times 10^{-12}$, the requirement from Ref. 1. This number appears under $\Delta f/f$ in Table 1, although it has nothing to do with a temperature step. From Ref. 6 we can deduce $c = f_0^2\ \sigma_y^2/\ln 4$, where $f_0$ = 100 MHz, and thence the standalone $S_\phi(10)$. Since 10 Hz is well within the loop passband, we have not attempted to compute the receiver $S_\phi(10)$. To obtain the *receiver* $\sigma_y(1)$, one realizes that the VCO phase noise, highpass filtered by the loop, is stationary, with variance approximately equal to

$$\sigma_E^2 = \int_{50}^{\infty} \frac{c}{f^3}\, df = 5.8 \times 10^{-12}$$

One can then use Eq. (A-1) to get $\sigma_y(\tau)$ by substituting $\sigma_E^2$ for $S_\phi(0)b_L$.

*100-MHz distribution amplifier* (module F). Let us first remark that the contribution of this amplifier to all outputs (multiplied to 100 MHz) is about

$$-\frac{14}{14.2}\ n_{F7} + n_{Fi}$$

where $i$ = 13, 8, or 19. This is why the receiver values of $\Delta f/f$ and $S_\phi(10)$ are double the standalone values. Reference 12 gives a phase change 0.08 deg/°C at 10 MHz; this scales to $\Delta\phi$ = 4 deg at 100 MHz for a 5°C change. Following R. Meyer's advice, we computed

$$\frac{\Delta f}{f} = \frac{2}{3}\ \frac{\Delta\phi}{2\pi f_0 \tau}\ , \qquad \text{(A-3)}$$

with $f_0$ = 100 MHz, $\tau$ = 2000 s. The idea is that the phase response to the temperature step has a time constant about 2000 s. This is a rough estimate, of course. For $S_\phi(f)$ we use

the measurements of Meyer and Sward (Ref. 7), which, scaled to 100 MHz, give $S_\phi(f) = 8 \times 10^{-11}/f$. Equation (A-2) gives $\sigma_y(\tau)$.

*Divider – cleanup filter – DA* (modules H-L). The performance breakdown of these combinations is given in Table 2, which, as usual, assumes that the output has been multiplied to 100 MHz. Flicker PM is assumed for all components. Sources are Refs. 7 and 13 for the divider, Ref. 14 for the cleanup filter, and Refs. 7 and 12 for the DA. There are two receiver entries in Table 1 for the 20-MHz module *H* because this module feeds backward into the loop and also forward into its own output.

# References


1. Meyer, R. F., "Hydrogen Maser Frequency Standard: Receiver Configuration and Stability Requirements," *The Deep Space Network Progress Report*, Technical Report 32-1526, Vol XVIII, pp. 66-72. Jet Propulsion Laboratory, Pasadena, Calif., December 15, 1973.

2. Finnie, C., Sydnor, R., and Sward, A., "Hydrogen Maser Frequency Standard," *Proceedings of the 25th Annual Frequency Control Symposium*, pp. 348-351. U.S. Army Electronics Command, Fort Monmouth, New Jersey, April 28, 1971.

3. Sward, A., "Measurement of the Power Spectral Density of Phase of the Hydrogen Maser," *JPL Quarterly Technical Review*, Vol. I, Number 4, pp. 30-33. Jet Propulsion Laboratory, Pasadena, Calif., January 1972.

4. Kuhnle, P. F., "Hydrogen Maser Implementation in the Deep Space Network at the Jet Propulsion Laboratory," Proc. 11th Annual PTTI Applications & Planning Meeting, Goddard Space Flight Center, Greenbelt, Md., Nov. 1979.

5. Greenhall, C. A., "Removal of Drift from Frequency Stability Measurements," *TDA Progress Report 42-65*, Jet Propulsion Laboratory, Pasadena, Calif., 1981.

6. Barnes, J. A., et al., Characterization of Frequency Stability, NBS Technical Note 394, National Bureau of Standards, Washington, D.C., 1970.

7. Meyer, R., and Sward, A., "Frequency Generation and Control: The Measurement of Phase Jitter," *The Deep Space Network, Space Programs Summary 37-64*, Vol. II, pp. 55-58. Jet Propulsion Laboratory, Pasadena, Calif., August 31, 1970.

8. MacConnell, J., Meyer, R., "L-Band Frequency Multipliers: Phase Noise, Stability, and Group Delay," *The Deep Space Network Progress Report*, Technical Report 32-1526, Vol. X, pp. 104-109. Jet Propulsion Laboratory, Pasadena, Calif., August 15, 1972.

9. Lutes, G., MacConnell, J., and Meyer, R., "Hydrogen Maser: Low Phase Noise, L-Band Frequency Multiplier," *The Deep Space Network Progress Report*, Technical Report 32-1526, Vol. VII, pp. 81-83. Jet Propulsion Laboratory, Pasadena, Calif., February 15, 1972.

10. Sward, A., "New Developments in the Hydrogen Maser Frequency Standard," *The Deep Space Network Progress Report*, Technical Report 32-1526, Vol. II, pp. 72-74. Jet Propulsion Laboratory, Pasadena, Calif., April 15, 1971.

11. Greenhall, C. A., Internal Noise of a Phase-Locked Receiver with a Loop-Controlled Synthesizer, *DSN Progress Report 42-52*, pp. 41-50, Jet Propulsion Laboratory, Pasadena, Calif., 1979.

12. Lutes, G., "Frequency Generation and Control: Distribution Amplifiers for the Hydrogen Maser Frequency Standard," *The Deep Space Network, Space Programs Summary 37-61*, Vol. II, pp. 68-72. Jet Propulsion Laboratory, Pasadena, Calif., January 31, 1970.

13. Lutes, G., "Improved Frequency Dividers," *The Deep Space Network Progress Report*, Technical Report 32-1526, Vol. II, pp. 56-58. Jet Propulsion Laboratory, Pasadena, Calif., April 15, 1971.

14. Lutes, G. F., "Phase-Stable, Low-Phase-Noise Filters for Reference Signals," *The Deep Space Network Progress Report*, Technical Report 32-1526, Vol. XII, pp. 44-46. Jet Propulsion Laboratory, Pasadena, Calif., December 15, 1972.

N82-11298



# Removal of Drift From Frequency Stability Measurements

C. A. Greenhall
Communications Systems Research Section

*This article gives a method of estimating frequency drift rate and removing its effect from Allan variance plots. When tried on a test of hydrogen masers, the method gives consistent results. An error in the previous Allan variance computation algorithm is corrected.*

## I. Drift Removal—Before and After

Imagine a frequency standard whose only problem is a steady frequency drift. Its phase error in radians has the form

$$\phi(t) = 2\pi f_0 \left(a + bt + \frac{1}{2} ct^2\right) \tag{1}$$

where $a$, $b$, and $c$ are constants, and $f_0$ is the nominal frequency of the oscillation. The drift rate of the relative frequency error $\Delta f/f_0 = \dot{\phi}(t)/(2\pi f_0)$ is $c$ per second. The two-sample Allan deviation, the usual measure of $\Delta f/f_0$, is

$$\sigma_y(\tau) = \frac{|c|\tau}{\sqrt{2}} \tag{2}$$

for this simple case. (In this report, Allan *deviation* is the square root of Allan *variance*.)

Now observe the behavior of three hydrogen masers, called DSN1, DSN2, and DSN3, which were tested at the JPL Interim Frequency Standard Test Facility for eight days at the end of 1980. Figure 1, a rough rendering of the frequency strip charts, shows $\Delta f/f_0$ vs calendar date for the three possible pairs of masers beating against each other. *The three curves have different $\Delta f/f_0$ scales.* In particular, the scale of the DSN2-DSN3 curve is expanded relative to the others because it does not exhibit the drift that dominates the other curves. The thickness of the DSN2-DSN3 curve is just a way of showing the size of the rapid (3-minute average) fluctuations of $\Delta f/f_0$. It is a good bet that DSN1 was drifting by itself at a rate about $-6 \times 10^{-19}$ per second, or $-5 \times 10^{-14}$ per day.

Figure 2 shows what this drift does to the Allan deviation. The usual $\sigma_y(\tau)$ is given by the "gross" curves, which, for the two pairs containing DSN1, become straight lines with slope one for the larger $\tau$. The dashed lines show the estimated drift component, Eq. (2), where the estimate of the drift rate $c$ is computed by a method explained below. The actual estimates of $c$ are

$6.15 \times 10^{-19}$/s DSN2-DSN1

$-6.32 \times 10^{-19}$/s DSN1-DSN3

$-4.13 \times 10^{-21}$/s DSN2-DSN3

all of which have standard deviation $3.6 \times 10^{-20}$. Evidently, DSN2-DSN3 has negligible drift.

The "net" curves in Fig. 2 show what happens when the estimated drift function $(2\pi f_0)\ (ct^2/2)$ is subtracted (in effect) from the phase data. The net Allan deviations, $\sigma_{y0}(\tau)$, for the DSN1 pairs look like the gross Allan deviation of DSN2-DSN3. All three curves have slope 0.77 for $\tau > 10^4$s. The effects of the random phase fluctuations, formerly masked by the drift, can now be seen.

## II. Method of Drift Estimation and Removal

### A. Quantities to be Estimated

It is convenient to work with the function

$$x(t) = \phi(t)/(2\pi f_0)$$

where $\phi(t)$ is the phase difference of the pair of oscillators being tested. The underlying assumption is that $x(t)$ is a mean-continuous stochastic process whose second differences

$$\Delta_\tau^2 x(t) = x(t) - 2x(t-\tau) + x(t-2\tau)$$

are stationary for each $\tau$. A deterministic example is given by Eq. (1); its second differences

$$\Delta_\tau^2\left(\frac{1}{2}ct^2\right) = c\tau^2$$

are constant. In fact, it is true in general that any such process can be written

$$x(t) = \frac{1}{2}ct^2 + x_0(t) \tag{3}$$

where $c$ is a constant, and the second differences of $x_0(t)$ have mean zero (Ref. 1). This decomposes the phase into a pure frequency drift term plus random fluctuations. (The term $x_0(t)$ might contain an $a + bt$ component, which goes away when second differences are taken. Anyhow, we do not care about constant phase and frequency offsets.) Our goal is to perform this decomposition on experimental phase data.

The usual Allan variance, called *gross AV* here, is defined by

$$\sigma_y^2(\tau) = \frac{1}{2\tau^2} E\ [\Delta_\tau^2 x(t)]^2 \tag{4}$$

and the *net AV* is defined by

$$\sigma_{y0}^2(\tau) = \frac{1}{2\tau^2} E\ [\Delta_\tau^2 x_0(t)]^2 \tag{5}$$

Since the second differences of $x_0(t)$ have mean 0, Eq. (3) gives

$$c = \frac{1}{\tau^2} E\ \Delta_\tau^2 x(t) \tag{6}$$

$$\sigma_{y0}^2(\tau) = \sigma_y^2(\tau) - \frac{1}{2}c^2\tau^2 \tag{7}$$

We want to estimate $\sigma_y(\tau)$, $c$, and $\sigma_{y0}(\tau)$.

### B. The Estimators

Let $x(t)$ be given for $0 \leq t \leq T$. We shall need the first four moments of the second differences. Define

$$m_i(\tau) = \frac{1}{r}\sum_{j=2}^{r+1} [\Delta_\tau^2 x(j\tau)]^i$$

The integer $r$, which depends on $\tau$, is the available number of second-difference samples. One can define it by saying that $(r+1)\tau$ is the largest multiple of $\tau$ that does not exceed $T$.

The usual estimator of gross *AV* is the time average

$$S^2(\tau) = \frac{1}{2\tau^2} m_2(\tau) \tag{8}$$

How shall we estimate $c$? Equation (6) suggests the unbiased estimator

$$\hat{c}(\tau) = \frac{1}{\tau^2} m_1(\tau) \tag{9}$$

Because a second difference is the difference of first differences, the implied summation in Eq. (9) telescopes, leaving us with

$$\hat{c}(\tau, \tau') = \frac{1}{\tau'}\left[\frac{\Delta_\tau x(\tau+\tau')}{\tau} - \frac{\Delta_\tau x(\tau)}{\tau}\right] \tag{10}$$

where $\tau' = r\tau$. The notation is expanded because Eq. (10) is more general than Eq. (9), in that $\tau'$ does not need to be an

integer multiple of $\tau$. The interpretation is that the average drift rate equals average frequency near the end of the record, minus average frequency near the beginning, divided by the length of the record (actually, by $\tau'$).

We wish to select just one estimator of $c$ for the given record length $T$. To do this, we might minimize the variance of Eq. (9) or Eq. (10) over $\tau$, where $\tau' = T - \tau$ in Eq. (10). This cannot be done in advance without knowing the spectrum $S_x(f)$ of $x(t)$. Since $S_x(f)$ determines Allan variance, we are asking for the outcome of our measurements before we do them. To escape this trap, we appeal to the past – a measurement made by Sward (Ref. 2) on hydrogen masers. His work gives the one-sided spectral density

$$S_x(f) = \frac{h_{-1}}{(2\pi)^2 f^3} \quad \text{(flicker frequency modulation)}$$

$$+ \frac{h_1}{(2\pi)^2 f} \quad \text{(flicker phase modulation)}$$

where

$$h_{-1} = 3.5 \times 10^{-29}, h_1 = 1.6 \times 10^{-25}\ \text{s}^2$$

and the second term is cut off at $f = 10^5$ Hz.

It turns out that the variance of $\hat{c}(\tau, \tau')$ can be read from formulas in Ref. 3, pp. 42-47, for different types of phase noise, including flicker FM and PM. Using the Sward spectrum, we find, for $\tau < 15$ s, that the flicker PM part of Var $\hat{c}(\tau, \tau')$ is dominant. For $\tau > 15$s, the flicker FM part takes over. Furthermore, as $\tau$ increases beyond 15 s, the variance becomes smaller than for any $\tau < 15$ s. Hence, we need only consider the flicker FM contribution to Var $\hat{c}(\tau, \tau')$, which is

$$\frac{h_{-1}}{\tau'^2}\left[(r+1)^2 \ln(r+1) - 2r^2 \ln r + (r-1)^2 \ln(r-1)\right] \tag{11}$$

Let us state the result of minimizing this.

*Assume that the normalized phase error $x(t)$ consists of flicker FM plus a constant frequency drift term $(ct^2/2)$. Then Eq. (10) gives a family of unbiased esimators of $c$. Let $T = \tau + \tau'$ be fixed. Then*

$$\min_{\tau} \text{Var}\, \hat{c}(\tau, T-\tau) = \frac{8.94\, h_{-1}}{T^2} \tag{12}$$

*The minimum is achieved for $r = (T - \tau)/\tau = 5.29$.*

In other words, *we should use a $\tau$ that is about one-sixth of the record length $T$.* Notice that the variance is like $1/T^2$ instead of $1/T$. This happens because the second differences of $x(t)$ have less power than white noise near zero frequency.

The minimum is broad enough to allow considerable departures from it. Although Eq. (10) is simple, the details of the data processing make it expedient to revert to the summed form Eq. (9). Moments of the second differences are accumulated only for a certain small set of $\tau$, and the actual estimator of $c$ is

$$C = \hat{c}(\tau_c)$$

where $\tau_c$ is the largest available $\tau$ such that $r_c = r(\tau_c)$ is at least 6. Of course, $\tau_c + \tau'$ is usually less than $T$. For the $\tau$-set actually used, $r_c$ falls between 6 and 16.

We can estimate Var $C$ from Eq. (12) (or Eq. (11) if $\tau_c$ is not quite optimal) if we have a value for the flicker FM constant $h_{-1}$, which satisfies

$$h_{-1} = \frac{1}{\ln 4}\sigma_{y0}^2 \tag{13}$$

(Ref. 3). This leads to the next goal, the estimation of $\sigma_{y0}^2(\tau)$. In view of Eq. (7), one might use the estimate

$$S^2(\tau) - \frac{1}{2}C^2\tau^2$$

The problem with this is that it can be negative. We prefer to start from Eq. (5). Given the data $x(t)$ and the estimate $C$ of $c$, an estimate of the "net data" $x_0(t)$ is $x(t) - (Ct^2/2)$ (except for a polynomial $a + bt$). Then, an estimate of $\Delta_\tau^2 x_0(t)$ is $\Delta_\tau^2 x(t) - C\tau^2$. This leads to our estimator of choice,

$$S_0^2(\tau) = \frac{1}{2\tau^2 r}\sum_{j=2}^{r+1}\left[\Delta_\tau^2 x(j\tau) - C\tau^2\right]^2$$

$$= S^2(\tau) - C\,m_1(\tau) + \frac{1}{2}C^2\tau^2 \tag{14}$$

for the net $AV\ \sigma_{y0}^2(\tau)$.

Since $C\tau_c^2 = m_1(\tau_c)$, $S_0^2(\tau_c)$ is just $\left(1/(2\tau_c^2)\right)$ times the variance of the sequence of second $\tau_c$-differences of $x(t)$. One can now estimate $h_{-1}$ from Eq. (13) by using $S_0^2(\tau_c)$ for $\sigma_{y0}^2$.

For the largest useful $\tau$, corresponding to $r = 5$ or less, an anomaly may appear. Either the estimated net $AV$ or the estimated drift contribution $C^2\tau^2/2$ can come out greater than the estimated gross $AV$. This should not be alarming. If net $AV$ is greater than gross $AV$, the gross $AV$ should still fall within the error bar of the net $AV$. This error bar is the next topic.

The net $AV$ estimate $S_0^2(\tau)$ is the average of the numbers

$$u_j = \frac{1}{2\tau^2}\left[\Delta_\tau^2 x(j\tau) - C\tau^2\right]^2$$

A computation with Gaussian flicker FM shows that the sampled process

$$\left[\Delta_\tau^2 x_0(j\tau)\right]^2$$

is almost white from zero frequency to the Nyquist frequency, even though the spectral density of $\Delta_\tau^2 x_0$ vanishes at zero frequency. Thus, it is reasonable to use the sample variance of the $u_j$ for estimating the error in the mean. Our one-sigma error estimate for $S_0^2(\tau)$ is $\delta$, given by

$$\delta^2 = \frac{1}{(r-1)r}\sum_{j=2}^{r+1}\left[u_j - S_0^2(\tau)\right]^2 \tag{15}$$

$$= \frac{1}{4\tau^4(r-1)}\left[m_4 - 4Dm_3 + 4D^2 m_2 - (2Dm_1 - m_2)^2\right]$$

where $D = C\tau^2$, and the $m_i$ are the moments $m_i(\tau)$ defined above. Finally, an error bar for $S_0(\tau)$ is $[S_0^2(\tau) - \delta]^{1/2}$ to $[S_0^2(\tau) + \delta]^{1/2}$. If $\delta > S_0^2(\tau)$, then the first number is replaced by zero.

## III. Correction of an Error

In the JPL frequency stability test setup, the phase data are written on magnetic tape. Later, they are processed into Allan variance by an offline computer program. The previous version of this program contains an error. To explain it, fix a $\tau$, and let

$$\xi(t) = \frac{1}{\tau\sqrt{2}}\frac{\Delta_\tau^2\phi(t)}{2\pi f_0} = \frac{\Delta_\tau^2 x(t)}{\tau\sqrt{2}}$$

be a scaled version of the second difference of phase. According to Eq. (8), the usual estimator of $\sigma_y(\tau)$ is

$$S(\tau) = \left[\frac{1}{r}\sum_{j=2}^{r+1}\xi^2(j\tau)\right]^{1/2} \tag{16}$$

a discrete rms time average of $\xi$, just as

$$\sigma_y(\tau) = \left[E\,\xi^2(t)\right]^{1/2} = \|\xi(t)\|_2$$

is the rms ensemble average, or $L^2$ norm, of the random variable $\xi(t)$.

The previous Allan variance program used

$$S_1(\tau) = \frac{1}{r}\sum_{j=2}^{r+1}|\xi(j\tau)| \tag{17}$$

to estimate the Allan deviation. Equation (17) is an unbiased estimator, not of $\sigma_y(\tau)$, but of

$$E\,|\xi(t)| = \|\xi(t)\|_1$$

the $L^1$ norm of $\xi(t)$. One can judge the size of this error by assuming that $\xi(t)$ is Gaussian with mean zero (no drift!), in which case

$$\frac{\|\xi(t)\|_1}{\|\xi(t)\|_2} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}|x|\,e^{-\frac{1}{2}x^2}\,dx = (2/\pi)^{1/2} = 0.7979$$

We compared the old values $S_1(\tau)$ and the new values $S(\tau)$ from another stability test run, which again measured three oscillator pairs. Figure 3 shows the comparison for one pair, a cesium standard and a hydrogen maser. For all three pairs, we computed the average of the ratios $S_1(\tau)/S(\tau)$. A ratio was included only if the standard deviation of $S_1(\tau)$ was less than 3 percent. The three averages, with error estimates, are

$$0.7992 \pm 0.0019$$

$$0.7905 \pm 0.0008$$

$$0.7968 \pm 0.0015$$

At least for this purpose, the Gaussian hypothesis seems justified. When drift is negligible, the old values for $\sigma_y(\tau)$ are 20 percent too low. This explains part of the difference between the JPL system and the Hewlett-Packard HP5390A frequency stability measurement system (Ref. 4), which does its own Allan variance computation.

It was necessary to correct this error before a drift-removal algorithm could successfully be installed in the JPL Allan variance program. Use of the $L^2$ norm makes it possible to remove the drift via Eq. (14) in one pass through the phase data.

## IV. Concluding Remarks

The drift removal method given above yields consistent results on the data from one frequency stability test run, in that the Allan deviation curves of the three hydrogen maser pairs look almost the same after the drift is removed by the analysis program. Although the drift estimation method assumes flicker FM, the actual Allan deviation plots do not become level at the higher $\tau$. We need more experience with the method before we can judge its robustness with respect to the flicker FM assumption. Perhaps one can find a method that tailors itself to the actual oscillator behavior. In the meantime, the present method appears to give useful results.

## References


1. Yaglom, A. M., "Correlation Theory of Processes With Random Stationary nth Increments," *American Math. Society Translations*, Ser. 2, Vol. 8, pp. 87-141, 1958.
2. Sward, A., "Measurement of the Power Spectral Density of Phase of the Hydrogen Maser," *JPL Quarterly Technical Review*, Vol. 1, No. 4, pp. 30-33, 1972.
3. Barnes, J. A., et al., *Characterization of Frequency Stability*, NBS Technical Note 394, National Bureau of Standards, Washington, D.C., 1970.
4. Sosa, E. N., "Station Stability Measurement," *TDA Progress Report 42-62*, pp. 35–42, Jet Propulsion Laboratory, Pasadena, Calif., 1981.

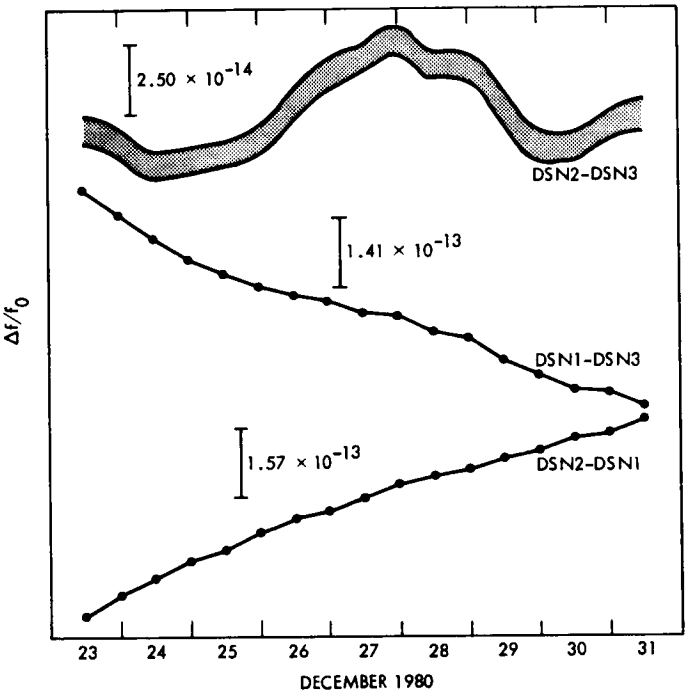



**Fig. 1. A plot of pairwise $\Delta f/f_0$ vs. time for a set of three hydrogen masers, called DSN1, DSN2, and DSN3. The $\Delta f/f_0$ scales are all different**

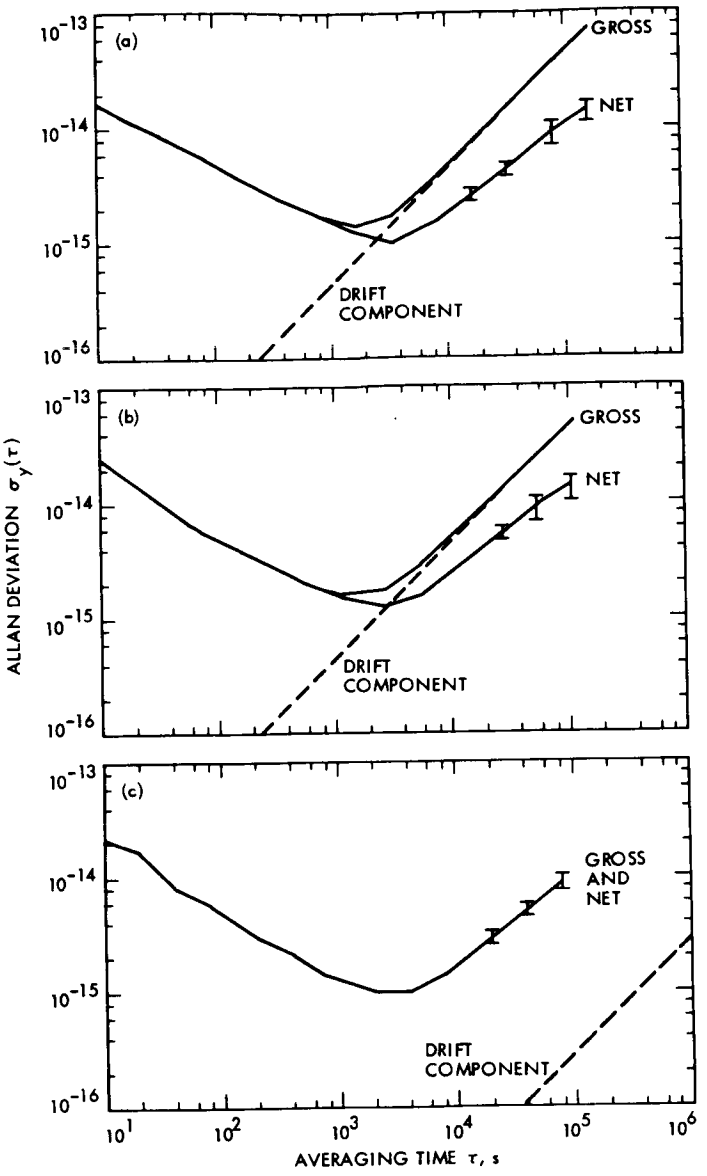



**Fig. 2. Pairwise Allan deviation of three hydrogen masers, before and after removal of drift from the measurements: (a) DSN2–DSN1; (b) DSN1–DSN3; (c) DSN2–DSN3**

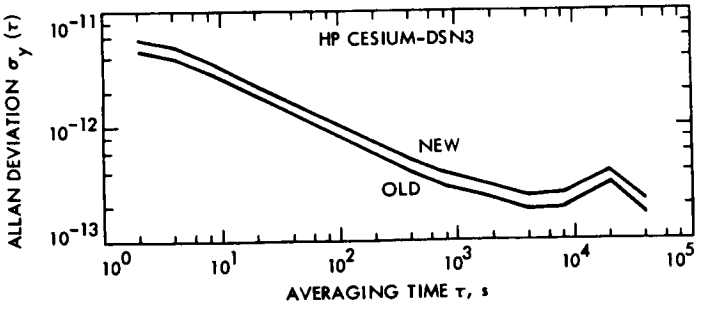



**Fig. 3 Allan deviation of a cesium-hydrogen maser pair, as computed by the old (incorrect) algorithm and the new (correct) algorithm. The old results are 20 percent too low.**

N82-11299



# Closed Form Evaluation of Symmetric Two-Sided Complex Integrals

R. Winkelstein
Communications Systems Research Section

*Evaluation of two-sided complex integrals of the form*

$$I_n = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} dS\, G(S)G(-S)$$

*and*

$$I_n = \frac{1}{2\pi i}\oint_{\substack{Unit\\Circle}} \frac{dZ}{Z}\, G(Z)G(Z^{-1})$$

*is often required when analyzing linear systems to determine signal variances resulting from stochastic inputs and system noise bandwidths. Presented are algebraic solutions of both the above integrals in a closed matrix equation form using coefficients of the numerator and denominator polynomials of the function G.*

## I. Introduction

In the analysis of linear systems for output and internal signal variance caused by noisy input signals, and for the analysis of noise bandwidths of such systems, the following integrals often require evaluation.

$$I_n = \frac{1}{2\pi i}\oint_{\substack{Unit\\Circle}} \frac{dZ}{Z}\, G(Z)G(Z^{-1}) \qquad (1)$$

$$I_n = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} dS\, G(S)G(-S) \qquad (2)$$

These integrals are two-sided integrals in the complex plane. Integral (2) arises in the analysis of continuous systems where $S$ is the Laplace transform complex frequency. $G(S)$ is the ratio of two polynomials in $S$, with the denominator polynomial of degree $n$ being at least 1 degree higher than the numerator polynomial. The zeros of the denominator polynomial are known as the poles of $G(S)$ and are assumed to be located in the left side of the complex $S$-plane.

Integral (1) arises in the analysis of sampled data systems where $Z$ is equal to $e^{ST}$, $T$ is the sampling time, and $e$ is the base of the natural logarithm. $G(Z)$ is the ratio of two polynomials in $Z$, with the denominator polynomial of degree $n$ being equal to, or greater than, the degree of the numerator polyno-

mial. The poles of $G(Z)$ are assumed to lie within the unit circle of the $Z$ complex plane.

Integrals in the complex plane of polynomial ratios are normally evaluated by factoring the denominator polynomial to determine the poles, and then summing the residues of the poles within the contour of integration. For the integrals (1) and (2), the residue method is highly laborious and essentially numeric in nature. However, advantage may be taken of the symmetry and assumptions in (1) and (2) to obtain an algebraic evaluation in matrix form using the coefficients of the polynomials of $G$.

## II. Sampled Data Systems

The function $G$ of Eq. (1) is given as

$$G(Z) = \frac{\sum_{i=0}^{n} b_{n-i} Z^i}{\sum_{i=0}^{n} a_{n-i} Z^i} \qquad a_0 \neq 0 \tag{3}$$

The solution system of equations as derived in Ref. 1 is

$$\sum_{i=0}^{n} (a_{i-r} + a_{i+r}) M_i = B_r$$

$$r = 0, 1, \ldots, n \tag{4}$$

$$M_0 = a_0 I_n$$

$$B_r = \begin{cases} \sum_{i=0}^{n} b_i^2 & r = 0 \\ 2 \sum_{i=0}^{n-r} b_i b_{i+r} & r = 1, 2, \ldots, n \end{cases}$$

$$a_m = 0 \begin{cases} m < 0 \\ m > n \end{cases}$$

Table 1 lists the matrix equations obtained from Eq. (4) for values of $n$ from 1 to 4. The algebraic value of $I_n$ may then be obtained using Cramer's rule (Ref. 2).

## III. Continuous Systems

The function $G$ of Eq. (2) is given as

$$G(S) = \frac{\sum_{i=0}^{n-1} b_{n-i} S^i}{\sum_{i=0}^{n} a_{n-i} S^i} \qquad a_0 \neq 0 \tag{5}$$

The solution system of equations is derived in Appendix A using a method similar to the derivation of Eq. (4) in Ref. 1.

$$\sum_{i=0}^{n-1} a_{n-2r+i} M_i = (-1)^r b_{n-r}^2 + 2 \sum_{k=1}^{\infty} (-1)^{r+k} b_{n-r-k} b_{n-r+k}$$

$$r = 0, 1, \ldots, n-1 \tag{6}$$

$$M_{n-1} = (-1)^{n-1} 2 a_0 I_n$$

$$a_m = 0 \qquad m < 0, m > n$$

$$b_m = 0 \qquad m < 1, m > n$$

Table 2 lists the matrix equations obtained from Eq. (6) for values of $n$ from 1 to 5. Similarly to the case of sampled data systems, $I_n$ may be obtained through the use of Cramer's rule.

## IV Conclusion

Algebraic closed matrix forms have been presented for the evaluation of integrals (1) and (2). The closed forms provide the possibility of obtaining some insight into parameter sensitivity in addition to greatly reducing the computational complexity required by the normal method of evaluation by residues.

# References

1. Jury, E. I., *Theory and Application of the Z-Transform Method*, Wiley, New York, 1964.

2. Apostol, T. M., *Calculus*, Vol. II, Wiley, New York, 1969.

**Table 1. Matrix equation solutions to sampled-data systems integral**

$$n = 1 \quad \begin{bmatrix} a_0 & a_1 \\ a_1 & a_0 \end{bmatrix} \begin{bmatrix} a_0 I_1 \\ M_1 \end{bmatrix} = \begin{bmatrix} b_0^2 + b_1^2 \\ 2b_0 b_1 \end{bmatrix}$$

$$n = 2 \quad \begin{bmatrix} a_0 & a_1 & a_2 \\ a_1 & a_0 + a_2 & a_1 \\ a_2 & 0 & a_0 \end{bmatrix} \begin{bmatrix} a_0 I_2 \\ M_1 \\ M_2 \end{bmatrix} = \begin{bmatrix} b_0^2 + b_1^2 + b_2^2 \\ 2(b_0 b_1 + b_1 b_2) \\ 2b_0 b_2 \end{bmatrix}$$

$$n = 3 \quad \begin{bmatrix} a_0 & a_1 & a_2 & a_3 \\ a_1 & a_0 + a_2 & a_1 + a_3 & a_2 \\ a_2 & a_3 & a_0 & a_1 \\ a_3 & 0 & 0 & a_0 \end{bmatrix} \begin{bmatrix} a_0 I_3 \\ M_1 \\ M_2 \\ M_3 \end{bmatrix} = \begin{bmatrix} b_0^2 + b_1^2 + b_2^2 + b_3^2 \\ 2(b_0 b_1 + b_1 b_2 + b_2 b_3) \\ 2(b_0 b_2 + b_1 b_3) \\ 2b_0 b_3 \end{bmatrix}$$

$$n = 4 \quad \begin{bmatrix} a_0 & a_1 & a_2 & a_3 & a_4 \\ a_1 & a_0 + a_2 & a_1 + a_3 & a_2 + a_4 & a_3 \\ a_2 & a_3 & a_0 + a_4 & a_1 & a_2 \\ a_3 & a_4 & 0 & a_0 & a_1 \\ a_4 & 0 & 0 & 0 & a_0 \end{bmatrix} \begin{bmatrix} a_0 I_4 \\ M_1 \\ M_2 \\ M_3 \\ M_4 \end{bmatrix} = \begin{bmatrix} b_0^2 + b_1^2 + b_2^2 b_3^2 + b_4^2 \\ 2(b_0 b_1 + b_1 b_2 + b_2 b_3 + b_3 b_4) \\ 2(b_0 b_2 + b_1 b_3 + b_2 b_4) \\ 2(b_0 b_3 + b_1 b_4) \\ 2b_0 b_4 \end{bmatrix}$$

**Table 2. Matrix equation solutions to continuous systems integral**

$n = 1 \quad [a_1]\,[2a_0 I_1] = [b_1^2]$

$$n = 2 \quad \begin{bmatrix} a_2 & 0 \\ a_0 & a_1 \end{bmatrix} \begin{bmatrix} M_0 \\ -2a_0 I_2 \end{bmatrix} = \begin{bmatrix} b_2^2 \\ -b_1^2 \end{bmatrix}$$

$$n = 3 \quad \begin{bmatrix} a_3 & 0 & 0 \\ a_1 & a_2 & a_3 \\ 0 & a_0 & a_1 \end{bmatrix} \begin{bmatrix} M_0 \\ M_1 \\ 2a_0 I_3 \end{bmatrix} = \begin{bmatrix} b_3^2 \\ -b_2^2 + 2b_1 b_3 \\ b_1^2 \end{bmatrix}$$

$$n = 4 \quad \begin{bmatrix} a_4 & 0 & 0 & 0 \\ a_2 & a_3 & a_4 & 0 \\ a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & a_0 & a_1 \end{bmatrix} \begin{bmatrix} M_0 \\ M_1 \\ M_2 \\ -2a_0 I_4 \end{bmatrix} = \begin{bmatrix} b_4^2 \\ -b_3^2 + 2b_2 b_4 \\ b_2^2 - 2b_1 b_3 \\ -b_1^2 \end{bmatrix}$$

$$n = 5 \quad \begin{bmatrix} a_5 & 0 & 0 & 0 & 0 \\ a_3 & a_4 & a_5 & 0 & 0 \\ a_1 & a_2 & a_3 & a_4 & a_5 \\ 0 & a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & 0 & a_0 & a_1 \end{bmatrix} \begin{bmatrix} M_0 \\ M_1 \\ M_2 \\ M_3 \\ 2a_0 I_5 \end{bmatrix} = \begin{bmatrix} b_5^2 \\ -b_4^2 + 2b_3 b_5 \\ b_3^2 - 2b_2 b_4 + 2b_1 b_5 \\ -b_2^2 + 2b_1 b_3 \\ b_1^2 \end{bmatrix}$$

# Appendix A
# Derivation of Eq. (6)

The integral to be evaluated is

$$I_n = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} dS\, G(S)G(-S) \tag{A-1}$$

where

$$G(S) = \frac{B(S)}{A(S)} = \frac{\sum_{i=0}^{n-1} b_{n-i}S^i}{\sum_{i=0}^{n} a_{n-i}S^i} \tag{A-2}$$

Using a partial fraction expansion, assuming nonrepeated roots of $A(s)$

$$G(S)G(-S) = \frac{B(S)\,B(-S)}{A(S)\,A(-S)} = \sum_{k=1}^{n}\frac{R_k}{S+p_k} + \sum_{k=1}^{n}\frac{T_k}{S-p_k} \tag{A-3}$$

where $R_k$ is the residue of pole, $-p_k$, in the left-hand complex $S$-plane, and $T_k$ is the residue of the symmetric pole, $p_k$, in the right-hand complex $S$-plane. From residue theory

$$I_n = \sum_{k=1}^{n} R_k \tag{A-4}$$

The relationship of $T_k$ to $R_k$ is found as follows.

$$R_k = \left.\frac{B(S)\,B(-S)}{A'(S)\,A(-S)}\right|_{S=-p_k} = \frac{B(-p_k)\,B(p_k)}{A'(-p_k)\,A(p_k)} \tag{A-5}$$

where $A'(S) \equiv d/dS\, A(S)$. Note that $d/dS\, A\,(-S) = -A'(-S)$. Therefore

$$T_k = -\left.\frac{B(S)\,B(-S)}{A(S)\,A'(-S)}\right|_{s=p_k} = -\frac{B(p_k)\,B(-p_k)}{A(p_k)\,A'(-p_k)} = -R_k \tag{A-6}$$

From Eqs. (A-3) and (A-6)

$$B(S)\,B(-S) = \sum_{k=1}^{n} R_k\,\frac{A(S)\,A(-S)}{S+p_k} - \sum_{k=1}^{n} R_k\,\frac{A(S)\,A(-S)}{S-p_k} \tag{A-7}$$

Eq. (6) is obtained by expanding both sides of Eq. (A-7) in polynomials of $S$ and equating coefficients. Starting with the right side of Eq. (A-7), define $c_{i,k}$ such that

$$\frac{A(S)}{S+p_k} = \sum_{i=0}^{n-1} c_{i,k} S^i \tag{A-8}$$

From Eqs. (A-2) and (A-8),

$$\sum_{i=0}^{n-1} c_{i,k} S^i = \frac{1}{S+p_k} \sum_{i=0}^{n} a_{n-i} S^i \tag{A-9}$$

which becomes

$$\sum_{i=1}^{n} c_{i-1,k} S^i + p_k \sum_{i=0}^{n-1} c_{i,k} S^i = \sum_{i=0}^{n} a_{n-i} S^i \tag{A-10}$$

For $i = n$

$$c_{n-1,k} = a_0 \tag{A-11}$$

Using Eq. (A-8)

$$\frac{A(-S)}{S-p_k} = -\frac{A(-S)}{-S+p_k} = -\sum_{i=0}^{n-1} c_{i,k} (-S)^i \tag{A-12}$$

From Eqs. (A-2) and (A-8)

$$\frac{A(S)A(-S)}{S+p_k} = \sum_{i=0}^{n-1} c_{i,k} S^i \sum_{l=0}^{n} a_{n-l} (-S)^l$$

$$= \left( \sum_{j=0}^{n-1} \sum_{i=0}^{j} + \sum_{j=n}^{2n-1} \sum_{i=j-n}^{n-1} \right) c_{i,k} a_{n-j+i} (-1)^{j-i} S^j \tag{A-13}$$

Similarly from Eqs. (A-2) and (A-12)

$$\frac{A(S)A(-S)}{S-p_k} = -\sum_{i=0}^{n-1} c_{i,k} (-S)^i \sum_{l=0}^{n} a_{n-l} S^l$$

$$= -\left( \sum_{j=0}^{n-1} \sum_{i=0}^{j} + \sum_{j=n}^{2n-1} \sum_{i=j-n}^{n-1} \right) c_{i,k} a_{n-j+i} (-1)^i S^j \tag{A-14}$$

Substituting Eqs. (A-13) and (A-14) into Eq. (A-7),

$$B(S)\,B(-S) = \sum_{k=1}^{n} R_k \left( \sum_{j=0}^{n-1} \sum_{i=0}^{j} + \sum_{j=n}^{2n-1} \sum_{i=j-n}^{n-1} \right) c_{i,k} a_{n-j+i} (-1)^{j-i} S^j$$

$$+ \sum_{k=1}^{n} R_k \left( \sum_{j=0}^{m-1} \sum_{i=0}^{j} + \sum_{j=n}^{2n-1} \sum_{i=j-n}^{n-1} \right) c_{i,k} a_{n-j+i} (-1)^{i} S^j$$

$$= \left( \sum_{j=0}^{n-1} \sum_{i=0}^{j} + \sum_{j=n}^{2n-1} \sum_{i=j-n}^{n-1} \right) S^j a_{n-j+i} (-1)^i (1 + (-1)^j) \sum_{k=1}^{n} R_k c_{i,k}$$

$$= \left( \sum_{r=0}^{[\frac{n-1}{2}]} \sum_{i=0}^{2r} + \sum_{r=[\frac{n+1}{2}]}^{n-1} \sum_{i=2r-n}^{n-1} \right) S^{2r} a_{n-2r+i} (-1)^{i} \, 2 \sum_{k=1}^{n} R_k c_{i,k}$$

$$= \sum_{r=0}^{n-1} S^{2r} \sum_{i=0}^{n-1} a_{n-2r+i} M_i \qquad \text{(A-15)}$$

where $a_m = 0, m < 0, m > n$

$$M_i = (-1)^{i} \, 2 \sum_{k=1}^{n} R_k c_{i,k}$$

From Eqs. (A-4) and (A-11)

$$M_{n-1} = (-1)^{n-1} \, 2 \sum_{k=1}^{n} R_k c_{n-1,k} = (-1)^{n-1} \, 2 a_0 I_n \qquad \text{(A-16)}$$

Using Eq. (A-2), the left side of Eq. (A-7) becomes

$$B(S)\,B(-S) = \sum_{i=0}^{n-1} b_{n-i} S^i \sum_{l=0}^{n-1} b_{n-l} (-S)^l$$

$$= \left( \sum_{j=0}^{n-1} \sum_{i=0}^{j} + \sum_{j=n}^{2n-2} \sum_{i=j-n+1}^{n-1} \right) b_{n-i} b_{n-j+i} (-1)^{j-i} S^j$$

$$= \sum_{r=0}^{n-1} S^{2r} \sum_{i=0}^{n-1} b_{n-i} b_{n-2r+i} (-1)^i \qquad \text{(A-17)}$$

where $b_m = 0, m < 1, m > n$.

It is convenient to express the inner summation in Eq. (A-17) as

$$\sum_{i=0}^{n-1} b_{n-i} b_{n-2r+i} (-1)^i = b_{n-r}^2 (-1)^r + 2 \sum_{k=1}^{\infty} (-1)^{r+k} b_{n-r-i} b_{n-r+i} \tag{A-18}$$

Therefore Eq. (A-17) becomes

$$B(S)\,B(-S) = \sum_{r=0}^{n-1} S^{2r} \left[ b_{n-r}^2 (-1)^r + 2 \sum_{k=1}^{\infty} (-1)^{r+k} b_{n-r-k} b_{n-r+k} \right] \tag{A-19}$$

where $b_m = 0, m < 1, m > n$

Equating the coefficients of $S^{2r}$ in Eqs. (A-15) and (A-19) gives

$$\sum_{i=0}^{n-1} a_{n-2r+i} M_i = (-1)^r b_{n-r}^2 + 2 \sum_{k=1}^{\infty} (-1)^{r+k} b_{n-r-k} b_{n-r+k}$$

which is Eq. (6). It is conjectured that the results are also valid for repeated roots.

D21 N82-11300



# Surface Resistivity Measurements of Candidate Subreflector Surfaces

E. H. Thom
TDA Engineering Office

T. Y. Otoshi
Radio Frequency and Microwave Subsystems Section

*Subreflectors on DSN antennas have historically been of formed and welded sheet metal construction, which is quite costly. Flame-spray metallized fiberglass-epoxy offers an alternative technique at much lower cost. This article presents results of a study to show that system noise temperature would be increased about 0.3 K at X-band when using the alternate material. Additional testing, however, is required to evaluate weathering and noise generation while diplexing at high power.*

## I. Introduction

In support of a study to find low-cost alternative materials for subreflector surfaces, a work order was issued to the Harris Corporation of Melbourne, Florida, to (1) purchase samples of flame-spray metallized fiberglass-epoxy materials, (2) construct a test cavity to JPL specifications, (3) perform measurements in accordance with JPL procedures, (4) reduce the data taken, and (5) prepare a test report. Except for Section VIII, this article is primarily a copy of the test report by C. W. Choi and G. J. Kirkpatrick of the Harris Corporation.

## II. Cavity and Samples

A $TE_{011}$ cavity resonator technique, developed by Clauss and Potter of JPL (Ref. 1), was the measurement technique used by Harris Corporation for the evaluations of low-cost alternative materials. The drawing supplied by JPL for the cavity resonator is shown in Fig. 1. Unavailability of a large block of copper made it necessary to find another suitable high-conductivity metal to make the cavity. Silver-plated brass was selected instead of aluminum owing to the higher cavity Q-values obtainable with this plated material. The top lid of the cavity is formed by laying a flat test sample on top of the cavity.

Flat sheet test samples were made from the following materials:

(1) Copper.

(2) Aluminum 6061-T6.

(3) Aluminum 2024-T3.

(4) Standard flame-sprayed aluminum on fiberglass substrate.

(5) Standard flame-sprayed copper on item 4 (resulting in a metallized surface consisting of copper over aluminum).

(6) Buffed, flame-sprayed aluminum on fiberglass substrate.

(7) Standard flame-sprayed copper on item 6 (resulting in a metallized surface consisting of copper over buffed aluminum).

Three sets of 4- × 4-in. samples were made for each of the materials in order to determine the repeatability of the Q-value measurements of each kind of sample. A silver-plated brass plate was used as the reference material in the calculation of the surface resistivity of each sample.

## III. Experimental Setup

A test setup block diagram for precision Q-measurement is available from a JPL report (Ref. 1). Since the available test equipment did not exactly match that listed in the JPL report, the revised test setup shown in Fig. 2 was adopted. This arrangement did not allow swept-frequency display, but because of the synthesized signal generator (HP 8672A) and precision attenuators (HP H382A), highly accurate data could be taken on a point-by-point basis.

Some pertinent electrical characteristics for major test apparatus are shown below:

(1) HP 8672A synthesized signal generator frequency resolution: 2 kHz at X-band.

Time base: internal 10 MHz ($<5 \times 10^{-10}$/day aging rate).

(2) HP H382A variable attenuator accuracy: ±0.2% of reading or 0.1 dB, whichever is greater.

(3) HP 8410B/8412A/8411A network analyzer accuracy: 0.08 dB/dB from midscreen. (Note that all the measurements were done at the midscreen employing two variable attenuators HP H382A.) The dual channel scheme of the network analyzer significantly reduces amplitude variation errors from the signal generator.

## IV. Measurement and Results

The loaded $Q$ of the cavity was obtained from

$$Q_\ell = \frac{F_0}{\Delta F} \tag{1}$$

where $F_0$ is the resonant center frequency and $\Delta F$ is the 3-dB frequency spread. The approximate center frequency and a pair of frequencies each for the 2-, 3-, and 4-dB points were recorded. The 3-dB point measurements were repeated and the peak power level checked again. All of these measurements were taken at the midscreen of the network analyzer using two precision variable attenuators to make accurate amplitude measurements. No attempt was made to utilize the phase information of the network analyzer since it is redundant. To insure the repeatability of the measurement, three sets of samples were tested. The center frequency was calculated from the average of the 3-dB points, since the frequency response of the test cavity near resonance is relatively flat. The insertion loss at the center frequency was measured, using one of the variable precision attenuators, after the frequency response of the cavity was measured.

Once these data were collected for each sample, the loaded $Q$ was calculated from Eq. (1) and the unloaded $Q$ was obtained from Eq. (15) of Ref. 1, given below as

$$Q_0 = Q_\ell \left( \frac{L^{1/2}}{L^{1/2} - 1} \right) \tag{2}$$

where $L$ is the transmission power loss ratio. The following expression given in Ref. 1

$$Q_0 = \frac{\left(\frac{\pi}{2}\right) Z_0 \left(\frac{\lambda_c}{\lambda_0}\right)^2 \left(\frac{\lambda_g}{\lambda_0}\right)}{\left[ R_{s_{cyl}} \left(\frac{\ell}{a}\right) + \left(\frac{\lambda_c}{\lambda_g}\right)^2 \left( R_{s_{end1}} + R_{s_{end2}} \right) \right]} \tag{3}$$

was employed to calculate the surface resistivity where

$Z_0 = 120\,\pi$, ohms

$\ell$ = length of cavity

$a$ = radius of cavity

$\lambda_c = a/0.610$

$\lambda_g = 2\ell$

$R_{s_x}$ is the surface resistivity of the location designated by $x$.

First the surface resistivity of the reference sample, a silver-plated brass plate, was calculated from Eq. (3). This value was subsequently used for $R_{s_{cyl}}$ and $R_{s_{end1}}$ in order to calculate the value of $R_{s_{end2}}$ for the other samples of each set. These were tabulated for each set and are shown in Tables 1 through 3 along with the measured data.

Examination of this data showed that the value of the surface resistivity of the silver-plated brass was reasonable. The value shown for copper falls between the two experimental figures given for copper in Ref. 1. However, the aluminum samples showed an unusually low value of resistivity, the 2024-T3 differing by about 20% from the sample JPL data on 2024 alloy of unknown T-value. The values for the flame-sprayed samples could not be compared owing to the lack of published data.

A possible explanation for the aluminum measurement could be a disparity between the surface characteristics of the reference silver-plated brass plate and the interior of the cavity.

To eliminate this possibility, a new brass plate was fabricated, and both the plate and cavity were polished to a uniform surface roughness, then simultaneously plated with about 10 skin depths of silver. The cavity coupling was increased to about 15 dB for the silver-plated brass plate, a 10-dB pad was inserted on each side of the test cavity for improved isolation, new data were obtained and the results were tabulated. Comparing the data to the original set, a rather uniform reduction in surface resistivity was observed, the average downward shift being about 0.00243 ohms. To insure the integrity of the sample surfaces, several of them were cleaned. Four of the aluminum plates were cleaned in trichloroethylene, nitric acid, and water while another two were cleaned in alcohol. Two copper plates were cleaned in trichloroethylene, hydrochloric acid, and water while another was cleaned in alcohol, as shown in Tables 4-6. The cleaning had a negligible effect on the results.

## V. Measurement Accuracy

As mentioned in the experimental setup section there are errors associated with the equipment. One source of error is the measurement of the 3-dB frequency spread. The precision variable attenuator has a possible error of 0.1 dB at the 3-dB point. In addition, the error involved in resetting to the midscreen of the network analyzer is on the order of 0.01 dB. With ±0.11 dB maximum error for the 3-dB measurement and the approximate slope of the frequency response of the test cavity, the error in the value of $Q_\ell$ can be bounded. The maximum error in the dB reading of the transmission loss ratio ($L$, dB) is 0.1 dB and an additional 0.01 dB due to resetability. Except for the error of 0.01 dB in resetting, the errors mentioned above are not random errors, but errors inherent in the equipment.

Experimentally, examining the slope of the frequency response, the error involved in the 3-dB frequency spread due to the possible 0.11-dB deviation was found to be less than 0.026 $\Delta F$. Thus the error involved in the loaded $Q$ value (Eq. 1) was less than 2.7%. The error involved with the measurement of the transmission loss ratio resulted in less than 0.28% error in the calculation of the unloaded $Q$ (Eq. 2) in any one of the samples. Additional unaccounted for errors such as variations in temperature and humidity may be estimated by examining Tables 1, 2, and 3 side by side. Maximum deviation of the values of unloaded $Q$ for any particular material is less than 1.36% for the first three sets of data shown in Tables 1, 2, and 3. The gross estimation of cumulative error due to all effects appears to be less than 10%. The 20% variance between the measured resistivity of aluminum and the published values remains to be explained. The three sets of seven samples were sent to JPL for possible verification.

## VI. Summary of Resistivity Measurements

Table 7 summarizes the surface resistivity measurement results of all six runs for each material tested. More detailed data can be found in Tables 1 through 6.

## VII. Microwave Performance Effects

Subsequent to receiving the resistivity results from Harris Corporation, JPL calculated the system temperature contribution which would be caused at X-band by the surface resistivity of each of the materials using the approximate formula for normal incidence angle (Ref. 2)

$$N_t \approx \frac{4R_s}{Z_0} T_p \tag{4}$$

where

$Z_0 = 120\,\pi$ ohms

$T_p = 290$ K

Table 8 shows these results.

Typically, DSN subreflectors have been made from 6061-T6 aluminum material which gives a system noise temperature contribution of 0.117 K at X-band. If buffed flame-sprayed aluminum on fiberglass substrate were used for the subreflector material, its contribution would be 0.391 K resulting in a net system temperature increase of approximately 0.27 K at X-band.

## VIII. Conclusion

Use of the low-cost materials would cause a small increase in system temperature while providing a large reduction in the procurement cost. However, before selecting these materials for use in the DSN, additional studies are required to ensure that the surface resistivity does not degrade with extended weathering. Another factor which requires evaluation for antennas which also transmit is possible noise that is generated while diplexing. In the final fabrication, quality assurance steps need to be taken to ensure that the plating thickness is at least 5 skin depths at the lowest S-band (or L-band) frequency at which the subreflector will be used.

## Acknowledgment

The authors thank D. Bathker of the Radio Frequency and Microwave Subsystems Section for encouragement and support as well as technical contributions to this article.

## References

1. Clauss, R., and Potter, P. D., "Improved R.F. Calibration Techniques – A Practical Technique for Accurate Determination of Microwave Surface Resistivity," Technical Report 32-1526, Vol. XII, pp. 54-67, Jet Propulsion Laboratory, Pasadena, Calif., Dec. 15, 1972.

2. Otoshi, T., *IEEE Trans. Instrumentation and Measurement*, Nov. 1972, p. 451 (Use Eq. (3) for zero diameter hole case.)

**Table 1. Run 1, first set of samples**

| Sample | $F_0$, MHz | ΔF, MHz | $Q_\ell$ | $L$, dB | $Q_0$ | $R_s$, Ω |
|---|---|---|---|---|---|---|
| Silver-plated brass | 8426.842 | 0.392 | 21497 | 21.8 | 23399 | 0.02777 |
| Copper | 8417.791 | 0.394 | 21365 | 21.7 | 23279 | 0.02795 |
| Aluminum 6061-T6 | 8428.158 | 0.440 | 19155 | 22.6 | 20689 | 0.03987 |
| Aluminum 2024-T3 | 8428.632 | 0.424 | 19879 | 22.3 | 21531 | 0.03582 |
| Standard flame-sprayed aluminum | 8426.427 | 0.854 | 9867 | 28.25 | 10264 | 0.1455 |
| Standard flame-sprayed aluminum and copper | 8417.341 | 0.682 | 12342 | 26.5 | 12955 | 0.1014 |
| Buffed flame-sprayed aluminum | 8417.545 | 0.790 | 10655 | 27.5 | 11124 | 0.1286 |
| Buffed flame-sprayed aluminum and copper | 8419.627 | 0.694 | 12132 | 26.6 | 12727 | 0.1045 |

**Table 2. Run 2, second set of samples**

| Sample | $F_0$, MHz | $\Delta F$, MHz | $Q_\ell$ | $L$, dB | $Q_0$ | $R_s$, Ω |
|---|---|---|---|---|---|---|
| Silver-plated brass | 8428.279 | 0.394 | 21392 | 21.6 | 23332 | 0.02787 |
| Copper | 8420.616 | 0.392 | 21481 | 22.6 | 23201 | 0.02814 |
| Aluminum 6061-T6 | 8428.683 | 0.438 | 19244 | 22.5 | 20804 | 0.03910 |
| Aluminum 2024-T3 | 8428.738 | 0.424 | 19879 | 22.2 | 21552 | 0.03550 |
| Standard flame-sprayed aluminum | 8427.374 | 0.856 | 9845 | 28.2 | 10244 | 0.1457 |
| Standard flame-sprayed aluminum and copper | 8416.108 | 0.680 | 12377 | 26.9 | 12962 | 0.1010 |
| Buffed flame-sprayed aluminum | 8427.564 | 0.780 | 10805 | 27.6 | 11275 | 0.1265 |
| Buffed flame-sprayed aluminum and copper | 8417.311 | 0.694 | 12129 | 26.5 | 12731 | 0.1041 |

**Table 3. Third set of samples**

| Sample | $F_0$, MHz | $\Delta F$, MHz | $Q_\ell$ | $L$, dB | $Q_0$ | $R_s$, Ω |
|---|---|---|---|---|---|---|
| Silver-plated brass | 8428.785 | 0.394 | 21393 | 21.7 | 23309 | 0.02790 |
| Copper | 8435.023 | 0.394 | 21409 | 22.0 | 23256 | 0.02832 |
| Aluminum 6061-T6 | 8428.758 | 0.440 | 19156 | 22.6 | 20690 | 0.03960 |
| Aluminum 2024-T3 | 8427.930 | 0.424 | 19877 | 22.3 | 21529 | 0.03551 |
| Standard flame-sprayed aluminum | 8423.660 | 0.856 | 9841 | 28.5 | 10225 | 0.1458 |
| Standard flame-sprayed aluminum and copper | 8414.552 | 0.688 | 12230 | 26.5 | 12838 | 0.1024 |
| Buffed flame-sprayed aluminum | 8427.631 | 0.786 | 10722 | 27.7 | 11183 | 0.1280 |
| Buffed flame-sprayed aluminum and copper | 8415.836 | 0.698 | 12057 | 26.7 | 12642 | 0.1051 |

**Table 4. Run 4, first set of samples**

| Sample | $F_0$, MHz | $\Delta F$, MHz | $Q_\ell$ | $L$, dB | $Q_0$ | $R_s$, Ω |
|---|---|---|---|---|---|---|
| Silver-plated brass | 8423.297 | 0.382 | 22051 | 15.4 | 26561 | 0.02449 |
| Copper[a] | 8417.284 | 0.380 | 22151 | 16.0 | 26337 | 0.02500 |
| Aluminum 6061-T6[a] | 8425.422 | 0.424 | 19871 | 16.5 | 23358 | 0.03568 |
| Aluminum 2024-T3[a] | 8425.776 | 0.416 | 20254 | 16.1 | 24017 | 0.03316 |
| Standard flame-sprayed aluminum | 8423.785 | 0.854 | 9864 | 22.2 | 10694 | 0.1449 |
| Standard flame-sprayed aluminum and copper | 8415.310 | 0.676 | 12449 | 20.5 | 13746 | 0.09965 |
| Buffed flame-sprayed aluminum | 8423.511 | 0.786 | 10717 | 21.6 | 11683 | 0.1278 |
| Buffed flame-sprayed aluminum and copper | 8415.649 | 0.686 | 12268 | 20.6 | 13530 | 0.1022 |

[a]Chemically cleaned.

**Table 5. Run 5, second set of samples**

| Sample | $F_0$, MHz | $\Delta F$, MHz | $Q_\ell$ | $L$, dB | $Q_0$ | $R_s$, Ω |
|---|---|---|---|---|---|---|
| Silver-plated brass (polished) | 8422.908 | 0.384 | 21935 | 15.4 | 26453 | 0.02481 |
| Silver-plated brass (unpolished) | 8424.378 | 0.384 | 21938 | 15.4 | 26414 | 0.02497 |
| Copper[a] | 8415.636 | 0.388 | 21690 | 15.0 | 26381 | 0.02482 |
| Aluminum 6061-T6[a] | 8425.615 | 0.430 | 19594 | 16.6 | 23019 | 0.03704 |
| Aluminum 2024-T3[a] | 8425.428 | 0.416 | 20253 | 16.1 | 24016 | 0.03315 |
| Standard flame-sprayed aluminum | 8421.993 | 0.846 | 9955 | 22.0 | 10809 | 0.1426 |
| Standard flame-sprayed aluminum and copper | 8412.003 | 0.682 | 12334 | 20.2 | 13670 | 0.1003 |
| Buffed flame-sprayed aluminum | 8424.718 | 0.780 | 10801 | 21.6 | 11787 | 0.1262 |
| Buffed flame-sprayed aluminum and copper | 8411.249 | 0.690 | 12190 | 20.3 | 13498 | 0.1023 |

[a]Chemically cleaned.

**Table 6. Run 6, third set of samples**

| Sample | $F_0$, MHz | $\Delta F$, MHz | $Q_\ell$ | $L$, dB | $Q_0$ | $R_s$, Ω |
|---|---|---|---|---|---|---|
| Silver-plated brass | 8420.424 | 0.380 | 22159 | 15.8 | 26434 | 0.02479 |
| Copper[a] | 8432.156 | 0.380 | 22190 | 16.4 | 26174 | 0.02595 |
| Aluminum 6061-T6[a] | 8426.163 | 0.426 | 19780 | 17.0 | 23055 | 0.03692 |
| Aluminum 2024-T3[a] | 8425.732 | 0.412 | 20451 | 16.6 | 24001 | 0.03322 |
| Standard flame-sprayed aluminum | 8423.786 | 0.848 | 9934 | 22.4 | 10749 | 0.1438 |
| Standard flame-sprayed aluminum and copper | 8411.620 | 0.664 | 12668 | 21.0 | 13915 | 0.09755 |
| Buffed flame-sprayed aluminum | 8424.558 | 0.780 | 10801 | 21.4 | 11799 | 0.1260 |
| Buffed flame-sprayed aluminum and copper | 8412.476 | 0.684 | 12299 | 20.4 | 13606 | 0.1011 |

[a]Cleaned with alcohol.

**Table 7. Summary of resistivity data at 8.43 GHz (see Tables 1–6 for actual frequencies)**

| Sample | Lowest $R_s$, Ω | Highest $R_s$, Ω | Average $R_s$, Ω |
|---|---|---|---|
| Silver-plated brass | 0.02449 | 0.02790 | 0.02627 |
| Copper | 0.02482 | 0.02832 | 0.02670 |
| Aluminum 6061-T6 | 0.03568 | 0.03987 | 0.03804 |
| Aluminum 2024-T3 | 0.03315 | 0.03582 | 0.03439 |
| Standard flame-sprayed aluminum | 0.1426 | 0.1458 | 0.1447 |
| Standard flame-sprayed aluminum and copper | 0.09755 | 0.1024 | 0.1004 |
| Buffed flame-sprayed aluminum | 0.1260 | 0.1286 | 0.1272 |
| Buffed flame-sprayed aluminum and copper | 0.1011 | 0.1051 | 0.1032 |

**Table 8. Average resistivity and noise temperature contributions at 8.43 GHz**

| Material | Average resistivity, Ω | Noise contribution, K |
|---|---|---|
| Silver-plated brass | 0.02627 | 0.081 |
| Copper sheet | 0.02670 | 0.082 |
| 6061 T6 aluminum | 0.03804 | 0.117 |
| 2024 T3 aluminum | 0.03439 | 0.106 |
| Standard flame-sprayed aluminum | 0.1447 | 0.445 |
| Standard flame-sprayed aluminum and copper | 0.1004 | 0.309 |
| Flame-sprayed aluminum (buffed) | 0.1272 | 0.391 |
| Flame-sprayed aluminum and copper (buffed) | 0.1032 | 0.318 |

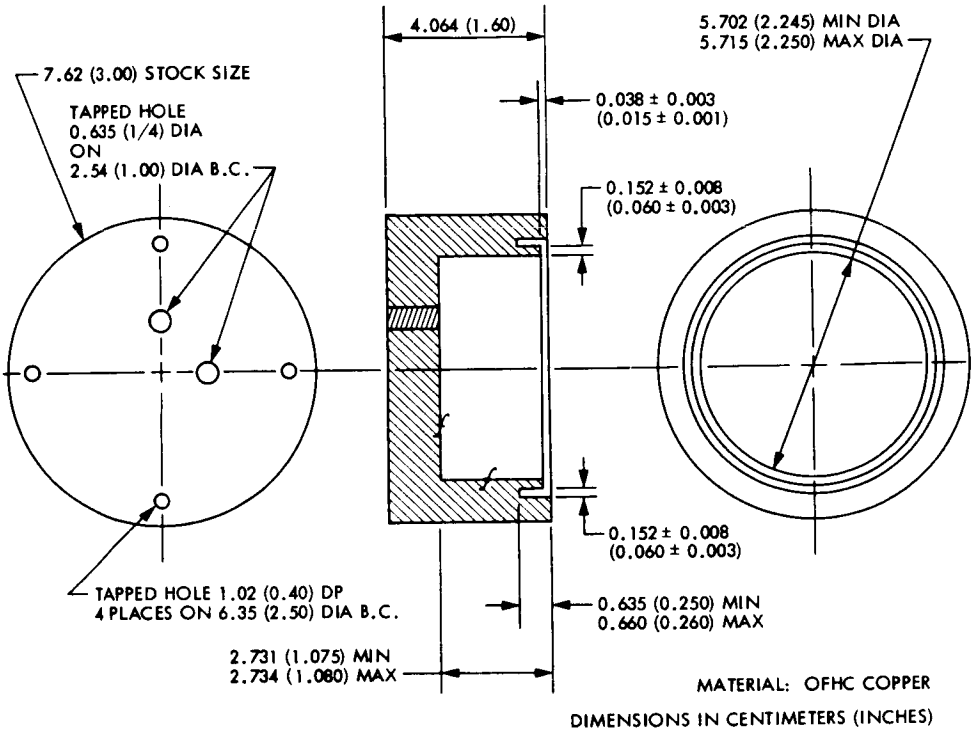



**Fig. 1. Cavity dimensions. Note: Required inside surface finish is 40 microcm (16 microinch) or better**

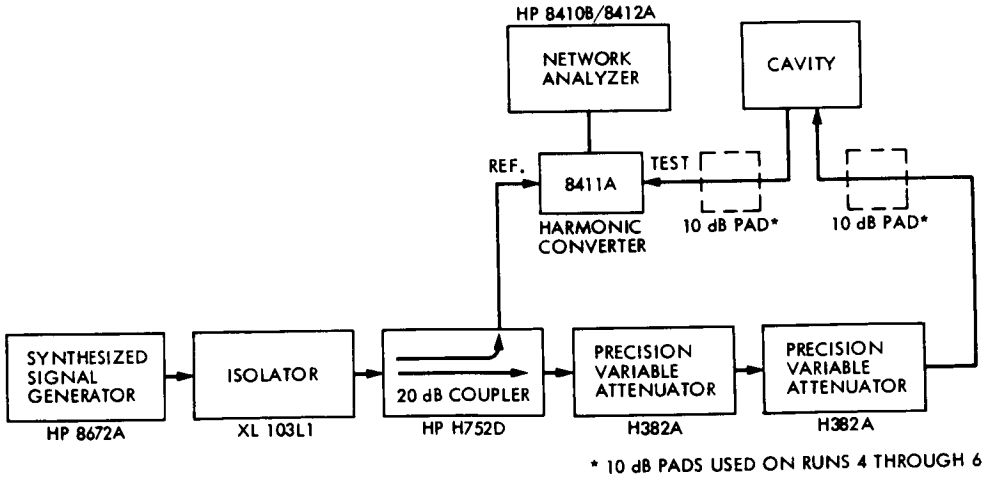



**Fig. 2. Test setup**

N82-11301 D22



# Rehabilitation of 64-Meter-Antenna Radial Bearing

H. McGinness and G. Gale
DSN Engineering Section

*The performances of the radial bearings on three 64-meter antennas are presented. Explanations for the distortion of the bearing at DSS 14 are made and the repairs are described. Recommendations for future tests and action are given.*

## I. Introduction

The radial bearing of the 64-m antenna is composed of two basic assemblies, namely, the track assembly which is attached to the antenna foundation, and the three wheel assemblies which are equally spaced on the alidade structure. The wheel assemblies are radially preloaded against the vertical cylindrical surface of the track, thus forming a vertical axis bearing capable of resisting horizontal forces applied to the antenna. The track assembly base or runner is a mild steel ring having a mean diameter of 8.96 m and a cross section 114 mm wide by 508 mm high. This ring was fabricated as four equal circular arc pieces, having mitered ends, which were permanently joined by five shrunk-in shear pins and four tension bolts. Ten alloy steel wear strips with mitered ends and counterbored holes are bolted to the mild steel ring. The wear strips are 35 mm thick and 310 mm high and their outer surfaces form the 8.144-m-diameter bearing race.

Each wheel assembly is composed of two wheels mounted in a pivoted frame attached to the antenna alidade base structure. Each wheel is 914 mm in diameter and 292 mm wide.

After being assembled on the ground, the track ring was hoisted up and placed around the collar at the top of the antenna pedestal. Alignment bolts, as shown in Fig. 1, were adjusted so as to obtain level, circularity, and concentricity. At DSS 14 the grout used between the ring and pedestal was EMBECO. After the installation of the track the rest of the antenna was built over it. The details of the parts shown in Fig. 1, as well as the wheel assembly parts, are shown on JPL drawing 9437995.

## II. Brief History of Radial Bearing Runner Symptoms

The 64-m antenna at DSS 14 was completed in 1966. The preload between each truck (an assembly of two wheels) and the track was set at 1,468,500 N. The following observations and changes were made in the years noted:

*1968*

The wear strips displayed a constantly changing wheel contact pattern.

*1975 July*

The wear strip contact pattern suggested areas of localized stiffness behind the runner. The pattern suggested that the alignment bolts were firm and were supporting the radial loads. Further investigation showed deterioration of the contact between the runner and the grout. The truck preload was reduced to 979,000 N.

*1976 July*

The gap between runner and grout had increased so that 0.50-mm feeler gage could be inserted to a depth of 150 mm.

*1976 November*

The gap between runner and grout had increased to 1.00 mm, and bolts holding the wear strips to the runner were starting to break.

*1976 December*

The wear strip bearing surface was not vertical but was sloped inward and upward by 0.24 degrees.

## III. Ascertainment of the Cause of Runner Displacement

In 1977 a section of grout was cored out so that the inside of the runner could be seen. A large amount of scale was present, especially at the bottom where its thickness was 2 mm or more (see Fig. 2). It was then definite that the corrosion was caused by moisture coming into contact with the runner/EMBECO grout interface. The conical shape of the runner was judged to be due to the graduation of corrosion, varying from a very small amount at the top to an appreciable amount at the bottom. It was impossible to know what change had occurred in the runner radius, since original dimensions were not known more precisely than were given by dimensional tolerances. It could only be assumed that the conical shape of the runner was caused primarily by an outward radial displacement of its lower edge equal to the product of the measured slope and runner height. An analysis of the runner was made to determine its hoop force corresponding to the estimated amount of radial displacement. Then an estimate of the joint pin shear stresses, corresponding to this hoop force, was made. When the shear pin stress was compared to what was believed to be the breaking stress, it was judged that the margin of safety was too low. It was believed also that the corrosion on the back side of the runner would cease if oxygen were excluded from the region. Therefore the decision was made to strengthen the runner by welding the joints and to exclude oxygen by filling all voids with an epoxy cement and sealing the upper and lower edges of the runner with an appropriate substance (see Fig. 3).

## IV. Action Taken to Stop Corrosion and Restore Runner Strength

In 1977 the runner joints at DSS 14 were welded in sequence according to the following procedure:

(1) All wear strips were removed. This increased the load on the joint shear pins by an estimated amount of from 2 to 5%.

(2) A weld was placed along the mitered edge of the joint so that a penetration of 26 mm was obtained. This produced an additional shear area of approximately 0.030 $m^2$. The combined shear area of the five shear pins was 0.0057 $m^2$.

(3) Approximately 0.152 m was torch-cut off the end of the runner segment. In so doing one shear pin was sacrificed. The cut end was then welded to its mating runner segment across the 0.114-m thickness of the runner. The additional weld area was approximately 0.006 $m^2$.

(4) Step 3 was repeated at the joint's other end, thereby sacrificing another shear pin.

(5) Weld material was added alternately to the top and bottom end welds until the spaces occupied by the cut off parts were filled. The total weld area per joint was approximately 0.062 $m^2$.

(6) The outer surface of the runner at the joint was ground smooth and the wear strips replaced. Holes in the wear strips had to be enlarged in order to replace the holding bolts.

It had been decided to fill the voids behind the runner with a special epoxy resin which had been developed for filling cracks and voids in earthquake-damaged concrete structures. Since there existed a finite gap between the grout and upper edge of runner, the epoxy could enter at the gap and by elevation head pressure be made to fill all the voids, provided that a temporary seal was established at the lower edge of the runner. The epoxy filling was accomplished in 1978. The exothermic action of this particular epoxy reduces the initial viscosity enormously so that it is capable of entering hairline cracks. For this reason it may be assumed that the voids were filled. The subsequent application of a sealing paint to the outside insures that oxygen is excluded from the runner grout interface.

## V. Measurement of Runner Residual Stresses

Although it was believed that welding the runner joints and filling the grout voids would prevent further distortion, it was deemed prudent to prove that such would be true. This was done, first, by measuring the total stress at several critical places in the vicinity of welded joints, and second, by installing a reference ring from which any additional distorsion could be measured. Both of these procedures were carried out in the summer of 1979.

The total stresses were measured by the proprietary process developed by R. G. Sturm (Ref. 1). The method consists of making a very shallow indentation in a highly polished spot of the surface where the measurement is to be made. The indentation (0.25 to 1.0 mm in diameter and less than 0.10 mm deep) is made with a conical indentor under a precisely controlled load. If stress exists at the polished surface, the edge of the indentation will be elliptical rather than circular. The ratio of the major-to-minor diameter is a function of the total stress at the surface. By indenting a calibration bar made of the same material and subjected to a known bending stress, a good estimate of the stress in the subject material can be calculated.

Polished spots were prepared at 21 places, most of which were near the welded joints. Two of the spots were inadvertently placed on the weld material and spurious results obtained. The location of the spots and the calculated stresses are listed in Ref. 2. Of the 19 measurements considered valid, the highest and lowest tensile stresses were 8292 ($10^4$) and 429 ($10^4$) N/m$^2$. The highest and lowest compressive stresses were 7055 ($10^4$) and 1951 ($10^4$) N/m$^2$. Since the minimum tensile yield stress of the material is 24830 ($10^4$) N/m$^2$, the measured residual stresses are considered to be sufficiently low. If a hoop stress is calculated for the lower edge of the runner, based upon a radial displacement of 1.9 mm, which corresponds to the average of measured slopes, the value of 8480 ($10^4$) N/m$^2$ is obtained. This calculation for the hoop stress, $\sigma_H$, is from the expression:

$$\sigma_H = \frac{\Delta R}{R} E \qquad (1)$$

where $\Delta R$ is the change in radius, $R$ is the mean radius of the runner, and $E$ is the elastic modulus.

Since the upper edge of the runner has moved outward by a small amount compared to that of the lower edge, the estimate for the average hoop stress is half of that estimated for the lower edge, namely 4240 ($10^4$) N/m$^2$. The average shear stress, $\tau_{AV}$, on the joint plane, which is at 25° to the edge of the runner, would be:

$$\tau_{AV} = \sigma_{AV} (\sin 25°) \cos 25° = 0.383\, \sigma_{AV} \qquad (2)$$

if the weld covered the entire joint area. Since the weld area is only 47% of the entire area, the average weld shear stress is estimated to be:

$$\tau_{AV} = \frac{0.383}{0.47} \sigma_{AV} = 0.814\, \sigma_{AV} \qquad (3)$$

Taking the average hoop stress as 4240 ($10^4$) N/m$^2$, the average shear stress at the weld is 3455 ($10^4$) N/m$^2$, which may be compared to the minimum shear yield of the material of 12415 ($10^4$) N/m$^2$.

The foregoing calculations indicate that there are no excessive stresses in the runner as it now exists after having its joints welded. The measured surfaces stresses substantiate this conclusion.

## VI. Reference Ring

The cross section of the reference ring and its relationship to the radial bearing runner is shown in Fig. 4. Also shown in this figure is the fixture which measures changes in the horizontal distance between the reference ring and runner. The details of these parts can be obtained from JPL drawing 9470753.

The cross section of the reference ring is approximately 63 mm wide and 39 mm high. Its inside radius is 4.13 m, which is approximately 0.42 m less than the radius of the runner wear strip. Any temperature expansion errors will be proportional to 0.42 m and not to the absolute radius, which is more than 10 times as much. Since the temperature difference between the fixture and runner structure is not likely to exceed 3°C and since their expansion coefficients are virtually the same, a measurement error of only 0.014 mm is calculated by forming the product of the distance, temperature difference, and expansion coefficient, namely the product, 420(3)(11) $10^{-6}$ = 0.014 mm. This is a trivial amount compared to the radial displacements which have already occurred, such as 1.9 mm at the lower edge of the runner. However, the sensitivity and accuracy of the measurement system is capable of detecting displacements smaller than 0.014 mm.

The reference ring was made in eight equal circular arc pieces. It was machined on a very accurate vertical boring mill. The final cuts were made after the rough machined parts had been stress relieved. Upon being unclamped from the machine, very little distortion from the circular shape occurred. Locating pads for the ring segments are shown in Fig. 5. The pads were located from a circle of bench marks near the radial bearing runner. The pads were leveled to a common plane by using push-pull screws and surveying techniques. Then the pads were grouted against the concrete foundation with epoxy resin. The pads are within 0.25 mm of a common level plane. The radial buttons, against which the inside edge of the reference ring bears, were located with a precision measuring bar pivoted at the center of the instrument tower. The bar passed through holes in the tower wall which were on the same radial lines as the mounting pads. After being located and checked, the buttons were doweled to the pads. The reference ring segments were placed on the pads and held securely with dogs and screws. Thus the inner edge of the reference ring was circular

as placed. When it was placed, there was a preload on the radial bearing trucks of approximately 890,000 N. A deflection analysis of the concrete structure composed of the ring portion and the disk portion shows a radial inward deflection of 0.061 mm at the trucks and an outward deflection of 0.058 midway between the trucks. This should be considered in calculating the circularity of the radial bearing runner.

The position of the radial bearing wear strip with respect to the reference ring may be seen in Fig. 4. The fixture which measures the horizontal distance from the inside of the reference ring to the outside of the wear strip appears to be two bars forming a right angle. The end of the horizontal bar contains oval-headed setscrews which bear against the top and inside surfaces of the reference ring. The vertical bar contains two linear displacement transducers spaced 254 mm apart. The setscrews shown near the end of the horizontal bar are actually two pairs of screws spaced approximately 0.50 m apart in the direction perpendicular to the plane of the figure. Halfway between the transducers on the vertical bar is a single setscrew. This arrangement allows the fixture to hang in a stable, statically determinate position. The transducer zero values are obtained through the use of a reference gage fixture, which is a welded steel structure duplicating a short section of the radial bearing wear strip and reference ring. The procedure essentially is as follows:

(1) The reference gage fixture is leveled by adjusting set screws on its three supporting feet. A precision machinist level is used to detect the level position (see Fig. 6).

(2) The fixture containing the transducers (hereinafter called the tool) is hung onto the reference gage fixture (see Fig. 7).

(3) The precision level is placed on the horizontal leg of the tool and the adjusting screw between the transducers is turned until the level reads zero. The transducers are moved axially until they both read zero.

(4) The tool is hung on the reference ring and slid circumferentially a small amount until it touches one of many locating pins in the reference ring. These pins establish the exact azimuth angles at which the measurements will always be made.

(5) The precision level is placed on the horizontal leg of the tool and the adjusting screw turned until the level indicates zero. The transducers are read and recorded.

The particular machinist level used has a sensitivity of 0.000042 radians per division. With care, estimates of half divisions can be made. A consideration of the configuration of Fig. 5 shows that the relationship between the transducer displacement $\delta$ and the angular level error $\epsilon_\theta$ is approximately:

$$\delta = 430\,\epsilon_\theta \tag{4}$$

where $\delta$ is in millimeters.

Taking $\epsilon_\theta$ as $\frac{0.000042}{2} = 0.000021$ gives

$$\delta = 430\,(0.000021) = 0.0090 \text{ mm} \tag{5}$$

Experience indicates that if a reading is carefully made and the tool removed, replaced, releveled, and read again, the transducer readings can be repeated to within 0.009 mm. The slightest amount of surface damage, foreign particle inclusion, or change in level temperature can increase the transducer error considerably. Using the tool with the amount of care that is practical, the repeatability error of the transducers is approximately 0.025 mm.

The first measurements of the runner with respect to the reference ring were made in October 1979. In Table 1 some of these are compared to the latest ones made in May 1981. Complete records are available from Section 355.

Detailed instructions for the use of the above described instrumentation can be obtained from JPL Section 355 personnel.

## VII. Radial Bearing Runner Condition at DSS 43 and DSS 63

A reference ring was installed at DSS 63 in January 1981. The general condition of the runner was good, in that there were no gaps between it and the grout. Initial measurements were taken and periodic measurements will be made.

A reference ring was installed at DSS 43 in March 1981. Near the preloaded trucks the gap between runner and grout was near zero, but gradually increased to 1.30, 0.65, and 0.63 mm at the three midpoints between trucks. Circumferential walking of the runner had been observed. Also, cracks and spalled off pieces of concrete had occurred below the runner in the vicinity of a joint.

The walking action can be explained by the fact that the periphery of the inside of the runner is greater than the periphery of the grout. The preloaded wheels force the runner against the grout, producing a rolling action which causes the runner to advance in one turn by the amount of $2\pi h$, where $h$ is the average gap. Taking $h$ as 0.43 mm, the advance in one turn would be approximately 2.7 mm. Because of various asymmetries, reverse motion would not exactly cancel the circumferential displacement, thus allowing a possible net displacement greater than 2.7 mm.

The reason why a gap between runner and grout has developed at DSS 43 and not at DSS 63 is not now known. The age of the two antennas is approximately the same. It is considered that a shrinkage of 0.43 mm over the grout thickness of 76 mm is very large for the dry pack type of portland cement grout used. A stress analysis indicates that the runner alignment bolts could support the runner against the wheel preload. If the grout had shrunk, there should be a gap at the wheel position where the runner would be supported by the alignment bolts. Rather it is more likely that the entire concrete pedestal structure has shrunk radially by the amount of 0.43 mm over the radial distance of 4340 mm. The reason why the DSS 43 pedestal might have shrunk and the DSS 63 pedestal remained dimensionally stable may involve the chemical properties of the different aggregates used.

Table 2 shows a comparison of measurements made at DSS 43.

## VIII. Design of Replacement Radial Bearing

It was decided to procure one or more new radial bearings to replace existing ones in case repairs on the present one are not successful. If repairs appear to be successful, the new one would serve as a spare and its existence would substantially shorten the bearing replacement time if such became necessary.

A replacement runner cannot be installed as were the original ones, because most of the antenna was constructed after the runner was installed as a complete ring. Therefore, the design of the replacement runner must be different. If the runner is made in four segments, as was done with the original runners, and its height reduced from 508 to 445 mm, each segment can be slid under the alidade base structure and then lifted up to its proper position on the pedestal. The four segments would be joined by close fitting bolts at mitered joints. Then the assembled ring would be aligned to the reference ring and grouted into place. The joint design is depicted in Fig. 8 and the details are shown on JPL drawing 9474446. The joint is mitered and uses eight special bolts, having precision-ground sections of 38-mm diameter which serve as shear pins. This is to be compared to the original design which employed five shrunk-in pins of the same diameter. The installation of a new runner would consist of the following steps:

(1) Remove the old wear strips.

(2) Torch-cut the old runner into sufficiently small lengths for removal.

(3) Chip off old grout and renovate alignment bolts. If old bolts cannot be turned in the anchor nuts, they must be torch-cut off and oversized holes tapped in the anchor nuts.

(4) Bring in the four segments of the new runner and assemble into one ring by installing the 32 special bolts, nuts and lock plates.

(5) Align the new runner assembly to the reference ring.

(6) Grout the runner in place.

(7) Install the new wear strips.

## IX. Recommended Repairs

The repair of the DSS 14 runner has been described above. The bearing surface is now conical instead of cylindrical and the wheel preload has not been restored to its original value because of the reduced contact area between the wear strip and wheel. It is proposed that the wheel assemblies be tilted to conform to the average slope of the conical surface of the wear strip. This can be done easily by shimming the wheel bearing housings and taking advantage of the self-aligning characteristics of the spherical roller bearings now in use. If the contact areas in the worst places are sufficiently improved to allow restoration of the original preload, and the tendency of the wheel to run off the track is not excessively increased, then this would seem to represent a satisfactory solution. If these conditions are not met, the new bearing should be installed.

At DSS 43 the gap could be filled with the same type of epoxy resin used at DSS 14. With the wheel preload removed, the circularity of the wear strip surface would be measured and calculated with respect to the reference ring. If the calculated circularity required improvement, it may be possible to achieve it by driving wedges between the runner and grout. Then the epoxy could be applied as it was at DSS 14.

The adherence of the epoxy to runner will probably reduce the tendency for the runner to walk circumferentially. If walking persisted, keys could be inserted by coring out the grout at four equally spaced places and inserting a vertical key containing several tapped holes of 25-mm diameter. Matching holes would be bored and counterbored in the runner and the keys attached to the inside of the runner with bolts. Grout would then be put around the keys. The keys and grout should be designed to withstand a shear force equal to the product of the wheel preload and the coefficient of friction between the runner and grout.

So far the DSS 63 azimuth bearing is without any of the above described defects. However, periodic measurements will be made with respect to its reference ring.

# References


1. Sturm, R. G., "Total Stress Gage, a Method and a Device for Determining Total Stress by Allowing the Material to Indicate Nondestructively the Stress Within It," Technical Note TN-001, Sturm-Stress, Inc., November 1970.

2. Liaw, B. D., Lock, R., Jr., and Sturm, A., "Total Stress Measurement in the DSS 14 64-m Antenna Bearing Ring," prepared under JPL contract (Purchase Order U-698621), July 1979.

**Table 1. Radial bearing measurements at DSS 14**

| Date | Azimuth angle station, deg | Transducer measurements, distance from vertical reference, mm | | Slope from vertical, deg | Mean slope from vertical, average of slopes at 15 azimuth stations, deg |
|---|---|---|---|---|---|
| | | Top | Bottom | | |
| 10-01-79 | 350 | −0.06096 | 1.4734 | 0.3461 max. | 0.3014 |
| | 200 | −0.18491 | 0.8948 | 0.2436 min. | |
| 11-09-79 | 350 | −0.07620 | 1.4486 | 0.3439 max. | 0.3021 |
| | 200 | −0.17348 | 0.8936 | 0.2407 min. | |
| 2-2-80 | 350 | −0.09169 | 1.4491 | 0.3476 max. | 0.3023 |
| | 200 | −0.13868 | 0.9548 | 0.2466 min. | |
| 4-7-80 | 350 | −0.08890 | 1.4376 | 0.3443 max. | 0.3023 |
| | 200 | −0.14783 | 0.9256 | 0.2421 min. | |
| 7-1-80 | 350 | −0.04775 | 1.5006 | 0.3493 max. | 0.3030 |
| | 200 | −0.06071 | 0.9749 | 0.2336 min. | |
| 10-25-80 | 350 | −0.02108 | 1.5400 | 0.3521 max. | 0.3076 |
| | 200 | 0.01270 | 1.0587 | 0.2359 min. | |
| 5-10-81 | 350 | −0.05131 | 1.5265 | 0.3559 max. | 0.3080 |
| | 200 | 0.01270 | 1.0571 | 0.2350 min. | |

**Table 2. Radial bearing measurements at DSS 43**

| Date | Azimuth angle station, deg | Transducer measurements, distance from vertical reference, mm | | Slope from vertical, deg | Mean slope from vertical, average of slopes at 15 azimuth stations, deg |
|---|---|---|---|---|---|
| 4-7-81 | 75 | −0.18313 | 0.24232 | 0.09597 max. | 0.0000618 |
| | 330 | −1.31293 | −1.44272 | −0.02928 min. | |
| 7-10-81 | 75 | 0.44653 | 0.82448 | 0.085256 max. | 0.0000544 |
| | 330 | −0.89510 | −1.02591 | −0.029507 min. | |

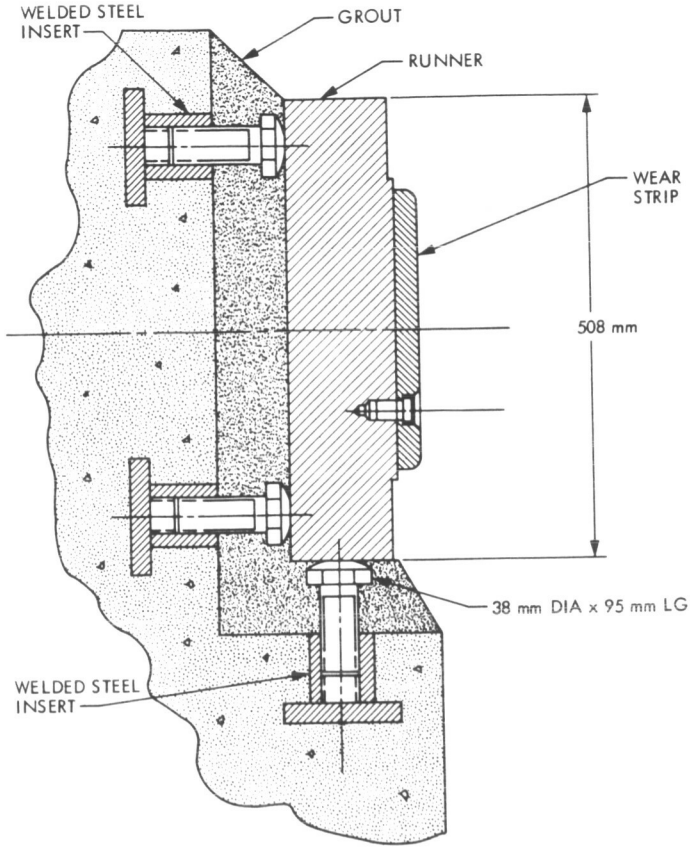



**Fig. 1. Cross section of runner showing alignment bolts and grout**

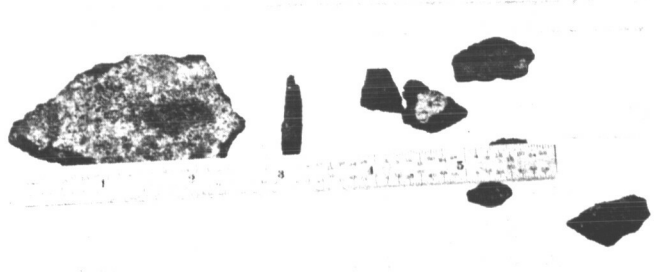


**Fig. 2. Scale removed from runner at grout interface, DSS 14**

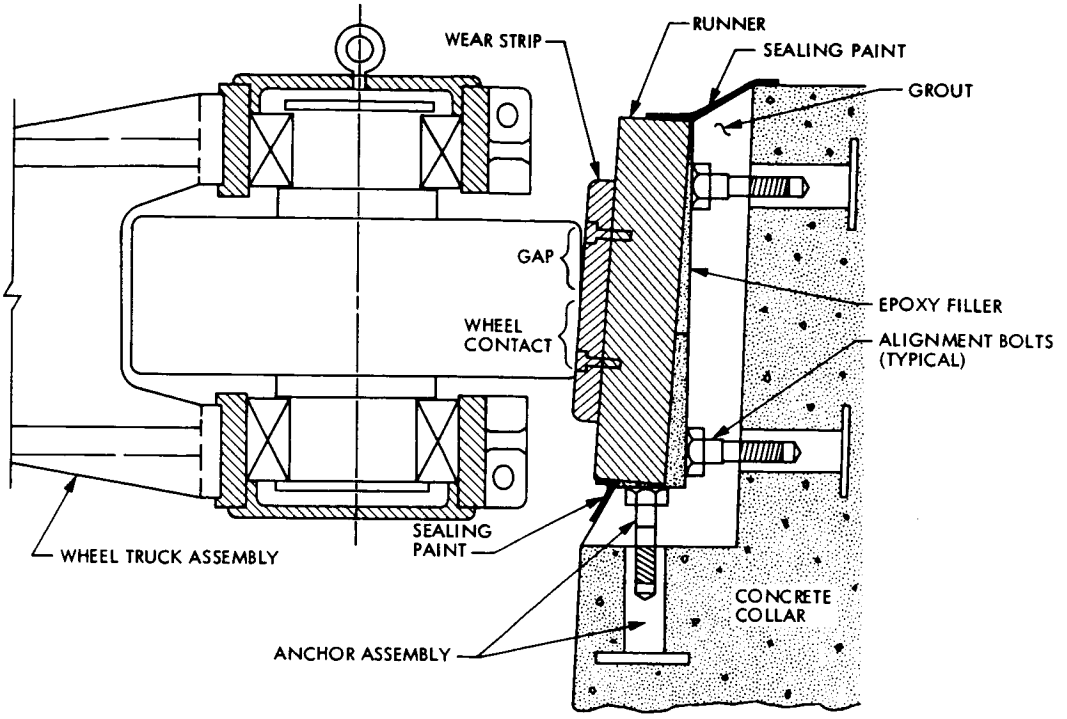



**Fig. 3. Distorted runner at DSS 14; voids filled with epoxy resin**

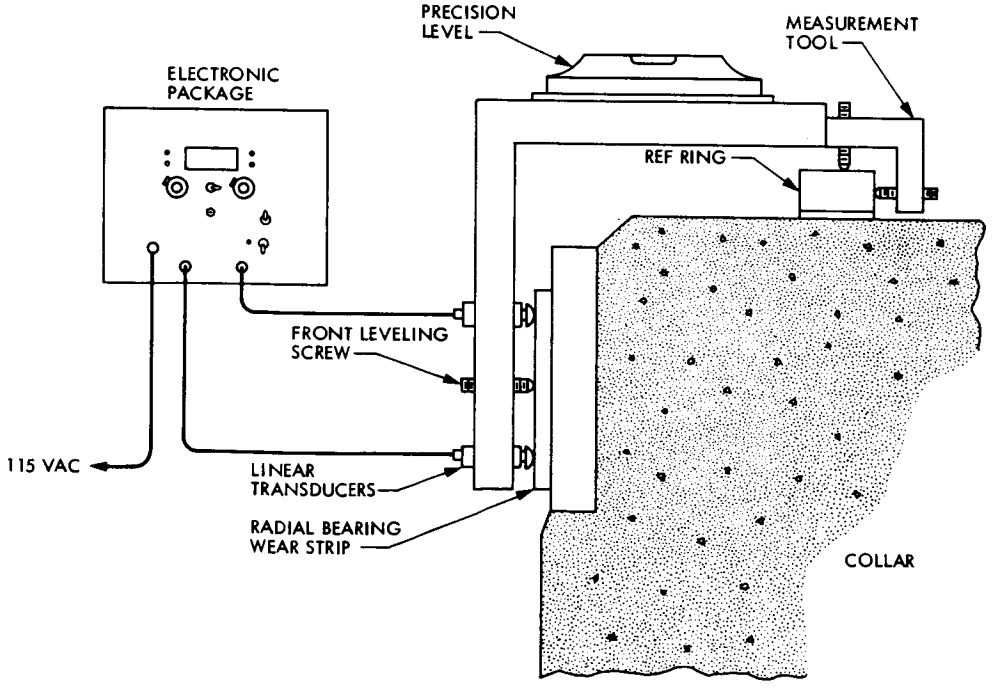



**Fig. 4. Reference ring and its relationship to wear strip**

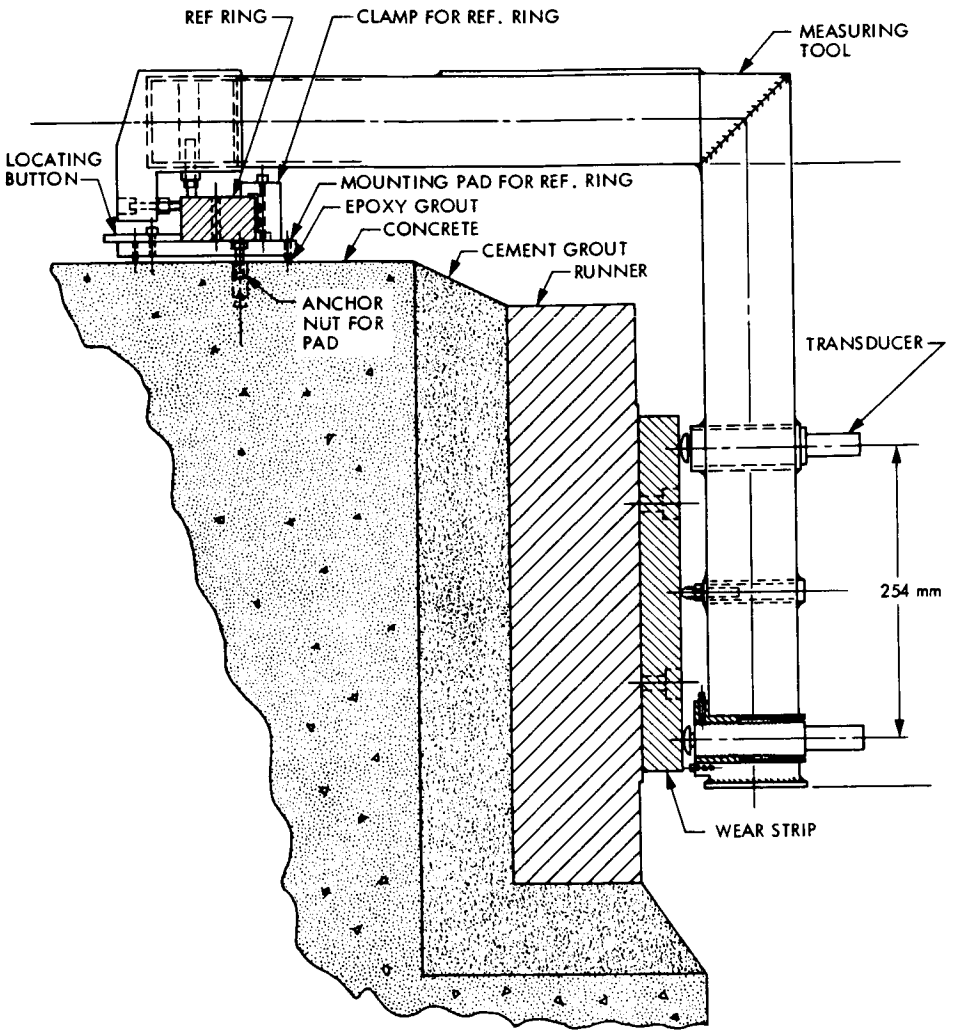



**Fig. 5. Reference ring location and attachment**

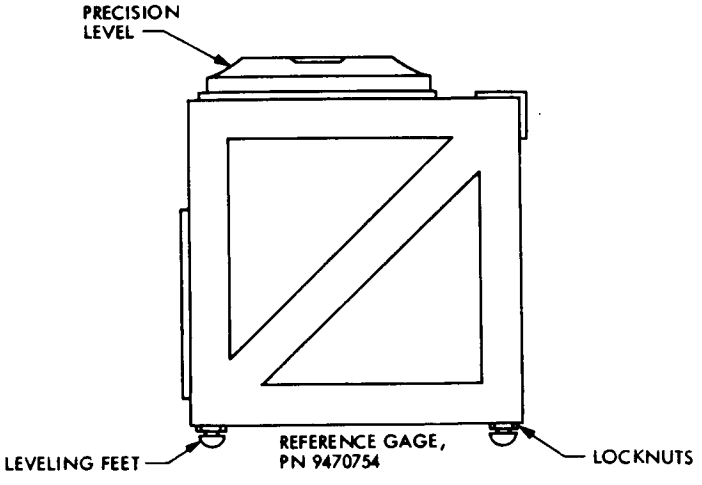



**Fig. 6. Reference gage zeroing fixture**

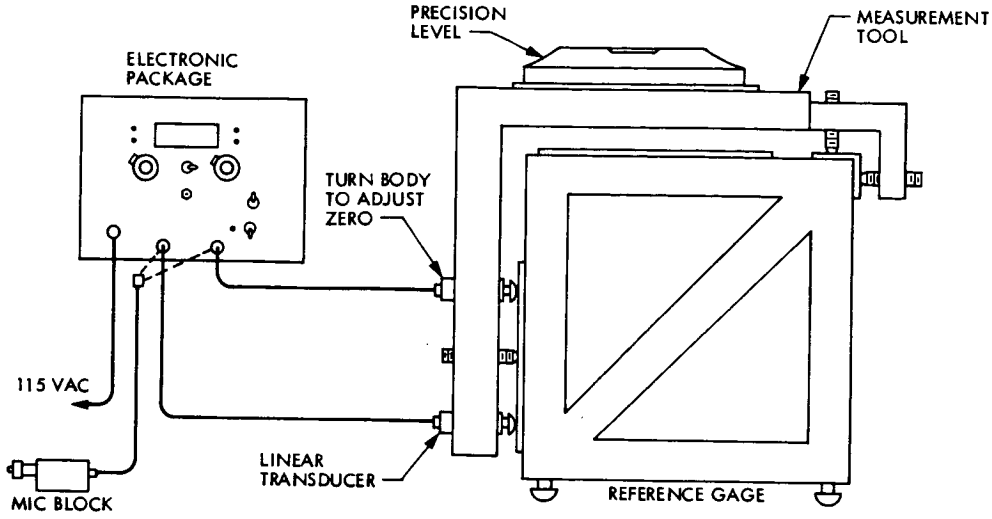



**Fig. 7. Measuring tool on reference gage zeroing fixture**

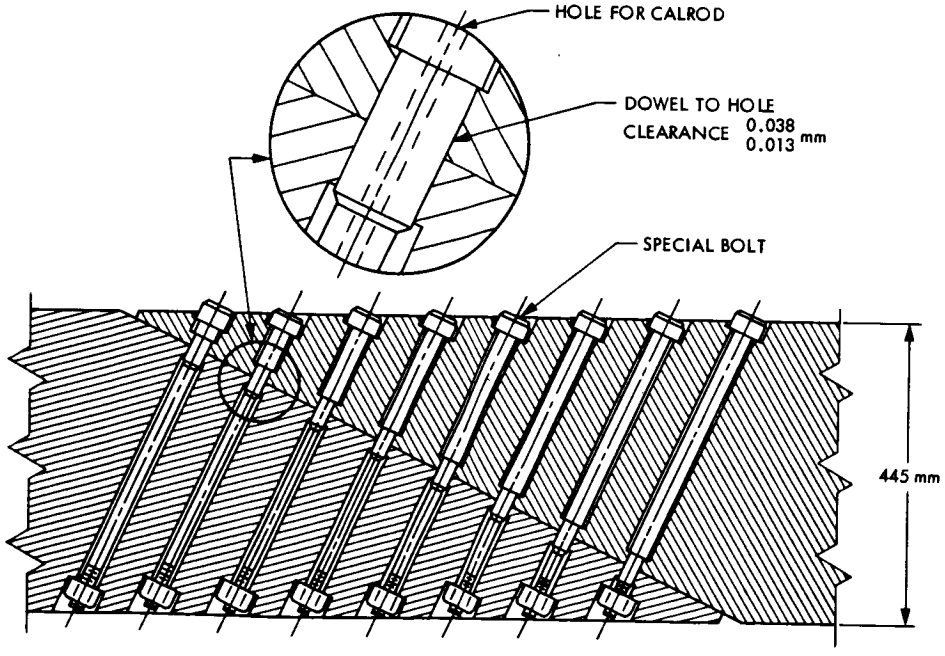



**Fig. 8. Joint design for spare radial bearing runner**

N82-11302



# The Quality Assurance Role in the Deep Space Network

R. L. Sirpilla
Quality Assurance – Ground Data Systems Section

*The following is presented as an informative description of the Quality Assurance role, to the Cognizant Engineering and Project personnel assigned to the Deep Space Network.*

## I. Introduction

The Quality Assurance – Ground Data Systems Section, has played an important role in providing quality engineering and inspection support to the Deep Space Network operations since 1964, as illustrated in Figs. 1 through 15.

In this role, the Quality Assurance Section provides skilled personnel who perform the following functions: quality hardware engineering, quality software engineering, inspection and testing, and an independent assessment to quality, engineering and project requirements. This has proven to be cost effective to the TDA-DSN operations.

By proper training, quality engineering, inspection, auditing, and verifying conformance in a timely manner, the Quality Assurance Section has performed a meaningful task in providing technical capabilities to assure equipment reliability for performance in the Deep Space Network.

## II. Program Plan

Document PP15151-1 (JPL internal document), *Quality Assurance Program for the Deep Space Network*, approved by the Office of Tracking and Data Acquisition, defines the Quality Assurance Section's role and responsibilities.

## III. Budget Guidelines/Resources

The TDA Budget Guidelines provide the description and resources for the Quality Assurance Section to support hardware and software requirements.

## IV. Project Management Plans

A Quality Assurance Task Manager is assigned to develop and implement a Quality Assurance Management and Program Plan in support of each project for the Deep Space Network.

## V. TDA Representation

The Quality Assurance Section Manager attends the weekly TDA meeting as a participating member in representing the Quality Assurance Section for problem solving recommendations.

## VI. Quality Assurance Engineering and Planning

The Quality Assurance Section provides specifications, workmanship standards, process specifications and quality assurance procedures in support of DSN equipment.

In addition, the Quality Assurance Section reviews DSN Engineering documents to ensure that appropriate and adequate quality provisions are included (Ref. 1).

## VII. Procurements/Work Orders

The DSN Quality Assurance Representative reviews statements of work, purchase requisitions and work orders for appropriate quality provisions for implementation.

To be cost effective and without compromising quality, quality requirements are tailored to fit the item being manufactured or purchased.

## VIII. Vendor Surveys

Quality Assurance personnel perform surveys with Engineering and Procurement to determine contractor capabilities for all types of hardware; approvals, conditional approvals or disapprovals are documented.

## IX. Cables/Harnesses

Cables and harnesses are inspected at the source and JPL. Electrical tests are performed 100 percent. These tests have proven to be effective by assuring that cables are reliable, thus reducing the requirement for spares in the Deep Space Network.

## X. Mechanical Inspection

Quality Engineering and inspection support vendors, JPL, and the DSN Stations. In-house mechanical inspection equipment is utilized in checking dimensions and characteristics with Engineering drawings.

At the DSN Stations some of the most important factors are checking alignments of panels, and inspection of structures to verify adherence to requirements.

The Mechanical Inspection Department also has metrology equipment and provides calibration of mechanical tooling.

## XI. Quality Hardware Engineering

The Quality Assurance Section provides capable and qualified Engineering personnel to assist technical engineers in determining design and fabrication requirements.

## XII. Inspection Personnel

The Quality Assurance Section provides inspection personnel to perform inspections in accordance with quality engineering and technical engineering requirements.

## XIII. Quality Software Engineering

Quality Assurance Software Engineering provides independent audit of design, audit of code versus design, audit of test plans versus requirements, and audit of test results.

## XIV. Receiving/On-Lab Inspections

The Quality Assurance Section provides inspection support for the receiving inspection areas and the various labs where DSN hardware is fabricated and tested.

## XV. Transfer Agreements

The Quality Assurance Section reviews all transfer agreements for both hardware and software.

Quality assurance supporting inspection documentation is utilized to verify conformance to engineering and quality requirements in support of transfer agreements.

## XVI. Audits

Audits for conformance to quality requirements are performed periodically at both the contractor and JPL facilities. Corrective actions are requested when required based on findings.

## XVII. Workmanship Assurance (W.A.) Program

The Quality Assurance Section provides a coordinator to the W. A. Program Office for purposes of auditing to JPL requirements and to ensure conformance to TDA-DSN requirements.

## XVIII. Documentation

The Quality Assurance Section maintains records for all activities performed in support of DSN requirements. Records are microfilmed after one year for retention. AODC and ACS are utilized in distributing various reports to cognizant personnel.

## XIX. Training and Certification Programs

The JPL Quality Assurance Training and Certification Program provides real-time courses and video related programs with qualified and certified instructors in the proper techniques of hand soldering, cabling and harnessing, wire wrap, cable repair, I. C. replacement, P. C. module repair, and flat ribbon cable.

A certification is issued to each successful attendee.

Instructors at the Deep Space Stations are certified and qualified by the JPL Quality Assurance Instructors.

Video tapes are provided to the DSN Stations and to other NASA centers.

## XX. Progress Reports

Monthly progress reports are distributed to TDA and DSN cognizant personnel to inform them of current status of hardware and software.

## Reference

1. DSN Standard Procedure, DP512360. Jet Propulsion Laboratory, Pasadena, Calif. (JPL internal document).

Fig. 1. Receiving inspection of DSN equipment

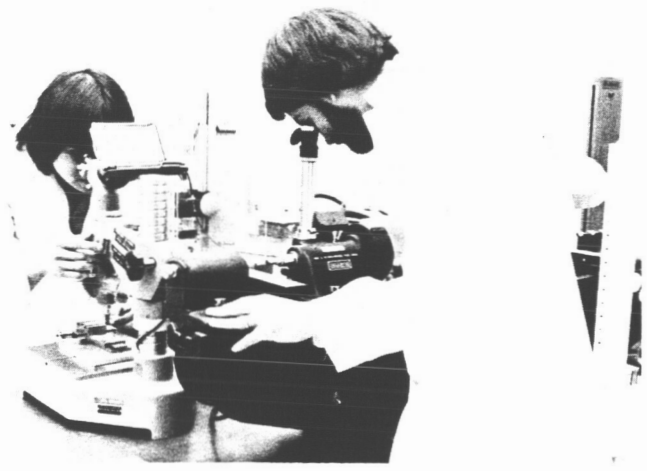


Fig. 3. Metrology and gage calibration

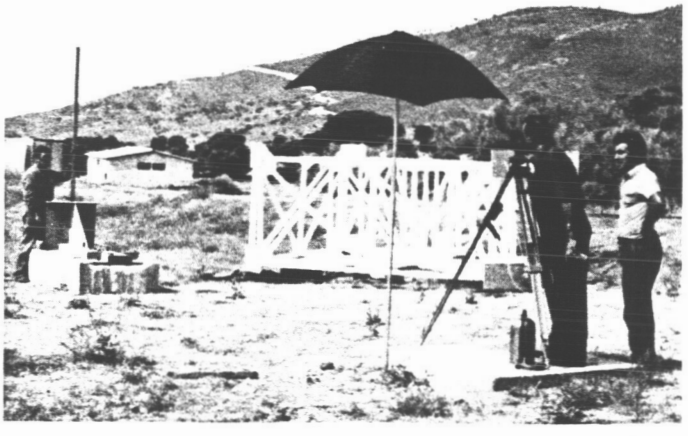


Fig. 2. Leveling out blocks for site erection of quadripod in Spain

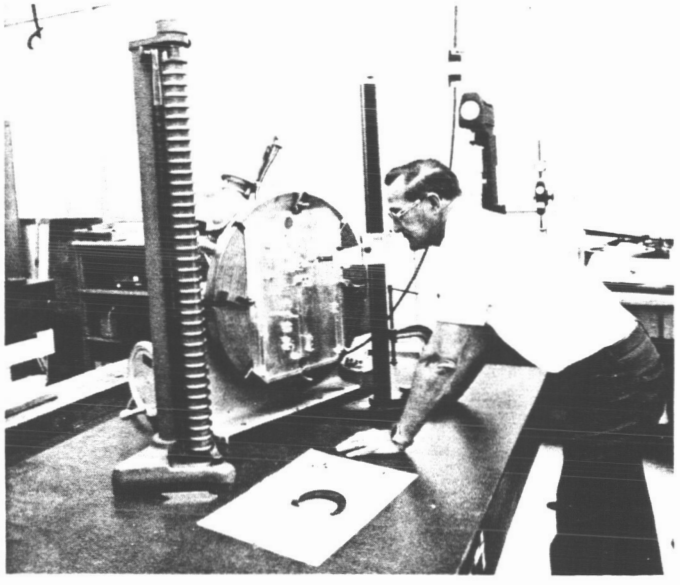


Fig. 4. Rotab plate-check inspection of in-house fabricated part

ORIGINAL PAGE IS
OF POOR QUALITY

Fig. 5. Checking precision piece parts on Intra-Ron for roundness, flatness, and perpendicularity

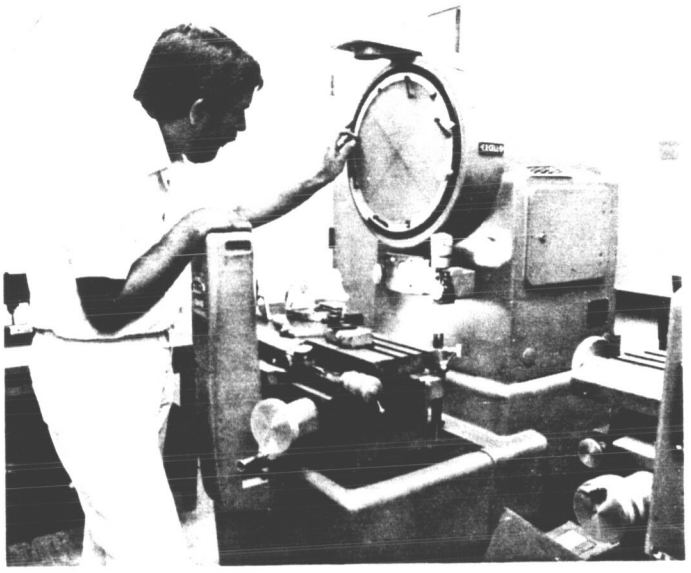


Fig. 7. Use of optical comparator for inspection of precision machine parts

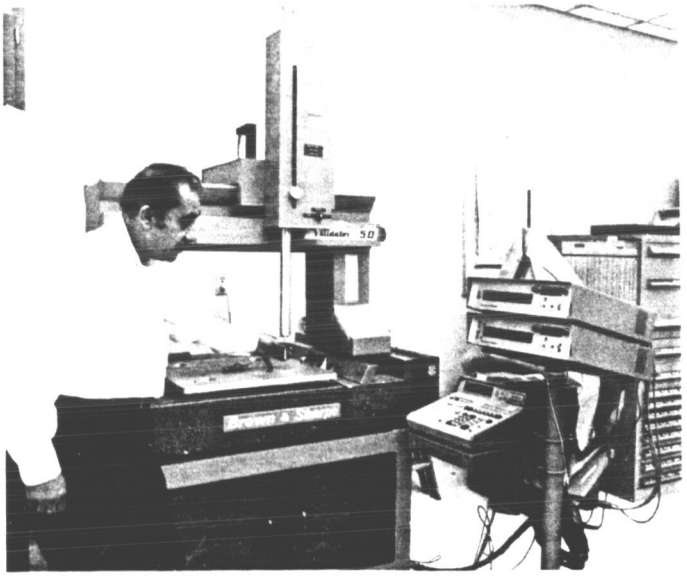


Fig. 6. Use of XYZ automated measuring equipment

Fig. 8. Electrical testing of cables

ORIGINAL PAGE IS
OF POOR QUALITY



C-3

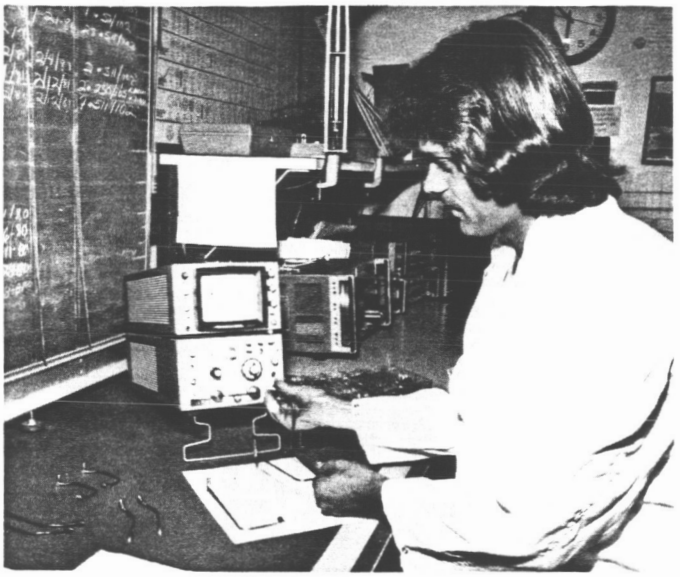


Fig. 9. Electrical testing of hardlines

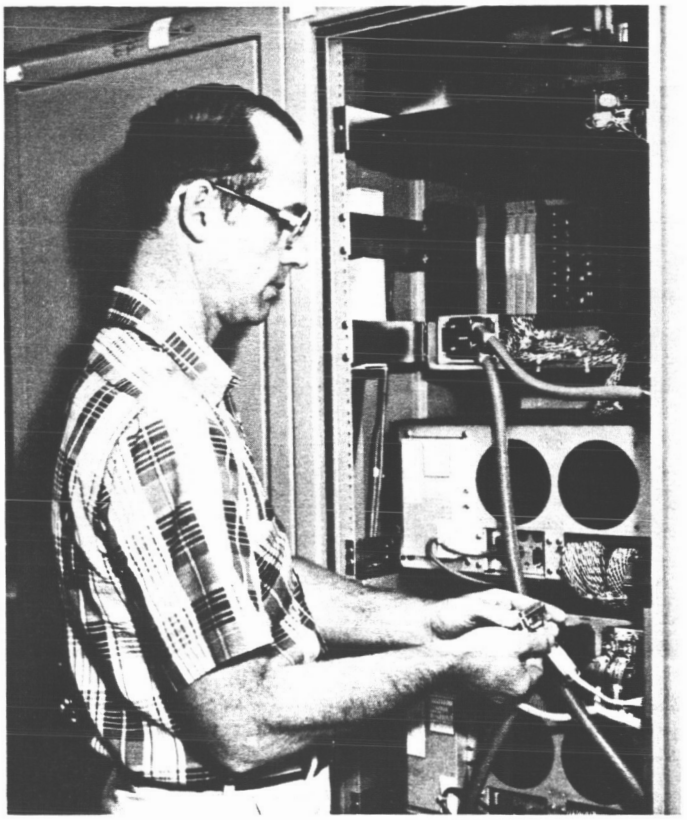


Fig. 11. Inspection of cabinet assembly at CTA-21

Fig. 10. Screening of purchase requisitions, drawings, and specifications for implementation of quality assurance requirements

Fig. 12. Auditing software code to engineering documentation

ORIGINAL PAGE IS
OF POOR QUALITY

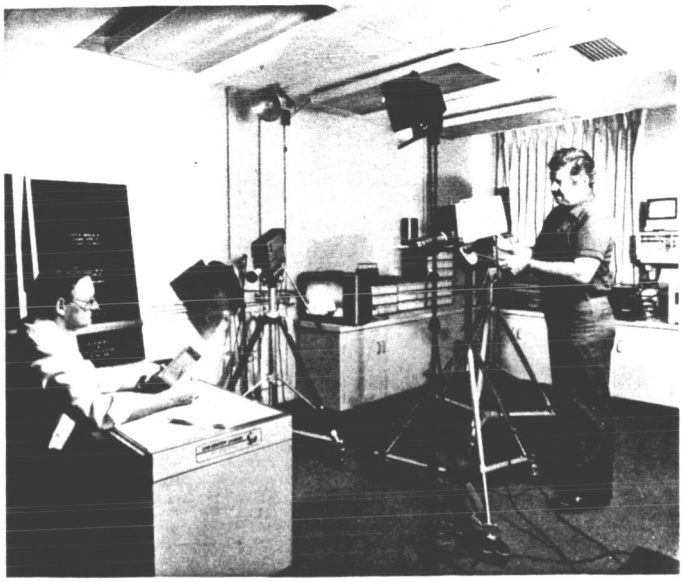


Fig. 13. Developing and producing video training and certification programs in electrical and mechanical assembly for JPL and DSN stations

Fig. 14. Instructor demonstrating correct method of electronic assembly

Fig. 15. Training classes are periodically audited for conformance to requirements by supervision and management

N82-11303



# Networks Consolidation Program: Maintenance and Operations (M&O) Staffing Estimates

J. P. Goodwin
Network Support Section

*This article describes the method used to make the initial estimation of M&O staffing for the DSN Mark IV-A design.*

## I. Introduction

The current Deep Space Network (DSN) configuration, designated Mark III, will be undergoing major changes during its evolution to a network design to be known as Mark IV-A. The Mark IV-A will consolidate Deep Space and High Elliptical Earth Orbiter (HEEO) missions tracking and will implement centralized control and monitoring at the Deep Space Communications Complexes (DSCC). One of the objectives of the new network design is to reduce Maintenance and Operations (M&O) costs. To determine if the system design meets this objective it was necessary to develop an M&O staffing model for Goldstone which could be used to estimate the staffing levels required to support the Mark IV-A configuration. Once a validated model was developed, the staffing estimations for Mark IV-A could be compared to staffing required for the current DSN and STDN configuration to determine if the new design could reduce M&O costs. The study was performed only for the Goldstone Complex and the Program Office translated these estimates for the overseas complexes to derive the Network estimates.

The approach used in developing the staffing model was to first look at the changes to DSCC configurations. As these changes were defined, the impact of the new design on Network Control Center and Network Support staffing was also determined. This developed an image of the new configuration which, measured against maintenance and operations of the DSN Mark III configuration, was used to estimate staffing levels. As will be described later in the article, the maintenance and operations tasks were separated, and the model estimated numbers of persons required for each task. This was just a means of breaking down the total job and sizing it, and should not be construed as being the method the DSN will use in performing the job. Each DSCC will decide what is the best approach for its local conditions and will allocate tasks to the position descriptions justified locally.

The maintenance model used in estimating manpower changes from Mark III to Mark IV-A was validated by using historical results from the Mark II to Mark III Data System transition.

## II. Deep Space Communications Complexes

### A. Operations

The baseline system design of the Mark IV-A configuration is described in Ref. 1. This design contains some equipment

from the current Mark III configuration, some from the STDN configuration, and some new equipment. All of the new equipment, and much of the older equipment, has been instrumented for centralized control and monitoring in the Mark IV-A design. A typical physical layout of the equipment is indicated in Fig. 1. This layout would locate the consoles for the centralized monitor and control adjacent to each other; these are the link monitor and control consoles and the complex monitor and control console. In addition to staffing these consoles, personnel are needed in the Signal Processing Center (SPC) equipment room to operate the locally controlled equipment, and in the recorder area for their operation.

The design process of Mark IV-A, which evolved to the configuration described in Ref. 1, had a continuing participation by operations personnel. As this configuration evolved, it was studied for operational feasibility and possible optimization of operational tasks. The equipment control areas, workload, operational scenarios, and assumed methods of operating equipment were all considered in determining the operational staffing levels. This led to the estimation of nine persons per shift to operate the DSCC as follows: one person to supervise all the complex activity; one person to control the complex resources from the complex monitor and control console; one person for each of four links to operate the monitor and control console devoted to a project; one person to support the locally operated equipment; and two persons to operate and do periodic maintenance of the recorders. The usual operating positions for these personnel are indicated in Fig. 1.

The station's operations personnel requirements were established from a zero base, and we consequently had to add two personnel to the complex staff to assume clerical and documentation work that had been done by deleted station staff. Other operations personnel such as management, clerical, documentation, fire and safety, facilities support, etc., were considered to be the same as required by the current configuration.

## B. Maintenance

The staffing necessary for the maintenance tasks at a DSCC were estimated through the following steps:

(1) Established maintenance complexity factors for each of the current subsystems.

(2) Established maintenance complexity factors for the Mark IV-A design using maintenance study team concepts.

(3) Calculated the ratio of Mark IV-A (complexity and quantity) to Mark III (complexity and quantity) to be used as a multiplier factor (MULT).

(4) Weighted (WT) each subsystem by its portion of current overall maintenance workload within that major function: digital, RF, antenna, servo, and communications.

(5) Multiplied factors (MULT) times weighted subsystem workload (WT) to derive equivalency factors (EQ).

(6) Multiplied equivalency factors (EQ) times current staffing levels (MKIII) to derive estimate of Mark IV-A staffing (MKIV).

Table 1 indicates the maintenance complexity factors and quantity factors which were assigned to the current configuration and estimated for the Mark IV-A configuration. The column identified as "MULT" is the ratio of the "Total" column of the Mark IV divided by the "Total" column of the Mark III portion of the table. It can be seen from the "MULT" column of this table that, in general, the maintenance of Mark IV-A is considered to be more difficult than the Mark III system.

Table 2 indicates how the equivalency factors were derived. Where a major function covered the maintenance of more than one subsystem, each subsystem was apportioned the percentage of the total workload it created for that major function. For instance, under "Digital" the command subsystem accounts for 18 percent of the digital workload, the telemetry subsystem accounts for 25 percent of the digital workload, etc. Some major functions such as antenna, servo, and communications consisted of only one subsystem so the equivalency factor was the same as the multiplier factor. For the digital and RF major functions, the weighted subsystem percentages were multiplied by the subsystem multiplier factors to produce the equivalency factor. The individual equivalency factors in the digital and RF functions were summed to give one equivalency factor for the digital function and one for the RF function.

The equivalency factor times the current staffing for each function produced the staffing estimate for the Mark IV-A maintenance task.

The algorithm produced a 33 percent increase in the on-site maintenance organization. In addition to the on-site maintenance at each complex which, in general, requires substitution of a malfunctioning module, board, etc., with an operable one from spares, there is another level of maintenance required at the complex which involves the repair of the module, board, etc. This level of maintenance is performed at the Complex Maintenance Facility (CMF), whose staffing was estimated by multiplying the current CMF staffing by the percentage change (+33 percent) in technicians performing on-site maintenance.

The number of personnel handling logistics for the complex was estimated by multiplying current staffing by the estimated parts percentage increase. The number of parts handled was derived by taking the fraction of total parts used by a major function in the Mark III era and multiplying it by the equivalency factor for that function. The summation of these products is the estimation of the ratio of Mark IV-A parts compared to the current number of parts. This calculation resulted in a predicted 13 percent increase in the Mark IV parts over Mark III and was used to calculate the increase in logistic parts handlers.

## III. Control Center Operations Section

The Mark IV-A design has not added or changed any work stations at the Network Control Center (Section 371), but it has changed the workload due to the added set of HEO missions. The Network Operations Planning Engineers (NOPE) and the Scheduling and Analysis Groups required additional staffing to support the added missions. The Operations Control Group needed support because the new mission set creates a much higher quantity of passes per day. The Analysis Group's requirements in radio metric support increased owing to the added missions and numbers of antennas.

## IV. Network Support Section

The new configuration support requirements (Section 377) can be absorbed by current staffing once the transition is completed, except that documentation support will increase by one (1) to accommodate the added technical writing and editing required for the HEO missions Network Operations Plans (NOP).

## V. Summary

The NCP staffing requirements discussed in this article were developed from the baseline design and were presented in the January 1981 NCP Fourth Formal Review. The staffing study projected that Goldstone (DSN) will require an additional two personnel, the Operations Control Center (Section 371) an additional 10 personnel, and DSN Support (Section 377) an additional person to support the Mark IV-A baseline configuration.

## Reference

1. Gatz, E. C., "Networks Consolidation Program System Design," in *The Telecommunications and Data Acquisition Progress Report 42-63, March and April 1981*, p. 150-153, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1981.

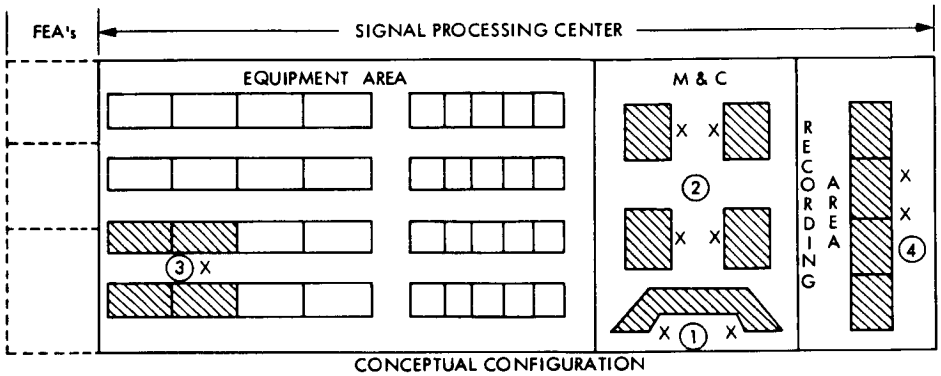
**Table 1. NCP subsystem complexity**

| Subsystem | Mark III | | | | Mark IV-A | | | | | | | MULT |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 64 m | 34 m | 26 m | Total | SPC | 64 m | 34A* | 34B* | 34C* | 9 m | Total | |
| Ant | 10 | 1 | 1 | 12 | – | 10 | 1 | 1 | 1 | 1 | 14 | 1.17 |
| Servo (APS) | 3 | 1 | 1 | 5 | – | 3 | 1 | 1 | 1 | 1 | 7 | 1.4 |
| RF | | | | | | | | | | | | 1.53 |
| RX/EX | 6 | 2 | 2 | 10 | 20 | – | – | – | – | – | 20 | 2.0 |
| TX | 3 | 1 | 1 | 5 | – | 3 | 1 | – | – | 1 | 5 | 1.0 |
| UWV | 4 | 3 | 2 | 9 | – | 4 | 3.5 | 1.5 | 1 | 2 | 12 | 1.33 |
| Digital | | | | | | | | | | | | 0.95 |
| CMD | 2 | 2 | 2 | 6 | 4 | – | – | – | – | – | 4 | 0.67 |
| TLM | 2 | 2 | 2 | 6 | 6 | – | – | – | – | – | 6 | 1.0 |
| TRK | 1 | 1 | 1 | 3 | 3 | – | – | – | – | – | 3 | 1.0 |
| FTS | 1.5 | 1 | 1 | 3.5 | 2 | – | – | – | – | – | 2 | 0.57 |
| RS | 1 | – | – | 1 | 1 | – | – | – | – | – | 1 | 1.0 |
| GCF (tapes) | 2 | 2 | 2 | 6 | 4 | – | – | – | – | – | 4 | 0.67 |
| M&C (SSC) | 1 | 1 | 1 | 3 | 4 | – | – | – | – | – | 4 | 1.33 |
| TSS | 1 | 1 | 1 | 3 | 2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 3 | 1.0 |
| TFC (MMA) | 1 | – | – | 1 | 2 | – | – | – | – | – | 2 | 2.0 |
| Comm term/uwave | 1 | 1 | 1 | 3 | 2 | – | 1 | – | – | – | 3 | 1 |

*34A = transmit/receive
34B = DSN-X, VLBI and STDN-S
34C = DSN-X and SETI

**Table 2. Maintenance and integration, manpower derivation**

| | | | Mark III | Mark IV-A |
|---|---|---|---|---|
| **Digital** | | | | |
| SS | WT MULT EQ | | | |
| CMD | 0.18 × 0.67 = 0.12 | | | |
| TLM | 0.25 × 1.0 = 0.25 | | | |
| TRK | 0.21 × 1.0 = 0.21 | | | |
| FTS | 0.08 × 0.57 = 0.05 | | | |
| RS | 0.04 × 1.0 = 0.04 | | | |
| GCF | 0.08 × 0.67 = 0.05 | | | |
| M&C | 0.08 × 1.33 = 0.11 | | | |
| TSS | 0.04 × 1.0 = 0.04 | | | |
| TFC | 0.04 × 2.0 = 0.08 | | | |
| | 1.0 0.95 | | | |
| | EQ × MKIII = MKIV-A | | | |
| Technicians: | 0.95 × 9 = 8.55 = 9 = +0 | Technicians | 9 | 9 |
| Engineers: | 0.95 × 4 = 3.8 = 4 = +0 | Engineers | 4 | 4 |
| **RF** | | | | |
| SS | WT MULT EQ | | | |
| RX/EX | 5/12 × 2.0 = 0.84 | | | |
| TX | 3/12 × 1.0 = 0.25 | | | |
| UWV | 4/12 × 1.33 = 0.44 | | | |
| | 1.53 | | | |
| | EQ × MKIII = MKIV-A | | | |
| Technicians: | 1.53 × 12 = 18.36 = 19 = +7 | Technicians | 12 | 19 |
| Engineers: | 1.53 × 3 = 4.59 = 5 = +2 | Engineers | 3 | 5 |
| **Ant** | | | | |
| | EQ × MKIII = MKIV-A | | | |
| Technicians: | 1.17 × 8 = 9.36 = 10 = +2 | Technicians | 8 | 10 |
| Engineers: | 1.17 × 2 = 2.34 = 3 = +1 | Engineers | 2 | 3 |
| **Servo** | | | | |
| | EQ × MKIII = MKIV-A | | | |
| Technicians: | 1.4 × 4 = 5.6 = 6 = +2 | Technicians | 4 | 6 |
| Engineers: | 1.4 × 1 = 1.4 = 2 = +1 | Engineers | 1 | 2 |
| **Comm** | | | | |
| | EQ × MKIII = MKIV-A | | | |
| Technicians: | 10 × 1.0 = 10 = +0 | Technicians | 10 | 10 |
| Engineers: | 2 × 1.0 = 2 = +0 | Engineers | 2 | 2 |




| AREA | EQUIPMENT |
|---|---|
| 1 | COMPLEX M & C CONSOLE |
| 2 | LINK CONSOLES |
| 3 | RDA, SDA, ODA, 9-M UWV, GSTDN BFS |
| 4 | TAPE RECORDERS |

**Fig. 1. Signal Processing Center conceptual configuration**

N82-11304



# NOCA Upgrade

G. D. Gould
Control Center Operations Section

*The Network Operations Control Area (NOCA) of the Space Flight Operations Facility (SFOF) is being upgraded in both hardware and software capabilities to provide the visibility and operability necessary for the operations control function to be performed within the present technological environment of the tracking network. This article addresses the hardware implementation currently in progress that provides a standard operating console with increased display capacity and improved ergonomics as a replacement for the existing, outmoded, circular consoles original to the SFOF.*

## I. Introduction

The Network Operations Control Area (NOCA) is located in the Space Flight Operations Facility (SFOF) of the Laboratory and occupies the west half of the main operations room of the SFOF. The SFOF, especially the main operations room, has traditionally been a showplace of the Laboratory and was designed to accommodate and impress visitors and spectators as well as to provide a functional facility for operational control activities of the Laboratory. The first operational use of the main operations room was in support of the Ranger VII mission in July of 1964. The nature of operations control at that time was people-intensive and lent itself to the additional purpose of the facility, to meet spectator expectations of a space center. Within a few years the functional capabilities of the main operations room became inadequate for the requirements of the larger, more complex JPL missions and, in 1968, the DSN Network Control task was moved into the main operations room. The present NOCA includes the principal function performed in the main operations room of the SFOF today. Advancements in the state-of-the-art to collect and communicate vast quantities of data by the Network, and the addition of several new distinct DSN Systems as well as increases in the sophistication and complexity of systems operations have far exceeded the capability of the 1964 design to provide a reasonably efficient operations control function to the Network with any degree of cost-effectiveness. The following sections of this article explain the hardware approach to meeting the requirements of functionalism, operability, and aesthetics.

## II. Functional Description

The NOCA provides the environment for a central point of operational direction and control to the facilities of the DSN in real time and a single point of access to the DSN for Network users in real time. The major tasks accomplished in performing these functions are the coordination of Network activities, monitoring of Network performance, providing facilities with supporting information, resolving conflicts at the operational level, and allocating resources based on real-time priorities.

### A. Coordination

The Network Operations Control Team (NOCT) operating within the NOCA coordinates Network activities and user requirements by voice communication with the Deep Space Stations, Mission Operations Centers, data processing facilities,

and the Ground Communications Facility in real time. The voice communication involved in the interfacility and intrafacility coordination performed by a single Operations Controller utilizes a total of twenty-eight dedicated circuits.

### B. Monitoring

The NOCT insures the integrity of DSN performance by monitoring facility configuration and systems performance parameters of the various systems in real time. Performance and status data are provided to the NOCA by the Deep Space Station subsystems and by real-time system monitoring computers in the SFOF. The monitoring of status and performance by a single Operations Controller utilizes a total of thirty-five discrete display channels.

## III. Hardware Improvements

### A. Functionalism

The basic convex structure of the circular console (Fig. 1) limited the available viewing angle of an operator to an area that – when combined with the low profile – would permit the observation of only two nine-inch CRT displays. The first improvement was to reverse the structure (Fig. 2) to one of a concave presentation to the operator, which increased the operator viewing angle. The second improvement in functionalism was to increase the height of the console to accommodate vertical installation of more than one digital television (DTV) display. In addition to increasing single operator viewing from two to seven simultaneous displays, the new dimensions provided increased operator working area and enabled the colocation within the console racks of displays and associated selectors.

The integral construction of the original circular consoles would not permit horizontal expansion. The replacement console is modular in construction and allows for further expansion if required with no sacrifice to basic design.

The increased mounting surfaces of the replacement console permitted the installation, within operator attention zone, of equipment that was previously peripheral to the console; (i.e., GMT clock, special purpose telephones, and a warning timer).

### B. Operability

In the design of the original circular console, operability was sacrificed for aesthetic appeal. This is understandable considering that human factors engineering has only recently become a popular consideration. The human factors guides and prior SFOF-peculiar studies applied to the development of the replacement console are given at the end of this article for reference. The entire approach to the replacement of the circular consoles was one of prototype development and, consequently, a period of three months was allowed for use and evaluation of the prototype by the operators that would eventually use it.

Improvements realized with the replacement console are most obvious when seated at the old, then moving to the new, but the ease of reading the new fourteen-inch displays as opposed to the previous nine-inch displays and being able to observe six or seven simultaneous displays as opposed to the previous two, are the most noticeable improvements. Improvements were made in the operator seating posture, reach requirements, button access, and speaker locations.

One compromise was made to appearance; the DTV tilt angle was increased slightly beyond optimum to reduce the height of the console slightly and provide a more appealing profile.

### C. Aesthetics

The improvement of the aesthetic appeal of the main operations room of the SFOF to the general public as a space center facility was beyond the scope of the NOCA upgrade.

Although emphasis was placed on functionalism and operability in the design of the replacement console, the basic criterion was maintained that any resulting change to the appearance of the SFOF must conform to a general perception of a space center. The color chosen, for example, for the new console is slightly lighter in shade than the recommended optimum for DTV viewing enhancement, but blends well with the overall operations room and background displays. It was determined that this tradeoff was acceptable due to the normally reduced lighting in the room. The prototype console was assembled and installed for viewing in the main operations room for several days to collect a wide variety of observers' opinions on the aesthetic result, which proved highly favorable.

The arrangement of the individual consoles in the final configuration was determined to some extent by human factors considerations, but also by aesthetic appeal.

Since the permanent installation of the first replacement console in the NOCA, it has become a favorite backdrop for film crews and public relations photographers, which is probably a satisfactory endorsement.

## IV. Summary

As of this reporting period, all hardware required for the replacement of the original SFOF circular consoles with the

NOCA upgrade standard operating consoles has been procured and is available to begin replacement immediately following Voyager 2 Saturn encounter critical activities. No direct cost savings are expected to be realized from the hardware replacement, but the improved operability of the NOCA should help to minimize any future costs that would result from the Networks Consolidation increased loading. The largest part of the NOCA upgrade is concerned with software implementation and has not been addressed in this article. It is currently planned that most of the NOCA upgrade software will be implemented as part of the Networks Consolidation Program schedule.

## References


1. Vancott, H. P., and Kinkade, R. G., *Human Engineering Guide to Equipment Design*, McGraw-Hill, Revised Edition, 1972.
2. Woodson, W. E., and Conover, D. W., *Human Engineering Guide for Equipment Designers*, University of California Press, Berkeley, 1966.
3. *Some NASA Contributions to Human Factors Engineering*, SP-5117 Technology Utilization. National Aeronautics and Space Administration, Washington, D. C.
4. *NCS Operations Effectiveness Study, 3rd module, Technical Studies on Network Operations Performance Analysis*, Dosher Associates, 1972.


Fig. 1. Original circular consoles

ORIGINAL PAGE IS
OF POOR QUALITY


**Fig. 2. Replacement console**

ORIGINAL PAGE IS
OF POOR QUALITY

N82-11305



# Relocation of the Deep Space Network Maintenance Center

K. F. Beutler
Deep Space Network Support Section

*The Jet Propulsion Laboratory maintains a Deep Space Network (DSN) Maintenance Center (DMC), whose task is to engineer and manage the repair and calibration program for the electronic and mechanical equipment used in the tracking stations located at Madrid, Spain, and Canberra, Australia. The DMC also manages the Goldstone Complex Maintenance Facility (GCMF), whose task is to repair and calibrate the Goldstone electronic and mechanical equipment. The combined DMC and GCMF is operated by the Bendix Field Engineering Corporation under contract to JPL. Prior to January 1981, the DMC and GCFM were located at the Goldstone Echo site. This paper discusses the rationale for moving the facility to Barstow, California, and the benefits derived from the move.*

## I. Introduction

The DSN Maintenance Center (DMC) provides the centralized technical repair function for the Deep Space Network. Prior to January 1981, it was located at the Goldstone Deep Space Communications Complex, 72.4 km north of Barstow, California.

As early as 1970, a problem was recognized in relation to having the DSN Maintenance Center located near Deep Space Station 12. Specifically, the proximity of the DMC to DSS 12 resulted in interference to DSS 12 tracking operations by spurious radiation (RFI) of both S- and X-band signals emanating from RF test equipment used in the DMC. Owing to the interference, controls had to be placed on RF testing activities at the DMC. Variations in shift schedules to accommodate the tracking, test, and maintenance activities created extra expense and had adverse effects on employee morale. Because of these problems, a study was made of the feasibility of relocating the DMC and GCMF to an off-site area. This study was initiated in July 1979.

## II. Statement of Problem

During the early 1970's when Goldstone was developing an active RFI program, it was noticed that during the testing and repair of certain RF exciter modules at the DMC, DSS 12 would report spurious spikes on its received signal. Investigation proved that signal generators being used at the DMC, which was 250 meters northwest of DSS 12, were creating the interference. Furthermore, the DMC and GCMF were located in five separate buildings throughout the Goldstone Complex. This dispersement over a wide geographical area created both supervisory and communications problems which had a detrimental effect on operations. Decentralized activity required

extra transportation, created nonproductive time due to the increased travel requirements, and caused equipment backlogs and excessive overtime.

As a quick-fix approach, the DMC explored the feasibility of moving the RF portion of the facility to the Microwave Test Facility, which is located approximately 2 km from DSS 12 and has natural terrain as RF shielding. This approach was abandoned because of inadequate available floor space.

Another option was to contract the work out to a vendor, but that approach was deemed too costly because of additional spares which would have to be procured to keep the supply pipeline full. The requirement for additional floor space to support existing and planned increases in workload was also considered.

## III. Solution of Problem

An initial finding of the feasibility study was that no DSN operations requirements precluded the relocation of the DMC and GCMF to an area remote from the Goldstone Complex. Based on this finding, approval was obtained from the Telecommunications and Data Acquisition Office to relocate the facility in Barstow, California.

In designing the relocation plan, the following criteria were followed:

(1) The DMC/GCMF functional capabilities had to be preserved.

(2) Any relocation was to have a minimum adverse impact on personnel and morale.

(3) The DMC/GCMF was to provide support continuity during relocation.

(4) Energy conservation guidelines and goals had to be met.

(5) Cost savings must be realized from the relocation.

A suitable facility was found in central Barstow that had the necessary floor space and also met the required selection criteria. An extensive RFI compatibility study revealed that no mutual RFI condition would be created as a result of the DMC/GCMF operations in the Barstow area.

Based on facility availability and RFI compatibility, a final study was conducted and a presentation of the results made to cognizant JPL management personnel. The direct cost savings, cost avoidance, and energy savings justified the decision, made in early 1980, to relocate the DMC/GCMF as well as some of the Goldstone indirect tracking support functions. A formal relocation plan was developed, approved and placed into effect coincident with completion of facility modifications. The actual relocation occurred in March 1981.

## IV. Benefits

In addition to eliminating the DSS 12 RFI problem, relocation of the DMC/GCMF to the Barstow area provided the opportunity to effect significant cost savings and cost avoidance in the areas of personnel transportation, facility operations, fuel, and future Goldstone Facility construction. The more significant benefits are described below:

(1) The existing RFI problem between DSS 12 and the DMC was eliminated.

(2) All DMC/GCMF functions were consolidated into a common facility. This will improve operational effectiveness and reduce costs.

(3) The new facility benefits from commercial power rates as contrasted to the prime energy rates prevailing at Goldstone.

(4) A savings in transportation costs of $115,000 per year will be realized, and safety will be enhanced.

(5) Barstow, being a more desirable work location than Goldstone, will attract prospective employees.

(6) Sufficient space is available in the new building to allow relocation of other Goldstone functions that are not required in direct support of tracking operations.

## V. Summary

A total of 71 personnel were relocated from Goldstone to the Barstow Facility. Of the total relocated, 54 were DMC and GCMF personnel and 17 were Goldstone Complex support personnel identified as indirect support personnel.

Occupancy of the facility actually began in October 1980, after the offices and laboratories were completed. The final phase of the relocation took place in January 1981. An open house and dedication ceremony, to mark the completion of renovations and occupancy of the facility, was held in March 1981.

N82-11306



# Goldstone (GDSCC) Administrative Computing

H. Martin
Goldstone Operations Section

*The GDSCC Data Processing Unit provides various administrative computing services for Goldstone. Those activities, including finance, manpower and station utilization, deep-space station scheduling and Engineering Change Order (ECO) control are discussed.*

## I. Introduction

Realizing the need in many administrative areas to assimilate data into useful management tools, GDSCC management initiated actions utilizing existing manpower to analyze those areas and develop the necessary software for data collection and reporting. This proved so successful that, in 1978, the GDSCC Data Processing Unit (DPU) was formed to enable full-time dedication to this program. The DP Unit presently consists of three people who perform data entry, systems analysis, software development and sustaining engineering on existing software. The DP Unit utilizes both the Information Processing Center (IPC) Univac computer via telephone lines and dumb terminals, and a Northern Telecom 410 intelligent terminal. The intelligent terminal utilizes a 10-Mbyte hard disk, 8-in. flexible diskettes, 64K memory and a bidirectional matrix printer for local data processing. It will also communicate with the IPC via telephone lines at a 1200-baud rate. Software written for the 410 is in either ANSI COBOL or TAL-II, the terminal's general application language. Software residing on the IPC is written in MBASIC.

## II. Data Processing Unit Activities

### A. Engineering Change Order (ECO) Management

A system was developed and a vigorous program initiated to bring the ECO's for all Goldstone stations under central control with status reporting. This system was followed later by the Engineering Change Management (ECM) system residing on the IPC computers. The DP Unit performs data input and report generation services to the ECM System. Special programs to extract selected data from the system have also been developed (Fig. 1).

### B. GDSCC Deep-Space Station Schedules

Schedules are supported via computer tie-lines rather than the teletype system. Software has been developed to provide the stations and maintenance unit with preliminary schedules extracted from the DSN data base which resides on the IPC. The schedules are in a format which allows maintenance and training data to be written in by the different units. The data are then entered into the intelligent terminal off-line, formatted and transmitted to the IPC for the DSN scheduling group (Fig. 2).

Final schedules are later extracted from the DSN data base and distributed to the appropriate units. Once the scheduled week is completed, real-time changes which occurred are input to the data base and a history file is established. The weekly Station Utilization Report is computer-generated from this history data (Fig. 3).

As an adjunct to the scheduling, a preventive maintenance (PM) system was developed which was compatible with the

DSN schedule. When the preliminary schedule is extracted, a report of those PM's which are due is also generated (Fig. 4). PM's are then planned by the units, input along with the other schedule data, and made a part of the final schedule. After the history schedule is complete, the PM information is extracted by the computer and the PM data base is automatically updated to reflect those which were performed.

## C. Manpower Utilization

Manpower utilization is a system by which the man-hour expenditures by station operators and maintenance personnel are recorded and reported upon. These data provide management with an insight into how personnel time is being expended in various categories (Fig. 5).

## D. Financial System

A system was initiated in the mid 1970's to satisfy a need to provide expenditure information to a level of detail not provided by the JPL financial system. To accomplish this, each account was subdivided into cost centers to which expenditures could be charged. This permitted Complex management to ascertain how much each unit, such as the carpenter shop, the air-conditioning shop, etc., was spending. It also provided information when allocating budgets. Since the initial system covered only those items which were processed through the contractor purchase order system, many expenditures and charges were not reflected in the reports. This system has since been modified to accept data which are extracted from contractor data such as accounts payable, other direct charges (ODC) reports and award fees. Data are also extracted from the JPL SRM Resources Status Report detail backup reports. The system is not, however, operated on an accrual basis. Additional information is also accumulated as to projected man-hour, overtime hour and labor expenditures.

Extracted from these data are such reports as a monthly budget report (Fig. 6), a procurement accomplishment report (Fig. 7), a subcontractor utilization report, plots and reports which provide comparison data on projected versus actual expenditures in dollars, regular labor hours and overtime hours. Expenditure history reports for individual cost centers can also be extracted.

## E. The Vehicle Utilization System

This system became an important function to reduce man-hours which were being required to manually extract data necessary to monitor the expenditures, gas consumption and periodic maintenance of the General Services Administration (GSA) vehicles being leased by the Complex. This is especially true with ever-increasing energy costs. Data are utilized from the daily gas log sheets from the GSA gasoline pumps and from the GSA billing sheets. Reports are generated which provide information on vehicles due for maintenance, monthly mileage and cost summaries, and data required for annual reporting to GSA (Fig. 8).

*** GDSCC ECM DATA BASE EXCEPTION REPORT ***

AS OF 15 JUN 81

(ALL DATES IN YYMMDD FORMAT)

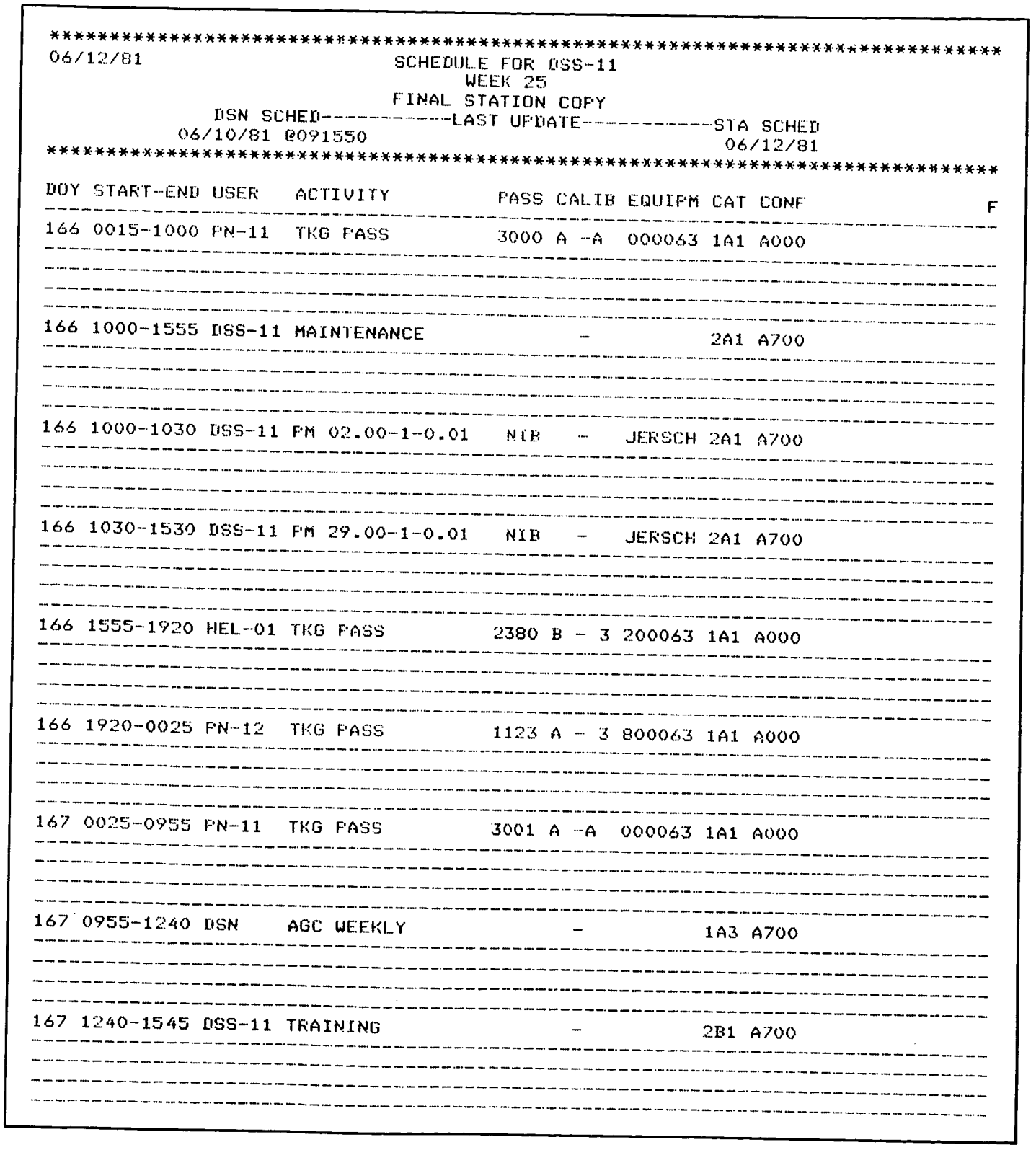
A. ECO'S SHIPPED BY THE COE BUT NOT YET RECEIVED BY THE FACILITY (30 DAYS +):

| | STATION NUMBER | PROJECT | ECONBR | SUBSYS | ECO TITLE | (SF DATE) SHIP DATE |
|---|---|---|---|---|---|---|
| 1 | 10 | MJS | 78.296 | 92.2 | ADD 168KB TRI-CHANNELER @43 Z | 790221* |
| 2 | 10 | DSN | 79.186 | 39.0 | GDSCC AIRPORT RUNWAY UPGRADEF | 800328* |
| 3 | 10 | DSN | 80.055 | 38.14 | FTS CS STD MON INTERFACE M | 810301* |
| 4 | 11 | DSN | 79.006 | 40.3A | SSA UPGRADE OBSOLETE PARTS T | 810128* |
| 5 | 11 | MJS | 80.123 | 34.91 | AUTO-UPLINK TUNE BY MDA S/W$R | 810227* |
| 6 | 11 | MJS | 80.209 | 14.92 | DYA-5081 S/W-50 PPS INTRPTS R | 801230* |
| 7 | 12 | MUL | 77.005 | 34.1A | MDA LOCK MONITOR MODS (MVPH) | 770304* |
| 8 | 12 | DSN | 77.127 | 37.0 | DISH DOCR SAFETY SWITCH-26M | 800420* |
| 9 | 12 | DSN | 80.055 | 38.14 | FTS CS STD MON INTERFACE M | 810301* |
| 10 | 12 | MJS | 80.123 | 34.91 | AUTO-UPLINK TUNE BY MDA S/W$R | 810227* |
| 11 | 12 | MJS | 80.209 | 14.92 | DYA-5081 S/W-50 PPS INTRPTS R | 801230* |
| 12 | 12 | DSN | 81.046 | 38.0 | DELETE VLF RCVR-PHAS RECRDR R | 810511* |
| 13 | 14 | DSN | 76.167 | 38.0 | REMOVE OBSOLETE FREQ AMPS | 770310* |
| 14 | 14 | DSN | 77.015 | 39.5 | SJ DOC FEEDCONE FIRE SENSORS | 791001* |
| 15 | 14 | DSN | 80.055 | 38.14 | FTS CS STD MON INTERFACE M | 810301* |
| 16 | 14 | MJS | 80.123 | 34.91 | AUTO-UPLINK TUNE BY MDA S/W$R | 810227* |
| 17 | 14 | MJS | 80.169 | 37.1 | ADD BU HI-PERFORM MASER-64M$T | 810420* |
| 18 | 14 | MJS | 80.209 | 14.92 | DYA-5081 S/W-50 PPS INTRPTS R | 801230* |
| 19 | 14 | DSN | 81.046 | 38.0 | DELETE VLF RCVR-PHAS RECRDR R | 810511* |

B. ECO'S INDICATING A MISSED COMPLETION DATE (FS = A, B OR C):

| | STATION NUMBER | PROJECT | ECONBR | SUBSYS | (SB) SUBSYSTEM TEST DATE | (SY) SYSTEM TEST DATE | PROJECT DUE DATE | (FS) FACILITY STATUS |
|---|---|---|---|---|---|---|---|---|
| 1 | 10 | DSN | 80.075 | 92.2 | 810503 | 810503 | 810501 | C |
| 2 | 10 | DSN | 80.100 | 94.3 | 801119 | NONE | 810301 | B |
| 3 | 11 | DSN | 77.127 | 39.6 | 810601 | 810601 | 790101 | B |
| 4 | 11 | DSN | 79.101 | 39.1 | 810610 | 810610 | 800601 | B |
| 5 | 11 | DSN | 80.075 | 92.2 | 810610 | 810615 | 810501 | B |
| 6 | 12 | MJS | 79.145 | 29.0A | 810601 | 810601 | 800201 | B |
| 7 | 12 | MJS | 79.145 | 39.6 | 810601 | 810601 | 800201 | B |
| 8 | 12 | MJS | 80.119 | 46.11B | 810411 | 810415 | 800226 | B |
| 9 | 14 | VIK | 77.270 | 29.0D | 810601 | 810601 | 790401 | B |
| 10 | 14 | MJS | 77.321 | 29.0 | 810601 | 810601 | 800101 | B |
| 11 | 14 | DSN | 78.150 | 35.14 | 810511 | 810610 | 810701 | C |
| 12 | 14 | DSN | 78.300 | 39.10 | 810429 | 810429 | 800601 | B |
| 13 | 14 | MJS | 79.012 | 39.6 | 810504 | 810504 | 800901 | B |
| 14 | 14 | DSN | 79.157 | 32.5 | 810531 | 810531 | 801101 | C |
| 15 | 14 | DSN | 79.247 | 31.2 | 810520 | 810520 | 810101 | B |
| 16 | 14 | DSN | 80.075 | 92.2 | 810610 | 810610 | 810501 | B |
| 17 | 14 | DSN | 80.164 | 32.8 | 810601 | 810615 | 810301 | C |
| 18 | 14 | DSN | 80.174 | 39.6 | 810503 | 810503 | 810601 | B |
| 19 | 14 | DSN | 80.189 | 39.92 | 810518 | 810630 | 801124 | C |

**Fig. 1. ECM data base exception report example**



```
**********************************************************************************
06/12/81                        SCHEDULE FOR DSS-11
                                      WEEK 25
                                FINAL STATION COPY
              DSN SCHED----------LAST UPDATE----------STA SCHED
           06/10/81 @091550                            06/12/81
**********************************************************************************

DOY START-END USER    ACTIVITY          PASS CALIB EQUIPM CAT CONF                F
-----------------------------------------------------------------------------------
166 0015-1000 PN-11   TKG PASS          3000 A -A  000063 1A1 A000
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
166 1000-1555 DSS-11 MAINTENANCE                -          2A1 A700
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
166 1000-1030 DSS-11 PM 02.00-1-0.01    NIB     -    JERSCH 2A1 A700
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
166 1030-1530 DSS-11 PM 29.00-1-0.01    NIB     -    JERSCH 2A1 A700
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
166 1555-1920 HEL-01 TKG PASS           2380 B - 3 200063 1A1 A000
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
166 1920-0025 PN-12  TKG PASS           1123 A - 3 800063 1A1 A000
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
167 0025-0955 PN-11  TKG PASS           3001 A -A  000063 1A1 A000
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
167 0955-1240 DSN    AGC WEEKLY                 -          1A3 A700
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
167 1240-1545 DSS-11 TRAINING                   -          2B1 A700
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
```

**Fig. 2. Deep Space Station schedule example**

**********

DSS-11 STATION UTILIZATION REPORT
FOR
WEEK 23, 1981

| | SOH | EUH | PER |
|---|---|---|---|
| I. DSN USER SUPPORT | | | |
| A. 1. SPACECRAFT TRACKING | | | |
| PIONEER-11 | 50.92 | 45.50 | 31.82% |
| PIONEER-12 | 27.17 | 22.92 | 16.98% |
| HELIOS-1 | 20.58 | 15.50 | 12.86% |
| VOYAGER-1 | 12.33 | 11.33 | 7.71% |
| 2. PROJECT RELATED SUPPORT | | | |
| 3. DSN PROJECT PREPARATION | | | |
| DSN | 10.08 | 10.08 | 6.30% |
| B. RADIO SCIENCE | | | |
| OSS | 3.75 | 3.00 | 2.34% |
| C. ADVANCED SYSTEMS | | | |
| D. SPECIAL | | | |
| SUB-TOTAL | 124.83 | 108.34 | 78.02% |
| II. FACILITY ACTIVITIES | | | |
| A. MAINTENANCE | | | |
| 1. PREVENTIVE | 31.33 | | 19.58% |
| 2. CORRECTIVE(DOWNTIME) | 0.00 | | 0.00% |
| 3. CORRECTIVE(NO DOWNTIME) | 0.00 | | 0.00% |
| B. PERSONNEL TRAINING | 3.83 | | 2.40% |
| C. DSN ENGINEERING | | | |
| 1. ENGINEERING SUPPORT | 0.00 | | 0.00% |
| 2. DEVELOPMENT OR TESTING | 0.00 | | 0.00% |
| 3. MINOR MODS | 0.00 | | 0.00% |
| III. OTHER ACTIVITIES | | | |
| A. MAJOR MODIFICATIONS | 0.00 | | 0.00% |
| B. HOST COUNTRY RADIO SCIENCE | 0.00 | | 0.00% |
| C. MISCELLANEOUS | 0.00 | | 0.00% |
| TOTAL HOURS | 160.00 | 108.34 | 100.00% |

**********

**Fig. 3. Deep Space Station utilization report example**

PREVENTIVE MAINTENANCE REPORT

DSS-11 WEEK 26
PM'S DUE FOR PERIOD 173/81 TO 179/81
AS OF 06/12/81

| PM NUMBER | HRS RQD | DUE | LAST DONE | LAST SCHED | OD | DESCRIPTION |
|---|---|---|---|---|---|---|
| 01.03-2-1.03 | 4.0 | 81162 | 81132 | 81132 | | SDA PHASE & GAIN ALIGNMENTS |
| 01.03-2-2.01 | 2.0 | 81175 | 81085 | 79136 | | POWER SUPPLY CHECKS |
| 02.00-1-0.01 | 0.5 | 81159 | 81152 | 81166 | | 10/20 KW TXR CHECKLIST |
| 02.00-2-2.01 | 4.0 | 81179 | 81089 | 81089 | | H.E. INSPECTION |
| 05.02-1-0.02 | 1.0 | 81161 | 81154 | 81168 | | HDDTR |
| 06.00-2-1.21 | 12.0 | 80347 | 80317 | 81128 | 6 | ANT HYDROMECHANICAL INSPECTION |
| 06.01-2-3.40 | 4.0 | 80156 | 79341 | 79341 | 2 | OUT OF BALANCE FORCE |
| 06.01-2-3.41 | 3.0 | 76181 | 76001 | | 10 | FOUNDATION LEVEL CHECK |
| 06.01-2-3.42 | 0.0 | 76181 | 76001 | | 10 | OPTICAL TRACKING ALIGNMENT CKS |
| 06.01-2-3.43 | 0.0 | 76181 | 76001 | | 10 | ACQUISITION AID ALIGNMENT CKS |
| 06.01-2-4.53 | 0.0 | 77001 | 76001 | | 4 | ORTHOGONALITY CHECK |
| 06.01-2-4.54 | 24.0 | 78277 | 77277 | | 2 | POLAR SHAFT ALIGNMENT |
| 06.01-2-5.60 | 72.0 | 78001 | 76001 | | 1 | DIS CURVATURE OPTICAL CKS |
| 06.02-2-1.18 | 1.0 | 81049 | 81019 | 81019 | 4 | LEVEL VIALS CK |
| 06.02-2-4.34 | 02.0 | 80001 | 79001 | | 1 | LUBE PRESSURE SWITCH CHECKS |
| 06.03-2-0.01 | 01.0 | 81147 | 81140 | 81140 | | SERVICE BULL AND PINION GEARS |
| 06.03-2-0.02 | 00.5 | 81144 | 81137 | 81137 | 1 | SERVICE MAIN DRIVE BRAKES |
| 06.03-2-0.03 | 02.0 | 81159 | 81152 | 81152 | | SERVICE LOW SPEED CLUTCH |
| 06.03-2-1.05 | 02.0 | 81136 | 81106 | 81106 | 1 | MAIN DRIVE GEAR REDUCER CK |
| 06.03-2-1.06 | 01.0 | 81154 | 81124 | 81124 | | LOW SPEED GEAR REDUCER CK |
| 06.03-2-1.07 | 00.5 | 81150 | 81120 | 81120 | | TACH DRIVE BELT CK |
| 06.03-2-1.08 | 00.5 | 81171 | 81141 | 81141 | | LIMIT SW DRIVE CHAIN CK |
| 06.03-2-1.09 | 06.0 | 80255 | 80225 | 80225 | 9 | LIMIT SWITCH CHECK |
| 06.03-2-1.10 | 01.0 | 81170 | 81140 | 81140 | | GEAR REDUCER BOX CK |
| 06.03-2-1.11 | 01.5 | 81103 | 81073 | 81073 | 2 | BULL-PINION GEAR BACKLASH (EAST) |
| 06.03-2-1.12 | 01.5 | 81103 | 81073 | 81073 | 2 | BULL-PINION GEAR BACKLASH(WEST) |
| 06.03-2-1.13 | 01.0 | 81103 | 81073 | 81073 | 2 | BULL-PINION GEAR BACKLASH(NORTH) |
| 06.03-2-1.14 | 01.0 | 81103 | 81073 | 81073 | 2 | BULL-PINION BEAR BACKLASH(SOUTH) |
| 06.03-2-1.17 | 01.0 | 80321 | 80291 | 80291 | 7 | DECLINATION SHIFT |
| 06.03-2-3.32 | 04.5 | 79289 | 79109 | 79109 | 3 | LOW SPEED CLUTCH CHECK |
| 06.03-2-3.33 | 01.0 | 81154 | 80339 | 81163 | | LUBE PINION SHAFT CARRIER BEARINGS |
| 06.03-2-3.34 | 01.0 | 81110 | 80295 | 81163 | | LUBE DEC SHAFT BEARINGS |
| 06.03-2-3.39 | 01.0 | 81007 | 80192 | 80192 | | LOW SPEED REDUCER DRAIN |
| 06.03-2-4.51D | 24.0 | 80067 | 79067 | 79067 | 1 | MAIN DRIVE BRAKE MAINT |
| 06.03-2-4.51H | 24.0 | 80067 | 79067 | 79067 | 1 | MAIN DRIVE BRAKE MAINT |
| 06.03-2-4.55 | 10.0 | 81053 | 80053 | 80053 | | PINION CARRIER BEARINGS |
| 06.03-2-6.70D | 10.0 | 79001 | 76001 | | | CLEAN MAIN DRIVE GEAR REDUCER |
| 06.05-2-1.19 | 03.0 | 81159 | 81129 | 81162 | | HYPERBOLA DRIVE MECHANISM |
| 29.00-1-0.01 | 06.0 | 81159 | 81152 | 81166 | | SUBSYSTEM HOUSEKEEPING |
| 29.00-1-1.15 | 4.0 | 81152 | 81122 | 81162 | | MANUAL MODE TESTS |
| 29.00-1-1.19 | 4.0 | 81165 | 81135 | 81163 | | AUTOMATIC TEST |
| 29.00-1-1.20 | 1.0 | 81165 | 81135 | 81163 | | RELIABILITY TEST |
| 35.04-1-0.04 | 1.0 | 81161 | 81154 | 81168 | | LINE PRINTER |
| 37.02-1-1.08 | 00.5 | 81171 | 81141 | 81171 | | S-BAND MASER GAIN AND B/W MEASUREMENTS |
| 37.02-2-1.01 | 3.0 | 81150 | 81120 | 81120 | | TWM/CCR MONTHLY INSPECTION |
| 91.06-1-0.02 | 00.3 | 81162 | 81155 | 81169 | | OPERATOR MAINTENANCE |

* NBR IN OD COLUMN - OVERDUE PERIODS SINCE DUE DATE...

END OF REPORT

Fig. 4. Preventive maintenance report example

WEEKLY MANPOWER UTILIZATION REPORT

GDSCC MAINTENANCE AND INTEGRATION UNIT

WEEK ENDING 10 MAY 81

NUMBER OF PERSONNEL=53 TOTAL M&I MANHOURS THIS WEEK= 2112.0, YTD= 40,514.9

| | | M&I MANHOURS AT DSS-11 | | | M&I MANHOURS AT DSS-12 | | | M&I MANHOURS AT DSS-14 | | | OTHER M&I MANHOURS | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PROJ OR TASK | WORK CODE | MAN HOURS THIS WEEK | % TOTAL MAN HRS THIS WEEK | % TOTAL MAN HRS YEAR TO DATE | MAN HOURS THIS WEEK | % TOTAL MAN HRS THIS WEEK | % TOTAL MAN HRS YEAR TO DATE | MAN HOURS THIS WEEK | % TOTAL MAN HRS THIS WEEK | % TOTAL MAN HRS YEAR TO DATE | MAN HOURS THIS WEEK | % TOTAL MAN HRS THIS WEEK | % TOTAL MAN HRS YEAR TO DATE |
| PROJ REL SUPP | 1A2 | | | | | | < .1% | | | | | | |
| DSN PROJ PREP | 1A3 | | | < .1% | | | | | | | | | |
| PREVENT MAINT | 2A1 | 116.0 | 5.5% | 4.7% | 130.0 | 6.2% | 5.4% | 393.0 | 18.6% | 20.3% | | | .6% |
| CORR MAINT | 2A2 | | | .7% | | | .9% | 80.0 | 3.8% | 4.3% | | | |
| CORR MAINT | 2A3 | 128.0 | 6.1% | 5.1% | 190.5 | 9.0% | 9.0% | 380.0 | 18.0% | 18.0% | | | .8% |
| PERS TRAINING | 2B1 | 8.0 | .4% | .3% | 16.0 | .8% | .6% | 8.0 | .4% | .8% | 332.5 | 15.7% | 9.6% |
| ENG SUPPORT | 2C1 | | | .2% | 4.0 | .2% | 1.3% | 8.0 | .4% | .5% | | | |
| ENG SUPPORT | 2C2 | | | .2% | 8.0 | .4% | .3% | 16.0 | .8% | .5% | | | |
| MINOR MOD | 2C3 | 21.0 | 1.0% | .5% | 23.0 | 1.1% | 1.7% | 38.5 | 1.8% | 2.5% | | | .2% |
| MAJOR MOD | 3A1 | | | | | | .8% | | | .1% | | | |
| MISC | 3C1 | | | < .1% | | | .1% | | | < .1% | | | .5% |
| ADMIN | | | | | | | | | | | 46.0 | 2.2% | 1.9% |
| ABSENCE | | | | | | | | | | | 148.0 | 7.0% | 5.5% |
| FIRST AID TNG | | | | | | | | | | | | | 1.3% |
| EYE EXAMS | | | | | | | | | | | | | .1% |
| OTHER PROJ | | | | | | | | | | | 11.5 | .5% | .2% |
| SAFETY | | | | | | | | | | | | | .3% |
| FORD LPO TNG | | | | | | | | | | | 6.0 | .3% | .1% |
| DSN MASTR CLK | | | | | | | | | | | | | |
| M&I TOTALS | | 273.0 | 12.9% | 11.8% | 371.5 | 17.6% | 20.0% | 923.5 | 43.7% | 47.1% | 544.0 | 25.8% | 21.0% |

M&I MANHOURS AT DSS-11 | M&I MANHOURS AT DSS-12 | M&I MANHOURS AT DSS-14 | OTHER M&I MANHOURS

**Fig. 5. Manpower utilization report example**

BUDGET REPORT FOR ACCOUNT 38126

FISCAL YEAR 81 THRU APR. 81
22 MAY 81

| COST CENTER | | | FY | DISBURSEMENTS | | | | | ACTUAL | OBLIG | PROJ | COMM | EST | EXPENDED | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CODE | TITLE | MONITOR | BUDGET | SIS | JPL | MATL | LABOR | ODC | BALANCE | $ | BALANCE | $ | BALANCE | TIME | AMT |
| 405B | IPC UTILIZATION(DPU) | EVERETT | 21561 | 0 | 13299 | 0 | 0 | 0 | 8262 | 0 | 8262 | 0 | 8262 | 58% | 62% |
| 405C | IPC UTILIZATION(MR) | TODD | 1310 | 0 | 359 | 0 | 0 | 0 | 951 | 0 | 951 | 0 | 951 | 58% | 27% |
| 405D | SMART TERMINAL MAINT | EVERETT | 2000 | 0 | 0 | 0 | 0 | 0 | 2000 | 0 | 2000 | 0 | 2000 | 58% | 0% |
| 500A | WATER TESTING | ALDERSON | 1273 | 0 | 0 | 544 | 0 | 0 | 729 | 0 | 729 | 0 | 729 | 58% | 43% |
| 500B | GDSCC TELEPHONE | POTTER | 11249 | 0 | 7485 | 0 | 0 | 0 | 3764 | 0 | 3764 | 0 | 3764 | 58% | 67% |
| 500C | GDSCC WATER SUPPLY | POTTER | 14447 | 0 | 10755 | 0 | 0 | 0 | 3692 | 0 | 3692 | 0 | 3692 | 58% | 74% |
| 500D | TELEPHONE - BARSTOW | MARTIN | 5599 | 0 | 1327 | 0 | 0 | 0 | 4272 | 0 | 4272 | 0 | 4272 | 58% | 24% |
| 500F | RODENT & PEST CONTROL | ALDERSON | 3342 | 0 | 0 | 1553 | 0 | 0 | 1789 | 0 | 1789 | 0 | 1789 | 58% | 46% |
| 500G | GDSCC GRNDS KEEPING | ALDERSON | 12365 | 0 | 0 | 7142 | 0 | 0 | 5223 | 0 | 5223 | 0 | 5223 | 58% | 58% |
| 500H | ELECTRICAL CONTRACT | ALDERSON | 2000 | 0 | 0 | 0 | 0 | 0 | 2000 | 0 | 2000 | 0 | 2000 | 58% | 0% |
| 500J | WELDING CONTRACT | ALDERSON | 2000 | 0 | 0 | 0 | 0 | 0 | 2000 | 0 | 2000 | 0 | 2000 | 58% | 0% |
| 922L | PLANT MAINT MISC | ALDERSON | 18172 | 0 | 13060 | 13943 | 0 | 6488 | -15319 | 0 | -15319 | 0 | -15319 | 58% | 184% |
| | ACCOUNT TOTALS | | 95318 | 0 | 46285 | 23182 | 0 | 6488 | 19363 | 0 | 19363 | 0 | 19363 | 58% | 80% |

Fig. 6. Budget report example

```
**********************************************************************************

                                    GDSCC
                        PROCUREMENT ACCOMPLISHMENT REPORT
                                     FOR
                                  JUNE, 1981
                                   06/15/81
==================================================================================

PART 1: SUMMARY OF ACTIVITIES
   A. FOR THE MONTH OF JUNE:

      PURCHASE REQUISITIONS RECEIVED:  18    TOTAL VALUE  $    6850.85
      PURCHASE ORDERS RELEASED      :  24    TOTAL VALUE  $  228408.07


   B. TOTALS:

       PURCHASE REQUISITIONS OUTSTANDING:    0    TOTAL VALUE $       0.00
       PURCHASE ORDERS NOT RECEIVED     :   36    TOTAL VALUE $   45512.09
       PURCHASE ORDERS NOT PAID         :  106    TOTAL VALUE $   26610.61
       PURCHASE ORDERS RELEASED THIS FY :  536    TOTAL VALUE $  733586.20

   C. 1. AVERAGE DAYS FROM TIME OF PR INITIATION TO PROCUREMENT RECEIPT:
         FOR ROUTINE ACTION :     8.5
         FOR PRIORITY ACTION:     5.2
         FOR SUBCONTRACTS   :     5.7


      2. AVERAGE DAYS FROM TIME PR IS RECEIVED BY PROCUREMENT TO PO RELEASE:

         FOR ROUTINE ACTION :     8.7
         FOR PRIORITY ACTION:     2.4
         FOR SUBCONTRACTS   :    20.3


      3. AVERAGE DAYS FROM TIME PO IS RELEASED UNTIL RECEIVED FROM VENDOR:

         FOR ROUTINE ACTION :    22.8
         FOR PRIORITY ACTION:    14.8
         FOR SUBCONTRACTS   :    14.7



----------------------------------------------------------------------------------
```

**Fig. 7. Procurement accomplishment report example**

GDSCC VEHICLE USE REPORT

PERIOD FROM 04/21/81 TO 05/20/81

PAGE 5 VANS

| VEHICLE CL TAG | ODOMETER BEGIN | ODOMETER END | TOTAL MILES | | GAL-OF-GASOLINE GSAPUMP | GAL-OF-GASOLINE CR-CARD | AVG MPG | NOTES | REMARKS |
|---|---|---|---|---|---|---|---|---|---|
| G42-30946 | 078402 | 079391 | 01489 | * | 116.6 | | 12.8 | | TRM 05/15/81 |
| G42-30947 | 072825 | 073282 | 00457 | * | 35.0 | | 13.1 | | TRM 05/12/81 |
| G42-30949 | 080966 | 082694 | 01728 | * | 100.1 | | 17.3 | | TRM 05/14/81 |
| G42-32376 | 064293 | 065216 | 00923 | * | 72.5 | | 12.7 | | |
| G42-32377 | 086141 | 087234 | 01093 | * | 108.7 | | 10.1 | | TRM 05/18/81 |
| G42-32378 | 089040 | 089958 | 00918 | * | 91.5 | | 10.0 | | TRM 05/20/81 |
| G42-32380 | 069742 | 071210 | 01468 | * | 142.5 | | 10.3 | | |
| G42-33277 | 054003 | 057327 | 03324 | * | 258.7 | | 12.8 | | |
| G42-33288 | 069947 | 071420 | 01473 | * | 117.5 | | 12.5 | | |
| G42-33290 | 056096 | 058509 | 02413 | * | 166.6 | | 14.5 | | |
| G42-33291 | 040916 | 042204 | 01288 | * | 112.6 | | 11.4 | | |
| G42-33292 | 050654 | 053011 | 02357 | * | 147.4 | | 16.0 | | |
| G42-3329[illegible] | 01[illegible] | 048392 | [illegible]627 | * | [illegible] | | 11.9 | | |
| G4[illegible]28150 | 011726 | 014066 | 02340 | * | 162.9 | | 14.4 | | |
| G43-29532 | 028646 | 029989 | 01343 | * | 106.3 | | 12.6 | | |
| SUBTOTAL | | | 45,935 | | 3,590.6 | | 12.8 | | |
| TOTAL | 47 VANS | | 87,943 | | 6,733.3 | | | | |
| GRAND-TOT | 112 VEHICLES | | 200,786 | | 13,608.8 | 51.6 | | | |

NOTES

NOTE(1) ODOMETER READING OF TOTAL MILES IS NEGATIVE
NOTE(2) TOTAL MILES EXCEEDS 4000
NOTE(3) THE END ODOMETER READING IS ZEROES
NOTE(4) THE AVG-MPG EXCEEDS 30.0
NOTE(5) NO GSA PUMP TRANSACTIONS THIS PERIOD

END OF REPORT

**Fig. 8. Vehicle use report example**

N82-11307



# Special Activity Utilization of GDSCC Antennas During 1980

E. B. Jackson
Goldstone Operations Section

J. W. Hudson
Bendix Field Engineering Corp.

*In addition to direct spacecraft project support, the GDSCC antennas also support "special" activities. These activities can be categorized as Advanced System Program, Crustal Dynamics Program, and Radio Astronomy Program activities.*

*This article briefly discusses the goals, and categorizes each of these types of activities that received tracking support at Goldstone. All Goldstone stations (DSSs 11, 12, 13, and 14) provided a total of 2814.5 tracking hours to "special" activities during the period January through December 1980.*

## I. Introduction

The antennas at the Goldstone Deep Space Communications Complex (GDSCC), in addition to providing telemetry, ranging, and commanding support to spacecraft projects, also support a number of other tasks that require tracking. These special activities are categorized into Advanced Systems Program, Crustal Dynamics Program, and Radio Astronomy Program. This article discusses these special activities that GDSCC antennas supported with tracking during the period January through December 1980.

## II. Advanced Systems Program

Activities in this category are funded by NASA through the Office of Space Tracking and Data Systems (OSTDS). This is the development work that culminates in provision, to the Deep Space Network (DSN) stations, of new capability with which each spacecraft project's needs can be met. These activities are organized into Research and Technology Operating Plans (RTOPs).

### A. RTOP 60: Radio Metric Analysis, Demonstration, and Instrumentation Development

The goal of this effort is the development and demonstration of advanced radio metric systems for navigation and radio science, with a specific angular accuracy goal of 50 nanoradians for delta VLBI measurements.

In general, using VLBI, this RTOP generates a catalog of suitable (strength and location) quasar calibration sources. Using this catalog of quasars, the ephemerides of the planets,

particularly Mars, is tied into the "quasar reference frame." Then, techniques are developed for differential measurements between spacecraft and quasars that locate the spacecraft within the quasar reference frame and hence the spacecraft with reference to the target planet(s). Table 1 lists the various activities within the RTOP that received tracking support from GDSCC during 1980.

### B. RTOP 61: VLBI Development and Analysis

This activity is closely associated with, and builds upon the results of, RTOP 60. Within this group of activities, experiments are conducted that demonstrate the capabilities of the Block I VLBI technology, which is planned to provide 50-nanoradian angular accuracy, 10-nanosecond clock synchronization between widely separated sites, and measurement of frequency standard stability with accuracies of parts in $10^{13}$, also between widely separated locations.

The activities within this RTOP that received tracking support at GDSCC are also listed in Table 1.

### C. RTOP 65: Antenna Systems Development

In recognition of the necessity of increasing antenna performance to provide increased planetary communications capability, this activity's goal is to improve antenna technology. Specifically, this RTOP looks at possible improvements in electronic – microwave capabilities, and mechanical – structural design. Application of technology to provide increased frequency operation of existing antennas, with improved techniques of antenna pointing, is also a part of this activity.

As part of the analysis performed in this area, techniques for measurement of antenna gain were refined and used to make extensive measurements at DSS 12 and DSS 14. These antenna gain measurements were used to quantify the improvements resulting from the modifications to the microwave system at DSS 14 and the increase in aperture at DSS 12.

The particular activity supported during 1980 was "Analytical Techniques and Procedures," which has as its objective the development of analysis software for application to antenna structural and mechanical performance. This activity also aims to extend the capability for automatic antenna structural design optimization and organize existing software to simplify maintenance and execution of improvements. Again, Table 1 illustrates tracking during 1980.

### D. RTOP 68: Station Monitor and Control Technology

Recognizing that a substantial portion of the operating cost for DSN stations is allocated to personnel, this RTOP seeks to develop automation technology that will allow entire station operation from a remote point and prove the reliability, safety, and efficiency of unattended operation. Additionally, this RTOP plans to develop the data base with which reliability, costs, and productivity can be monitored, using DSS 13 as a demonstration unattended station.

Table 1 lists only tracking support to this activity; substantial other work was also ongoing at DSS 13 during the year as the equipment used to make DSS 13 operate in the unattended mode evolved and additional systems were added to those controlled remotely.

### E. RTOP 70: High-Speed Signal Processing Research

The objective of this RTOP is to develop high-speed digital signal processing techniques for use in the DSN. A major part of this task is the development of a test bed as a research tool used to explore high-speed techniques. The test bed is to be used for demonstrations of various development capabilities, such as high-performance array processors, and wideband correlation subsystems. This work also demonstrates commercial LSI signal processing building blocks and utilizes promising commercial design systems.

The particular activity supported at GDSCC during 1980 was the Radar Data Acquisition, which uses various planets and asteroids as targets for demonstration of signal processing techniques.

In Table 1 are listed the various targets which were explored during 1980 as the demonstration test bed was used to develop signal processing techniques.

## III. Radio Astronomy Program

Activities in this category are either sponsored by the Office of Space Science (OSS) or authorized by the Radio Astronomy Experiment Selection (RAES) Panel.

Experimenters within NASA who desire OSS support of their activities submit an observing plan in the form of an RTOP. If approved, OSS requests OSTDS to provide support, and time is made available on suitable facilities of the Deep Space Network (DSN). Experimenters from universities who have observing plans that require NASA facilities submit their observing plan to the RAES Panel. If approved, time is made available through an agreement whereby NASA makes avail-

able a percentage of the operational time on DSN facilities. Table 2 depicts the support at GDSCC in this area during 1980.

## A. OSS Sponsored Activities

1. **Pulsar rotation constancy**. This experiment seeks to monitor short-term variations in the period of the relatively young VELA pulsar (PSR 0833-45). Additionally, by observing twenty-three older, more stable pulsars, this experiment hopes to obtain data that will provide further tests of the hypothesis that pulsars are neutron stars resulting from supernova explosions, which impart high velocities to the resulting pulsar.

2. **Interstellar microwave spectroscopy**. This experiment conducts fundamental research into questions of stellar and galactic evolution by observations of atomic and molecular spectral lines and related radio-continuum observations. Among the questions that these data may help to answer are (1) the distribution of interstellar gas/dust clouds, (2) the chemistry of dense interstellar clouds, and (3) the interactions of young stars with the surrounding clouds.

3. **ALSEP-quasar VLBI**. By means of astrometric observation of moon-based ALSEP transmitters and quasar radio sources using differential VLBI, the lunar orbit can be ascertained with reference to the inertial quasar coordinate frame. Such referencing of the lunar orbit to the quasar frame also has application to relativity and cosmological theories, in addition to its obvious space program application.

4. **Planetary radio astronomy**. This experiment has the dual objectives of studying the properties of the planet Jupiter's radio emission and measurement of the thermal emission from the atmosphere of the outer planets. Both of these observations will aid construction of improved models of Jupiter's radiation belt environment as well as atmospheric models of the outer planets, particularly Uranus.

## B. RAES Panel Sponsored Activities

1. **Quasar patrol (RA 137)**. By use of very long baseline interferometry (VLBI) techniques, this experiment systematically monitors time variation in the small-scale structure of quasars and radio galaxies, particularly 3C273, 3C279, and 3C 120. Past observations have shown changes in small-scale structure taking place in a period of a few weeks. These changes, in the case of the three objects mentioned above, seem to require highly relativistic velocities to explain the observed separation or expansion of the small components. In the particular case of 3C273, the observed expansion proceeds at apparent velocities in excess of the speed of light.

2. **Compact nuclei VLBI (RA 169)**. Using only United States based baselines, VLBI observations of many quasars, radio galaxies, spiral and elliptical galaxies have detected compact radio nuclei at the center. These nuclei typically have measured angular dimensions of less than 0.001 arc seconds and flux densities of approximately 0.1 Jansky. This experiment utilizes two 64-m antennas with an Australian – United States baseline for higher resolving capability. The experiment also uses both S- and X-band for greater sensitivity along with greater resolving power.

3. **M87 interstellar scintillation VLBI (RA 171)**. A dynamical study of the velocity dispersion of stars in and around the nucleus of M87 provides evidence that an extremely massive dark object, or group of objects, exists in the nucleus. Detection of interstellar scintillation from this nucleus would provide evidence of a possible black hole. These VLBI observations are made at S-band and utilize DSS 14 at GDSCC and DSS 63 in Madrid, Spain.

4. **VLBI investigation of SS-433 (RA 175)**. SS-433 has a bizarre optical spectrum that exhibits three sets of emission lines. One of these three emission lines shows near-zero radial velocity, while the other two show large and variable shifts to the blue and the red. The variability is periodic, with a period of about 164 days.

This experiment conducts regular VLBI observations of this source, using many simultaneous baselines. The multiple baselines should enable determination of the source angular size and structure, which will aid modelling. The involved stations at GDSCC include DSS 13 and DSS 14, along with a number of other radio observatories in the United States and Europe.

5. **VLBI investigation of "twin" quasi-stellar objects (QSOs) 0957 + 561A, B (RA 176)**. These QSOs, designated 0957 + 561A, B, are separated by only 6 arc seconds in angle and have equal optical redshifts and remarkable similiarity in their optical spectra. This similiarity is inconsistent with chance alone, so the objects are in some manner physically associated. One theory proposes that there is only one object, whose radiation is gravitationally bent about an intervening massive lens, a "gravitational lens."

This experiment uses VLBI observations with many simultaneous baselines to provide data with which detailed structure maps can be prepared. Repeated observations could monitor

changes in the radio brightness distribution. At GDSCC, the involved antenna is DSS 14.

## IV. Crustal Dynamics Program

A Crustal Dynamics Program goal is the demonstration of the capability of VLBI Systems to make highly accurate geodetic measurements. The particular activity supported at GDSCC during 1980 is Astronomical Radio Interferometric Earth Surveying (ARIES) data collection. The thrust of the ARIES activity is to demonstrate the geodetic performance of highly mobile VLBI stations while providing accurate data of significant geophysical interest. By moving around Southern California, and making VLBI measurements with base stations located at Goldstone and other fixed points, the ARIES activity is able to accumulate data on regional deformation and strain accumulation, particularly as associated with the San Andreas Fault.

These measurements are usually scheduled for periods in excess of 24 hours at a time, and a number of suitable radio sources are observed simultaneously and repeatedly by all the stations involved. During the period January 1 through December 31, 1980 DSS 13 (Venus Station) provided 367.75 hours of tracking support to these measurements of crustal deformation along the San Andreas Fault.

**Table 1. Advanced systems program utilization of GDSCC antennas during 1980**

| Activity | Station and tracking hours provided | | | |
|---|---|---|---|---|
| | DSS 11 | DSS 12 | DSS 13 | DSS 14 |
| RTOP 60: radio metric analysis, demonstration and instrumentation development | | | | |
| Clock synchronization VLBI | | | | 57.25 |
| Ecliptic source catalog development | | | 57.25 | 14.00 |
| Mark IV Differential One-Way Range (DOR) Systems development | | | 34.25 | 0.50 |
| Mark IV Radio Metric Systems development | | | 77.50 | |
| Viking Delta VLBI | | | 40.75 | |
| Voyager Delta VLBI | | | | 1.00 |
| Voyager VLBI | | | | 169.50 |
| RTOP 61: VLBI development and analysis | | | | |
| VLBI systems design and analysis | | | 54.25 | 7.00 |
| Microwave phase center calibration | | | 19.25 | 28.00 |
| Radio source catalog development | | | 22.25 | |
| RTOP 65: antenna system development | | | | |
| Analytic techniques and procedures | | 96.50[a] | 19.25 | 129.00[a] |
| RTOP 68: station monitor and control technology | | | | |
| DSS 13 S-X unattended systems development | | | 62.25 | |
| RTOP 70: high-speed signal processing research | | | | |
| Radar data acquisition | | | | |
| Asteroid Apollo | | | | 12.75 |
| Asteriod Toro | | | | 18.25 |
| Comet Bradfield | | | | 2.25 |
| Planet Mars | | | | 126.00 |
| Planet Mercury | | | 16.25 | 17.00 |
| Planet Venus | 93.75 | | 73.50 | 224.00 |
| TOTALS | 93.75 | 96.50 | 476.75 | 806.5 |

[a]Antenna gain measurements.

**Table 2. Radio astronomy utilization of GDSCC antennas during 1980**

| Activity | Station and tracking hours provided | | | |
|---|---|---|---|---|
| | DSS 11 | DSS 12 | DSS 13 | DSS 14 |
| OSS sponsored activities | | | | |
| Pulsar rotation constancy | 87.45 | 54.75 | 188.25 | 80.75 |
| Interstellar microwave spectroscopy | | | 19.25 | 39.50 |
| ALSEP-quasar VLBI | | | 53.25 | |
| Planetary radio astronomy | | 13.75 | 162.00 | |
| RAES panel sponsored activities | | | | |
| Quasar patrol – RA 137 | | | | 15.75 |
| Compact nuclei VLBI – RA 169 | | | | 11.50 |
| M87 interstellar scintillation VLBI – RA 171 | | | | 15.75 |
| VLBI investigation of SS-433 – RA 175 | | | 224.00 | |
| VLBI investigation of "twin" quasi-stellar objects (QSOs) 0957 + 561A, B – RA 176 | | | | 10.00 |
| TOTALS: | 84.75 | 68.50 | 646.75 | 173.25 |

N82-11308



# Goldstone High Power Radiation Control

B. A. Gaudian
Goldstone Operations Section

R. B. Cushman
Bendix Field Engineering Corporation

*The airspace near Goldstone is shared by many users in carrying out their respective missions. Coordination of Goldstone High Power Transmitter Radiation activity with mutual airspace users near Goldstone is an effective method to prevent aircraft from being exposed to radiation levels greater than 10 mW/cm$^2$, to allow Goldstone to transmit high power levels for Planetary Radar experiments and to provide the necessary information to both civil and military aircraft users to enable them to perform their necessary activities.*

## I. Introduction

The Goldstone Deep Space Communications Complex (GDSCC), the Department of Defense (DOD), and the Federal Aviation Administration (FAA) use the airspace near Goldstone to carry out their respective missions. Whenever Goldstone activates high power emitters the airspace near the antenna beam could be hazardous to air operations. To prevent aircraft from being exposed to levels of radiation greater than 10 mW/cm$^2$ Goldstone has two choices: operate in conjunction with a local surveillance radar to detect aircraft within the proximity of the transmitted beam or operate within a set of parameters mutually acceptable to airspace users. Goldstone has elected the latter of the two choices. The intent of this article is to provide information regarding the internal high power radiation policies, high power radiation constraints, agreements with external agencies, and the actions required to fulfill high power radiation agreements.

## II. Internal High Power Radiation Policies

The following policies are adhered to by all high power transmitter users. Any changes to these policies must be approved by the Goldstone Radio Spectrum Coordinator and the JPL Operations Support Group (OSG).

### A. Long Range Scheduling Information

The Mission Coordination Group shall provide an annual prediction of radio astronomy and special activities to the OSG and the Goldstone Radio Spectrum Coordinator for planning and coordination of high power emissions. These predictions should include any object which may be desired to be tracked, and the date, time, power and band of emission.

### B. Midrange Scheduling Information

The OSG shall enter a 12-week advance high power transmitter projection into the Univac 1100/81F data base on a weekly basis. This projection will include the year, week, day of year, calibration information, power of emission, and the source that will be tracked. The Goldstone Radio Spectrum Coordinator will query the data base (Fig. 1) each week to update the Goldstone schedule (Fig. 2).

### C. Published 7-Day Schedule

The OSG shall notify the Goldstone Radio Spectrum Coordinator of any scheduled radiation greater than 20 kW

originating from the GDSCC. The OSG shall transmit, via teletype (Fig. 3) by Tuesday, the high power radiation schedule for the following week. This message will include the project, Deep Space Station (DSS) number, power and band of emission, the day of year, track times in 30-minute intervals, azimuth, elevation, declination, and the zenith crossing time.

### D. Real-Time High Power Activation (Spacecraft Emergency)

Any high power activity (100 kW or greater) not appearing in the published 7-day schedule will be handled in real-time. The Goldstone Radio Spectrum Coordinator will take appropriate action during normal working hours. Outside normal working hours the OSG is responsible for informing the Federal Aviation Administration (Los Angeles Center) and the Edwards AFB Radar Approach Control (RAPCON) of any emergency high power radiation. After proper notification the JPL Network Operations Chief (NOC) may authorize the station to radiate at high power. This is not a standard mode of operation – it applies to spacecraft emergency only.

### E. Real-Time High Power Termination

In the event of an air emergency or personnel emergency in or adjacent to Goldstone the DSN may be requested to terminate any high power radiation (100 kW or greater). During normal working hours the Goldstone Radio Spectrum Coordinator will make the decision to terminate any high power emission. After normal working hours the OSG will make this decision according to the inputs received. These inputs may come from the DSN stations, RAPCON, CCF, FAA, or local land and airspace users.

## III. High Power Radiation Constraints

The majority of the DSN high power targets are within a specific tracking window. The following constraints were approved by the DSN, DOD, and FAA after examining our normal tracking targets. These constraints apply to all high power emissions of 100 kW or greater.

### A. Elevation Restriction

There will be no high power emission below 15 degrees elevation.

### B. Declination Restrictions

All high power emissions shall fall between ±25 degrees declination.

### C. Minimum Time Notifications

There will be a minimum of one week notification on high power emissions, except for maintenance at zenith. If high power transmitter usage exceeds 2 days of any week a minimum of 3 weeks notification is required.

### D. Exceptions

Any exceptions to the restrictions in pointing angles or notification will be handled by the Goldstone Radio Spectrum Coordinator on a case-by-case basis. Additional external coordination may be required to determine if the activity will impact local airspace users.

## IV. Agreements With External Agencies

Goldstone has agreed to minimize airspace usage by managing the GDSCC high power radiation program. GDSCC provides sector information to the Central Coordinating Facility (CCF) at Edwards AFB for dissemination to other agencies.

### A. Annual and Semiannual Projection of High Power

GDSCC and the R2508 Complex Control Board members review the long range schedule to ensure compatibility with their scheduled activities. To date, this action has not caused any cancellations of DSN high power transmissions, and it has created a spirit of cooperation.

### B. Advance Notification of High Power Emissions

GDSCC must provide a minimum of three working days notification to the CCF and Fort Irwin. Subsequent to this, CCF will notify FAA (Los Angeles Center) when the restricted areas are closed or when the Goldstone VFR Avoidance Zone is activated. CCF will also notify all military airspace users, with the exception of Fort Irwin, of the avoidance sectors. Goldstone will notify the Fort Irwin Frequency Manager of Goldstone High Power Activation. The NTC Frequency Manager will notify NTC air operations for the safety of NTC small, fixed-wing aircraft and helicopters.

### C. CCF Notification Responsibility

The CCF is responsible for notifying all users of the R-2508 including the FAA Los Angeles Center of all Goldstone high power radiation. Normally CCF will provide specific avoidance sectors to the military sponsors and will retain control of restricted areas to simplify general aviation avoidance. A 12-nautical-mile VFR avoidance zone centered at DSS 14 will be in effect whenever Goldstone is radiating high power.

## V. Specific Actions Required to Fulfill High Power Agreements

The Goldstone Radio Spectrum Office is responsible for policing the Goldstone High Power Radiation Control program and taking the necessary action to fulfill the written or verbal agreements with external agencies. To assist in this activity a computer program has been developed for the 9825A calculator which inputs the date, day of year, planet/mission, declination, emission power, band, start time of track, stop time of track, and the zenith crossing time. Figure 1 shows a sample computer printout depicting the input data, the NASA Goldstone High Power Avoidance Zones, the action required by the Radio Spectrum Office, the station authorization, and the required distribution. In the event the transmitter on/off times are not above the 15-degree elevation point a caution would be printed to alert the Goldstone Operations Supervisor.

### A. External Notification

After providing a Goldstone high power printout, the Avoidance Zone information is transferred to a map (Fig. 4) which depicts the transmitted beam avoidance zones and the time the beam remains in each sector. This information is then passed to the Central Coordinating Facility at Edwards AFB, and to the Spectrum Management Office at Fort Irwin.

### B. Internal Notification

After external notification is accomplished it is necessary to issue a signed authorization to transmit. This authorization is specific as to power and band of transmission, the JPL scheduling message authorizing the transmission, restrictions or exceptions, the Spectrum Coordinator's signature, and the date authorized (Fig. 5).

## VI. Summary

From the implementation of this High Power Radiation Control program in January 1980, until August 1981, there have been a total of 78 high power transmitter operations. All of these were conducted successfully by advance coordination with the military and the FAA.

The Goldstone High Power Radiation Control program has been operating smoothly and has proved that coordination is preferable to an aircraft tracking radar system. The military organizations have upheld their end of the agreement, and Goldstone has policed internal policies to prevent any conflicts between organizations. As long as this spirit of cooperation continues the Goldstone High Power Radiation Control program will be a success.

HIGH POWER TRANSMISSION SCHEDULE
FOR WEEKS 23 THRU 42
06/03/81

| YEAR | WEEK | DOY | TIME | CAL | POWER | | SOURCE |
|---|---|---|---|---|---|---|---|
| 81 | 23 | 158 | 1335-0145 | D A | X-BAND 400KW | | MERCURY |
| 81 | 24 | 161 | 1515-0245 | D A | X-BAND 400KW | | MERCURY |
| 81 | 24 | 165 | 1245-1815 | D A | S-BAND 400KW | (BISTATIC) | MERCURY |
| 81 | 25 | 169 | 1230-2210 | C A | X-BAND 400KW | | MERCURY |
| 81 | 25 | 171 | 1220-0055 | D A | S-BAND 400KW | (BISTATIC) | MERCURY |
| 81 | 26 | 174 | 1710-0145 | D A | X-BAND 400KW | | MERCURY |
| 81 | 27 | 182 | 1740-0110 | D A | S-BAND 400KW | (BISTATIC) | MERCURY |
| 81 | 28 | 189 | 1100-1730 | D A | X-BAND 400KW | | MERCURY |
| 81 | 28 | 190 | 1640-0040 | D A | X-BAND 400KW | | MERCURY |
| 81 | 29 | 198 | 1645-0015 | D A | X-BAND 400KW | | MERCURY |
| 81 | 30 | 207 | 1545-2345 | D A | X-BAND 400KW | | MERCURY |
| 81 | 31 | NOTHING SCHEDULED | | | | | |
| 81 | 32 | NOTHING SCHEDULED | | | | | |
| 81 | 33 | NOTHING SCHEDULED | | | | | |
| 81 | 34 | NOTHING SCHEDULED | | | | | |
| 81 | 35 | NOTHING SCHEDULED | | | | | |
| 81 | 36 | NOTHING SCHEDULED | | | | | |
| 81 | 37 | NOTHING SCHEDULED | | | | | |
| 81 | 38 | NOTHING SCHEDULED | | | | | |
| 81 | 39 | NOTHING SCHEDULED | | | | | |
| 81 | 40 | NOTHING SCHEDULED | | | | | |

END OF REPORT

**Fig. 1. High power transmission schedule**

1981 DUAL, BI, and TRI-STATIC High Power Transmitter Projection

| DATE | DOY | TIME (Z) | BAND /PWR | MISSION /PLANET | ZENITH TIME (Z) | 15 DEGREE RISE | 15 DEGREE SET | DEC ANGLE |
|---|---|---|---|---|---|---|---|---|
| Mar 02 | 061 | 0800-1305 | X200 | SATURN | 0940 | 0500 | 1430 | 001 |
| Mar 07 | 066 | 0900-1330 | X200 | SATURN | 0920 | 0434 | 1408 | 001 |
| Mar 09 | 068 | 0630-1100 | X400 | SATURN | 0912 | 0426 | 1400 | 001 |
| Mar 14 | 073 | 0630-1100 | X200 | SATURN | 0851 | 0404 | 1339 | 001 |
| Mar 20 | 079 | 0755-1225 | X200 | SATURN | 0826 | 0339 | 1314 | 000 |
| Apr 07 | 097 | 0525-1245 | X200 | SATURN | 0710 | 0221 | 1200 | 000 |
| Apr 16 | 106 | 0205-1145 | X400 | Mon-St JUP | 0633 | 0150 | 1125 | 000 |
| Apr 27 | 117 | 0000-1030 | X400 | Mon-St JUP C | 0543 | 0056 | 1038 | 001 |
| Apr 28 | 119 | 0000-1030 | X400 | Mon-St JUP C | 0540 | 0048 | 1030 | 001 |
| Apr 30 | 120 | 2250-1030 | X400 | Mon-St JUP E | 0540 | 0043 | 1026 | 001 |
| May 17 | 137 | 1715-0300 | X400 | Mon-St MERC | 2115 | 1515 | 0310 | 025 |
| May 28 | 148 | 1615-0310 | X400 | Mon-St MERC | 2120 | 1530 | 0315 | 024 |
| Jun 07 | 158 | 1335-0145 | X400 | Mon-St MERC | 2100 | 1515 | 0205 | 023 |
| Jun 10 | 161 | 1515-0245 | X400 | Mon-St MERC | 2055 | 1510 | 0245 | 022 |
| Jun 14 | 165 | 1245-1815 | S400 | Bi-St MERC | 2027 | 1450 | 0215 | 021 |
| Jun 18 | 169 | 1220-2210 | X400 | Mon-St MERC | 2000 | 1420 | 0145 | 020 |
| Jun 20 | 171 | 1220-2250 | S400 | Bi-St MERC | 2000 | 1420 | 0145 | 020 |
| Jun 23 | 174 | 1710-0145 | X400 | Mon-St MERC | 1945 | 1400 | 0115 | 020 |
| Jul 01 | 182 | 1740-0110 | S400 | Bi-St MERC | 1900 | 1325 | 0040 | 019 |
| Jul 08 | 189 | 1100-1730 | S400 | Mon-St MERC | 1840 | 1300 | 0020 | 019 |
| Jul 09 | 190 | 1640-0040 | X400 | Mon-St MERC | 1815 | 1245 | 0010 | 020 |
| Jul 17 | 198 | 1645-0015 | X400 | Mon-St MERC | 1758 | 1218 | 2338 | 021 |
| Jul 26 | 207 | 1545-2345 | X400 | Mon-St MERC | 1720 | 1145 | 2300 | 022 |
| Sep 20 | 263 | 1700-0135 | X400 | Mon-St MERC | 2115 | 1700 | 0135 | 349 |
| Sep 30 | 273 | 1710-0100 | X400 | Mon-St MERC | 2115 | 1710 | 0100 | 345 |
| Oct 30 | 303 | 1400-2250 | X400 | Mon-St MERC | 1825 | 1400 | 2250 | 354 |

**Fig. 2. 1981 high power transmitter projections**

GJGO38A
RR JGLD JGTS JOCC
DE JJPL 042A
07/2055Z
FM DSN SCHEDULING
TO JGLD/GOLDSTONE RADIO SPECTRUM COORDINATOR
INFO JGLD/E JACKSON/GOLDSTONE RFI COORD R CUSHMAN
JGTS/STADIR/STAOPS
JOCC/NOC
DLD/D ENARI/G M REED/J R HALL/J NASH/R SATTERLEE/R DURHAM/N LACY/
M HALAL/L LAPRADE

SUBJECT: DEEP SPACE NETWORK RADIATION SCHEDULE FOR GOLDSTONE DURING
WEEK 29, DATES: 13 JULY - 19 JULY 81
PROJECT: MERCURY RADAR MAPPING DSS 14 RADIATING AT 400 KW,
IN X-BAND

| D.O.Y. | G.M.T. | AZ | EL | DEC |
|---|---|---|---|---|
| 198 | 1645 | 0 | 90 | |
| | 1745 | 0 | 90 | |
| | 1845 | 198.5 | 75.5 | 21.39 |
| | 1915 | 221.3 | 72.4 | |
| | 1945 | 236.8 | 67.8 | |
| | 2015 | 247.3 | 62.3 | |
| | 2045 | 255.0 | 56.5 | |
| | 2115 | 261.0 | 50.6 | |
| | 2145 | 266.1 | 44.5 | |
| | 2215 | 270.6 | 38.4 | |
| | 2245 | 274.7 | 32.2 | |
| | 2315 | 278.6 | 26.2 | |
| | 2345 | 282.5 | 20.2 | 21.43 |
| | 0015 | 0 | 90 | |

07/2102Z JUL 81 JJPL

**Fig. 3. Goldstone weekly radiation schedule**

| DATE | TIME (Z) | A SECTOR | | | | | B SECTOR | | | | | C SECTOR | D SECTOR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | W1W2* | NW | NE | SE | SW | NW | NE | SE | SW | E1E2* | | |
| | | | | | | | 000-400 | 000-400 | | | 200-400 | | |
| | | | | | | | 000-400 | 200-400 | | | | | |
| | | | | | | | 000-400 | | 300-400 | 400-500 | | | |
| 7/17/81 | 1845-2128 | | | 000-400 | 400-500 | 400-500 | 000-400 | | | | | | |
| 7/17/81 | 2128-2238 | | 200-400 | 000-400 | | | 000-200 | | | | | | |
| 7/17/81 | 2238-2338 | 200-400 | 000-400 | 000-200 | | | 000-200 | | | | | | |
| | | | | | | | | | | | | | |

*NOTE: W1W2 IS AN EXTENSION OF A SECTOR AND E1E2 IS AN EXTENSION OF B SECTOR

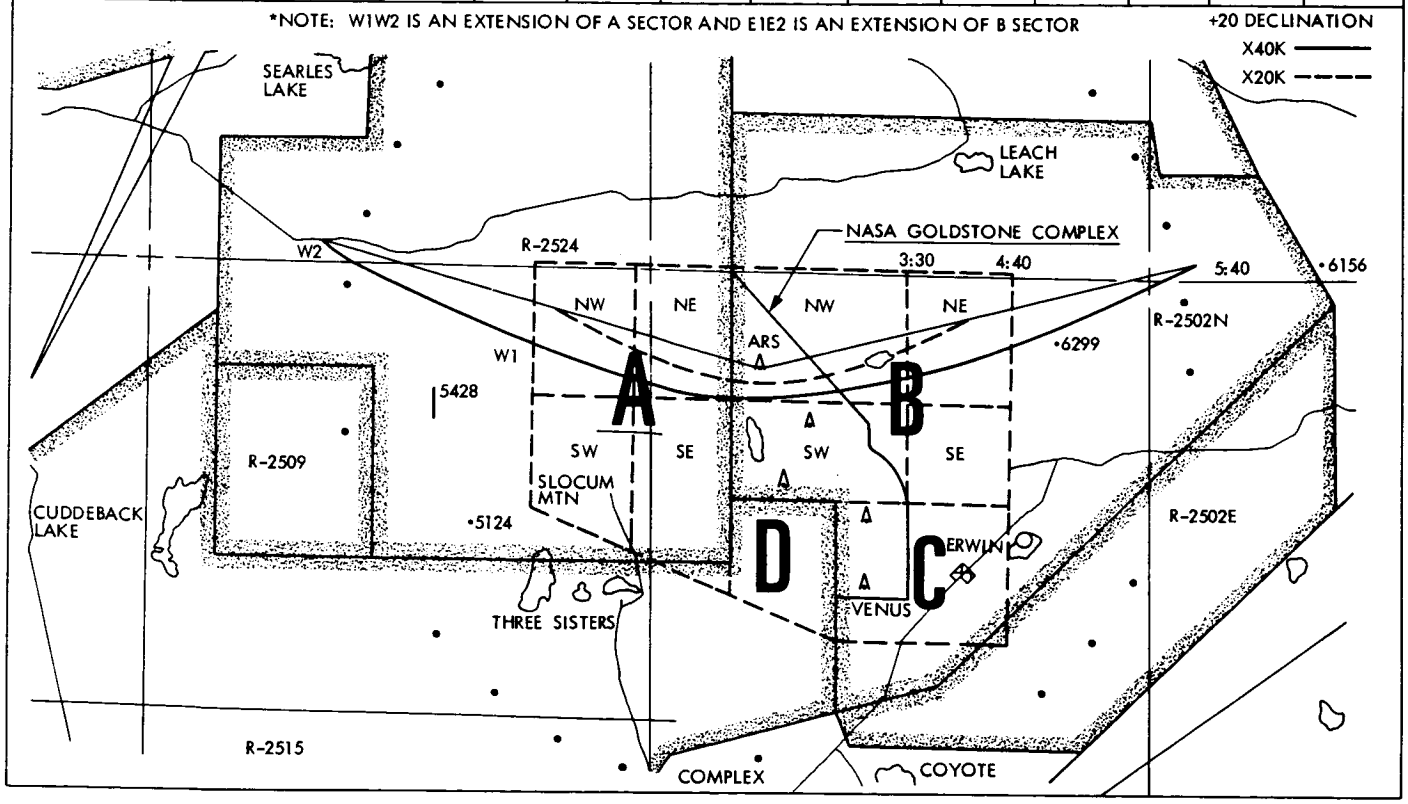



**Fig. 4. Goldstone high power avoidance zone**

GOLDSTONE HIGH POWER TRANSMISSION AUTHORIZATION

| Date | Day | Planet/ Mission | Dec. | Pwr/ Band | Start Time GMT | Stop Time GMT | Zenith Crossing |
|---|---|---|---|---|---|---|---|
| 7/17/81 | 198 | Mercury | 20 | 400X | 18:45 | 23:45 | 17:58 |

NASA GOLDSTONE HIGH POWER AVOIDANCE ZONES

| | | | A Sector | | | | | B Sector | | | | | C | D |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | DATE | | W1W2 | NW | NE | SE | SW | NW | NE | SE | SW | E1E2 | Sec | Sec |
| (4) | 7/17/81 | 18:45-21:28 | | | 0-4 | | | 0-4 | | | | | | |
| (5) | 7/17/81 | 21:28-22:38 | | 2-4 | 0-4 | | | 0-2 | | | | | | |
| (6) | 7/17/81 | 22:38-23:38 | 2-4 | 0-4 | 0-4 | | | 0-2 | | | | | | |

*** ACTION ***

(A) Notify CCF of High Power Avoidance Zones.
(B) Notify Ft. Irwin of High Power Avoidance Zones.
(C) Issue STATION AUTHORIZATION to transmit.

---

STATION AUTHORIZATION GRANTED FOR - 400KW X-BAND
PER JJPL 042A DTG 07/2055Z JUL 81

Signature Date

CAUTION: Transmitter on/off times are not within the 15 degree tracking window. The transmitter may be tested at zenith during the following times:
(b) 7/17/81 23:38-23:45

cc: R. Blanchard
L. Butcher
R. Kelley

**Fig. 5. Goldstone high power transmission authority**

N82-11309



# Planetary Radar

R. D. Shaffer and J. T. McCluskey
Control Center Operations Section

G. S. Downs
Communications Systems Research Section

*This article reports on the radar astronomy activities supported by the Deep Space Network during the third quarter of FY81. The planet Mercury was the principal object of interest.*

The 64-meter Goldstone station supported the observations of Mercury utilizing S- and X-band high-power transmitters (200 kW). The data have been edited, but have not been reduced to altitudes as yet. The estimate of the amount of useable data acquired vs the total obtained appears to be approximately 70%.

During the third quarter of FY81, the planet Mercury was observed on 10 different occasions during the latest inferior conjunction, primarily at 12.5-cm wavelength (S-band). This series of observations marks a return to Mercury after an absence of seven years and is the start of a long-term series of observations (1 to 10 years) to map the topography and scattering properties as completely as is physically possible. Currently, signal-to-noise considerations limit the coverage to between latitudes $\pm 11^{\circ}$, representing 18% of the surface.

Inferior conjunctions of Mercury occur about three times per year, allowing for considerable radar activity throughout the year. Due to considerable activity related to Venus and Mars, observations of Mercury will resume in April of 1982 with two observing sessions using two-element radar interferometer measurements designed to improve the precision with which the spin axis of Mercury is defined.

D31 N82-11310



# An Optimization Model for Energy Generation and Distribution in a Dynamic Facility

F. L. Lansing
DSN Engineering Section

*An analytical model is described using linear programming for the optimum generation and distribution of energy demands among competing energy resources and different economic criteria. The model, which will be used as a general engineering tool in the analysis of the Deep Space Network ground facility, considers several essential decisions for better design and operation. The decisions sought for the particular energy application include: the optimum time to build an assembly of elements, inclusion of a storage medium of some type, and the size or capacity of the elements that will minimize the total life-cycle cost over a given number of years. The model, which is structured in multiple time divisions, will employ the Decomposition Principle for large-size matrices, the Branch-and-Bound Method in mixed-integer programming, and the Revised Simplex Technique for efficient and economic computer use.*

## I. Introduction

The problem of allocating limited resources among competing activities in the "best" possible way has been always the prime concern to any organization. To make decisions after comparing the performance characteristics and life-cycle costs of existing versus new designs or between two new alternate designs is a common engineering practice. However, seeking better decisions with "optimum" designs rather than "working" designs is a superior engineering practice that has progressed only in the last few decades as a result of advances in computer technology. Solving large optimization problems with variables in the order of ten thousands is now feasible (Ref. 1).

In light of decreasing fossil energy resources, erosion of purchasing power due to inflation, rising operation, maintenance and utility costs, and the inevitable need to replace equipment because of obsolescence or wear, increased attention has been given to optimum solutions for energy systems. Optimum configurations of energy generation and distribution systems are sought from the supply (or material) end to the demand (or consumer) end with multiple competing processes and links in between.

Fortunately, all optimization problems are similar in structure (Ref. 2) whereby a measure of the system effectiveness, hereinafter called the "objective function", which involves many decision variables, is to be maximized (or minimized), subject to some limitations or constraints. The aim of any optimization problem is to seek the optimum decision variables that determine the "best" competing activities among limited resources, and determine the maximum (or minimum) objective function.

For a dynamic facility, the design of an efficient and economical energy network is essential, so that the available energy forms are efficiently converted and distributed to the consumer in a cost-effective manner. The development of an optimization model for such a network provides an excellent morphology for the optimal mixing of several different types of resources, multiple components, links, and energy consumers. The optimization model also provides for the optimal operation of such a network when installed.

Currently available simulation models (Refs. 3 to 6) for thermal systems are custom made to solve a preselected design configuration. A few of these models include some component optimizations, yet they have limited applications. An optimization problem with large complex constraints, decision variables, and many possible combinations of energy resources, power plants, and consumers render the model suitable for use with a computer. While the mathematical tools for solving these large optimization problems are available (Refs. 1, 2, 3, and 7), formulating the problem in the "standard" form represents the majority of the effort. As the dynamic facility changes its configuration with time, analytical modeling becomes necessary to operate the facility in its "best" economical condition. Variations in costs, energy prices, weather patterns, load profiles, the energy supply, and reliability constraints could be treated as deterministic, probabilistic, or a combination of both as determined by the analyst.

The Deep Space Network (DSN) facilities that are part of the NASA-owned facilities have identified under several projects the consequences of energy supply shortages and the important need for energy conservation and self-sufficiency. The prime concern has been to reduce operational costs and to improve the facilities reliability and maintainability for successfully supporting deep-space tracking missions. Under the DSN Energy Conservation Project, several feasibility studies are being carried out for the installation of system(s) that will provide reliable energy to the Deep Space Communication facilities in sufficient amount and at a competitive cost.

In Figs. 1 through 4, a few examples are given for the energy systems under study. Figure 1 shows a typical DSN facility in Spain where only diesel engines are used as prime movers. Possible ties to a nearby utility network and further coupling with high-performance fuel cells are being investigated. Decisions to "build" or "no-build," when to start construction, and what capacities should be built are part of the answers sought by the study.

Another example of using the optimization tool is in the search for the optimum-size solar concentrator, as shown in Fig. 2, for gas-fired heat pumps presently near commercialization. A dual utilization of solar energy for direct power via photovoltaic cells and for thermal energy via solar collectors is shown in Fig. 3 coupled to conventional gas and electric power in a building. A larger combination of links and processes is sketched in Fig. 4, where different forms of energy resources share loads with the conventional utilities in meeting the facility demands.

Conservation measures applied to equipment increase the efficiency of energy utilization, thus reducing the consumption of other conventional energy sources. Also, conservation measures applied to a building envelope (such as adding insulation, lighting reductions, etc.) reduce the consumption levels of original resources. A conservation measure treated as an "energy resource" will have a cost per unit of energy that depends on the measure's implementation cost and its potential reduction of demand.

The sample energy systems shown in Figs. 1 to 4 (or any combination of these systems) require an optimization model that can handle a large number of configurations yet be expressed in simple terms for low computational cost to enable wide user acceptance.

Chapman (Ref. 8) has outlined the first energy distribution model, which provided a starting point for this study. His objective was to determine which of several possible power plants and connecting links (decision variables) need to be constructed while satisfying the needs (constraints) of the energy consumers at a minimum total life-cycle cost (objective function). Each component was assumed to have a known cost function of its capacity, reflecting the initial capital cost plus the maintenance cost over the expected lifetime. Chapman's model was limited in scope and was not tested numerically. As a result of on-going energy conservation activities and changes in economic parameters, the facilities configuration is changing and an update of the initial Chapman model seemed necessary. The following section summarizes the present study objectives

## II. Study Objectives

The objectives for developing the energy generation-distribution model are listed below; they are the result of a careful review of the relevant literature.

(1) To provide a working engineering tool for the optimal design of mechanical systems in general, and, in particular, the optimal design of an energy network comprised of a mixture of several different types of energy sources, and serving different consumers. The tool should provide also the optimal construction time and operation of such a network when installed. The

model should include both a "concept optimization" and a "design optimization" within a given concept.

(2) To determine the size or capacity of each add-on element and the distribution pattern for each energy resource, without affecting the network availability, reliability, or continuity of service.

(3) To achieve the above two objectives by an optimum generation-distribution system that either must have a minimum life-cycle cost, or must be built under a limited budget while satisfying other essential constraints such as the energy demands of each consuming building, the capacity constraints of the distribution links, the constraints of equipment reliability and availability during critical tracking periods, the convenience in startup and shutdown, and the periods of restart after a period of shutdown.

(4) The energy model should consider future decreases of building loads as a result of the on-going efforts for energy consumption reduction. Conservation measures should be treated, in a broader view, as an "energy resource" to compete with other conventional and nonconventional resources.

(5) To perform a sensitivity analysis of the model's optimum solution(s) to variations in economic and noneconomic parameters. While near-term cost could be estimated with a reasonable degree of certainty, long-term costs and escalation rates will be probabilistic. This objective is desirable especially when considering the inflation trends experienced during the last decade. Although the study may seek a single optimal solution, sensitivity trends of cost escalations and operation and maintenance changes compared to those of existing systems have to be considered as well.

## III. Approach

A broad variety of energy generation-distribution systems and combinations utilizing several energy resources could be proposed, as shown in Figs. 1 through 4. The major operation is how one should approach the optimal system design. Optimization (sometimes called programming), whether it is in linear or nonlinear form, has been applied to many technical and management decisions. Linear optimization is used when both the constraints and the objective function are expressed by linear or piecewise linear functions of the decision variables. Decomposition techniques are employed if the problem is large and can be partitioned in special form (Ref. 1). The optimum feasible solution will be the particular selection of the decision variables that satisfy all the constraints while minimizing the total life-cycle cost.

In the present study, we will try to set up and formulate the problem for using the well-developed linear programming (LP) methods. Linear programming is known to be very efficient for solving large optimization problems (Ref. 1). In mathematical terms, we will seek to minimize the total life-cycle cost (TLCC) of the system (objective function), which is written in linear terms of $N_u$ unknowns (or decision variables), and set up the constraint equations bounding the solution. Hence, we seek to minimize

$$TLCC = \sum_s C_s X_s \tag{1a}$$

Subject to $N_g$ constraint equations of the equality type:

$$\sum_s a_{rs} X_s = b_r \tag{1b}$$

and

$$X_s \geqslant 0 \tag{1c}$$

where $s = 1, 2, \cdots N_u$ and $r = 1, 2, \cdots N_g$. Equation (1c) constrains all the unknown variables to be nonnegative. Inequality constraints with greater than or equal to ($\geqslant$) or less than or equal to ($\leqslant$) can be converted to equality-constraints by using additional "slack" variables. The format of Eq. (1) is a "standard" format for optimization problems. Several methods are available in the literature (Ref. 1, 2) for solving Eq. (1), of which the Simplex Algorithm is extremely efficient, especially when the number of variables is large. Several versions of the Simplex Algorithm, such as the Revised-Simplex, the Dual-Simplex, and the Revised-Dual Simplex, can be used to reduce the computer memory use and cost. The Dantzig's Decomposition Principle (Ref. 1) is particularly useful for large problems if the elements of the constraint matrix, $a_{rs}$, are diagonalized. In addition, if some of the unknown variables must be integers, the problem is called Mixed-Integer Programming and the Branch-and-Bound Technique or equivalent is further used. For more details on the various methods to solve Eq. (1), see Refs. 1, 2, 3, and 7.

Although the DSN ground stations are typically distributed facilities, selecting a single facility or grouping the facilities into a facility complex will be made for simplicity. Hence, the energy generation and distribution model will be restricted to only *one* consumer to represent one building, a group of buildings, or a complex facility. Existing as well as newly built power plants dealing with various energy types will be

examined for operation in optimum conditions to satisfy the building demands while simultaneously minimizing the total life-cycle cost of the energy network over a certain planning period. The computerized model should be modular and open-ended to accommodate a distributed network consisting of a number of energy sources, conversion processes, storage nodes, consumers, and the internal transmission network that links supply and demand. Requirements for this model will be to provide characteristics of the facility elements as a basis for load analysis, design improvement, and "build" or "no-build" decisions. The format of the model should: (1) be amenable to checking and future modification, (2) require user information that is normally available and clearly definable to avoid ambiguity, and (3) be capable of such efficient execution that undue demands are not placed upon the computer facilities and excessive run-time costs are not entailed.

The type of elements considered in the model are conversion processes, and storage distribution nodes and links. Identical mathematical representations are used to describe any element type within a given time step. However, the presence or absence of particular features for individual elements are accounted for by a set of dichotomous (binary) indices assigned to the individual assemblies. Further, the distribution model allows: (1) the transfer of electrical energy between power plants, (2) intermediate storage-distribution nodes at various sites between plants, and (3) the substitution of electrical energy for direct heating and cooling for any consumer.

Interactions between nodes, processes, and assemblies will include:

(1) The economical competition of a mixture of energy sources and conversion processes to supply the loads at particular consumer sites.

(2) The logistics of the energy transmission network and storage media considerations for load leveling or "shedding."

(3) Selection of component location to be either centralized or distributed with respect to the location of energy resources and consumer sites.

(4) Synthesis of a network to meet time-varying energy usage at the consumption locations and a reliable installed capability to meet peak consumption requirements.

(5) The effects of daily, monthly, and seasonal variations on power generation capability, on environmental influences upon demand, and on the total operation cost including the time of use.

## IV. Assumptions

Mathematically, the model is comprised of a collection of equations that describe the essential properties and controls of the system being modeled. Derivation of energy (or material) relationships is essential. The total system cost (objective function) will be the sum of the total life-cycle cost of individual subsystems, which in turn will include the initial capital cost plus the operation, maintenance, and replacement costs over the study period. A block diagram for any energy generation-distribution system is shown in Fig. 5., where conversion processes are symbolically denoted by circles and storage-distribution nodes are denoted by rectangles. The following assumptions are made:

(1) Type of energy sources: These are end nodes, represented by rectangles in Fig. 5, which include but are not limited to:

  (a) Electromagnetic waves in the solar band (solar radiation).

  (b) Kinetic energy of wind currents (wind energy).

  (c) Solid fossil fuel such as coal and shale.

  (d) Gaseous fossil fuel such as natural gas, methane, or any gaseous coal product.

  (e) Liquid fossil fuel such as diesel oil, gasoline, or any liquid coal product.

  (f) Direct electrical energy supplied by a utility company through an electrical network.

  (g) "Energy conservation," which is treated broadly as an "energy resource." Special care should be taken, however, when it is included in the analysis.

  (h) Water or process steam as a material or energy resource, respectively.

  (i) Air as a material resource for combustion processes.

(2) Energy (or material) conversion processes: These are designated by circles in Fig. 5, and include but are not limited to any of the following:

  (a) Low- and high-concentration-ratio solar collectors for converting solar radiation into thermal energy.

  (b) Photovoltaic cells for direct conversion of solar radiation into electrical energy.

  (c) Wind turbines for direct conversion of wind kinetic energy into electrical energy.

  (d) Solar ponds for heating, electrical power applications, or both.

(e) Internal or external combustion engines, using fossil fuel for electrical power generation, with or without waste-heat recovery.

(f) Boilers for hot water or steam generation using fossil fuel combustion.

(g) Fuel cells for direct conversion to electricity with and without waste-heat recovery.

(h) Mechanically or electrically powered heat pumps for both heating and cooling.

(i) Electrical resistance heaters.

(j) Heat-powered absorption or jet (ejector) refrigeration units for cooling.

(k) Mechanically or electrically driven vapor compression chillers.

(l) Power cycles for electrical power generation driven by such high-temperature heat engines as organic or steam Rankine, Stirling, and Brayton.

(m) Waste-heat recuperators, regenerators, or heat exchangers for heat recovery.

(n) Empty (or dummy) processes to by-pass a stage of nodes to another stage and simplify the mathematics of the problem.

(3) Storage-distribution nodes: Each is represented by a rectangle in Fig. 5, and assumed to consist of one type of energy form or storage media such as (but not limited to):

(a) Electrical energy.

(b) Fossil fuel.

(c) Low-temperature or high-temperature thermal energy, such as cold water or hot water tanks.

(d) Chemical energy (or material) as in oxygen gas, hydrogen gas, and water.

(e) An empty (or a dummy) node for the mathematical simplification of the problem.

Storage capability (such as a flywheel, a battery, or water tanks) could be added to each node. Storage-distribution nodes are controlled by conservation laws relating the influx and outflux of energy (or material). The first law of thermodynamics, the continuity equation, or Kirckhoff's law for electrical nodes could be applied, for example. However, by selecting each node to handle only one type of energy (or material), only one conservation equation is needed for each node. No lateral or cross exchange between nodes within the same stage of nodes is assumed. If lateral nodal linkage is desired, the system network should be organized by utilizing an additional number of empty (or dummy) elements.

(4) Staging: A set of conversion processes and storage-distribution nodes need to be constructed between the available energy sources (supply) on one end and the facility loads (demand), on the other. A large number of different possibilities arise as a result of different types of supply and demand. Grouping of conversion processes and storage-distribution nodes into "stages" is done for convenient sequence identification.

A stage of conversion processes will consist of a number of processes that have the same "level." Hence, all processes in the first stage have to be undergone before the processes in the second stage can begin, and so on. The same is applicable for storage-distribution nodes. Theoretically, there is no limit to the number of processes, nodes, or stages in any system configuration. The only limit is the memory of the available computer to solve the problem.

An assembly is defined here as an integral collection of conversion processes, storage-distribution nodes, and their associated links that could have a "stand alone" function. For example, an add-on photovoltaic array, with its electrical linkage to an existing electrical node is treated as an assembly. In the model, "build" or "no-build" decisions are associated with assemblies rather than elements due to the above integral function.

(5) Conservation equations: Energy (or material) balance of each storage-distribution node is mandatory. The choice to "build" or "no-build" a storage capacity for a given node could be determined by the model. Storage capacities, if a "build" decision is made, are either known in advance or considered part of the decision variables.

(6) Capacity Bounds: If a particular link, conversion process or storage node is to be constructed, its capacity must: (a) exceed the prescribed lower bound for design feasibility, (b) lie below the prescribed upper bound, which may not be necessary in some cases, and (c) provide enough margin above the peak power experienced during operation. These constraints are useful in forcing the mathematics of the optimum solution to yield a practical design for each element size and to meet peak demands without overloading.

(7) Reliability: The reliability of an add-on or new system should be higher than or at least equal to the reliability of the existing system. To satisfy this reliability condition, two approaches could be followed, The first

is to calculate the optimum element capacity, followed by a slight overdesign. The optimum capacity is divided then into a large number of small modules rather than a small number of large modules. The final "practical" system selected by this approach may not be the least expensive, since an overdesign is later incorporated after the analytical solution is obtained. The second approach is to modify the original problem statement by additional size and cost constraints to keep the system's reliability at least the same. The optimum solution in the second approach does not need a readjustment for reliability after it is obtained. The selection of one of the two approaches is optional. However, the first approach is assumed throughout due to its simplicity.

(8) Dummy Elements: The generation-distribution system should allow loading and unloading various links producing the same kind of output energy (or material) when all are connected at a distribution-storage node. Empty (or dummy) elements are permitted to form an empty set with a null effect on the flux passing through. Also, dummy elements should be assigned numbers and be treated the same way as active elements in forming the matrix of equation coefficients.

(9) Centralization: Centralized, distributed, and semicentralized semidistributed systems could be parameterized by the linkage distances, which are either known in advance or assumed to be part of the decision variables.

The present model identifies each element in the system simply by its location and number, and allows changes to be made in the original system configuration without much effort. The cost of installing any element in an assembly is itemized into several linearized parts, which are addressed below.

## V. Cost Analysis

Life-cycle cost (LCC) is an evaluation method that takes into account relevant cost over a selected time period of a system of elements, materials, and operation. It incorporates initial investment costs, future replacement costs, operation and maintenance cost, and salvage values, adjusting them to a consistent time basis and combining them in a single measure that makes it easy to compare alternative options (Refs. 9-13).

The discount rate $i'$ % for a certain time period is the rate of interest that reflects the investor's "opportunity cost" of money, excluding general inflation. The effect of discounting is to obtain the present value equivalent of future cash flow. Whether a discount rate of 7% is selected, as recommended for federal buildings (Ref. 9), or 2% as in DSN practice (Ref. 11) is immaterial at this point. Different discount values will change the present worth equivalent of future savings for projects displacing the use of a conventional fuel. The discount expression is given by:

$$PW \text{ equivalent} = \frac{\text{future cash flow occurring at period } m}{(1+i')^m}$$

Several criteria for LCC analysis are explained in detail in the literature (Ref. 9). Examples are the total life-cycle cost (TLCC), net savings, savings-to-investment ratio, simple payback period, discounted payback period, and annual levelized cost. In this study, the total life-cycle cost (TLCC) is used as the objective function. This sums all significant costs of a system over a certain period, discounted to present value at a selected base year. Detailed life-cycle cost elements could be classified into direct costs for work force and indirect costs for such items as services, travel, and computer programming. A broader classification is made according to the execution sequence (Ref. 12) as follows:

(1) Future research and development cost.

(2) Future planning cost including cost of feasibility studies.

(3) Implementation cost including cost of design reviews, testing, quality assurance, and installation.

(4) Maintenance and operations cost including cost of utility and preventive maintenance.

(5) Sustaining cost including the cost of replacement and modifications.

(6) Salvage cost.

Each of the above cost elements should be multiplied by an inflation (or discount) factor and an overhead factor. Cost estimates are provided in constant dollars computed at a base year, which is usually the year in which the study is performed.

The key elements of the LCC methodology used in this study follow the NBS rules (Ref. 9), which are utilized in Appendix A for the input data. The obligations that occur at different times should be adjusted to a common time basis, which might be: (1) the present, whereby all expenses are converted to an equivalent value occurring now, (2) annually (or levelized), whereby all expenses are converted to an equivalent value occurring in a uniform amount each year over the study period, and (3) the future, whereby all expenses are converted to an equivalent value occurring at some common

time in the future. For convenience, the present worth basis will be used.

The "economic efficiency" of a given system configuration can be improved by the optimal timing of investment decisions. To have a "wait option" means timing the investment so as to capture the largest possible long-run benefits in the face of technological advances and other future changes.

The monthly recurring fuel or nonfuel operation and maintenance costs are assumed to begin to accrue at the beginning of the base year. These are evaluated as lump sum amounts at the end of each *month* over the study period, starting with the end of the first month of a base year. Treating investment costs as a lump sum occurring at the beginning of the construction time is a simplified approach somewhat less accurate than a detailed analysis employing actual scheduling and accounting of costs. The difference in the two approaches is generally not large. Nonannually recurring replacement costs and salvage values are temporarily neglected in the analysis.

The time step-by-step method of calculating the present worth energy cost is suitable since both the quantity and type of energy in each element are expected to change periodically, for each hour, day and month. Due to the time changing of installation costs, the development of new technologies, the declining of a component efficiency with time, the increase of maintenance cost by aging, and the escalation of energy costs and maintenance costs as a result of inflation, it becomes necessary to have the model divided into multiple time steps. The structure details of the model are explained in the next section.

## VI. Equation Formulation

The relationships between the decision variables and the constraint equations are grouped by the headings that follow.

### A. Time Steps

In the model, a distinction should be made between an energy (or material) flux time step, $m$ (where $m = 1, 2, \cdots N_t$) and a construction time step $n$, (where $n = 1, 2, \cdots, N_{ct}$) as illustrated in Fig. 6. The first type time steps are assumed to be four for each representative day for each month to account for four different time-of-day uses for $N_y$ years under study. This means a total of $48\,N_y$ time steps. Construction time steps, on the other hand, indicate the time-frequency at which management could make the start-of-construction decisions. For instance, due to known budgetary cycles or management directions, the beginning of construction within the user's organization is assumed to take place once a year, once every 6 months, once every 3 months, or once a month. Hence, each year of the study period will be divided into 1, 2, 4 or 12 construction-time steps where decisions to "build" or "no-build" are made and construction monies are spent. The construction time steps are $N_y$, $2\,N_y$, $4\,N_y$, or $12\,N_y$, respectively. The selection of small periods for energy (or material) flux computations and large periods for construction time decisions forces the solution and the feasible region[1] to eliminate undesirable answers. Without the construction-time steps, a new assembly built when $m = 59$, for instance, means that construction must start in the second year of study, the month of March and off-peak period, which is not a practical time. Although the selection of $n$ is arbitrary, it is restricted in this model such that the start-of-construction decisions are not within a month. Since four consecutive energy (or material) flux time steps represent a one-month period, the number of decisions to "build" or "no-build" during $N_y$ years becomes:

$$\left.\begin{aligned} N_{ct} &= N_{t/4} = 12\,N_y \text{ for monthly construction decisions} \\ N_{ct} &= N_{t/12} = 4\,N_y \text{ for quarterly construction decisions} \\ N_{ct} &= N_{t/24} = 2\,N_y \text{ for semiannual construction decisions} \\ N_{ct} &= N_{t/48} = N_y \text{ for annual construction decisions} \end{aligned}\right\} \quad (2)$$

where $N_t = 48\,N_y$.

### B. Losses in Transmission Links

The energy (or material) flux travelling through the transmission links between a conversion process and a storage-distribution node are generally expressed as:

$$\overline{E}(k) = E(k) - \alpha(k)\,L(k)\,E^{\gamma}(k)$$

where $E(k)$ is the destination energy (or material) flux at the receiving end of the link $k$, $E(k)$ is the origin energy (or material) flux at the sending end of the link, $L(k)$ is the link length, $\alpha(k)$ is a proportionality loss coefficient, and $\gamma$ is an exponent between 1 and 2. The coefficient $\alpha$ is a characteristic link constant. To suit a linear programming format, we assume

$$\overline{E}(k) = T(k)\,E(k) \quad (3)$$

where $k$ is the number assigned to the link joining the process $i$ and the node $j$, and $T(k)$ is the transmission efficiency, taken

[1] This is the mathematical region which contains the feasible solutions of the problem that satisfy all the constraint equations.

as a constant. Note that the indices $i$, $j$, and $k$ will take the following values:

$$i = 1, 2, \cdots N_c, j = 1, 2, \cdots N_o, \text{and } k = 1, 2, \cdots N_\ell$$

where $N_c$, $N_o$, and $N_\ell$ are the number of processes, storage-distribution nodes, and connecting links, respectively.

## C. Conversion Efficiency

In handling the conversion efficiency expressions, the following convention is used. A simple process is defined as having one input link and one output link. A compound process involves, in general, more than one input link and more than one output link, and could be decomposed into a number of simple processes as shown in Fig. 7. For any link $k$, connected to a process inlet, a coupling identifier $p(i,j) = -k$ is assigned. For any outlet link, an identifier $p(i,j) = k$ is assigned. The conversion efficiency (or yield) $F$ is defined for a simple process, as sketched in Fig. 8, by the ratio of output flux to input flux in the output and input links, respectively, or

$$F(k_{in}, k_{out}) = \frac{E(k_{out})}{\overline{E}(k_{in})} \tag{4}$$

where the first index, $k_{in}$, refers to the process inlet link number, the second index, $k_{out}$, refers to the process outlet link number. Combining Eqs. (3) and (4) to eliminate $\overline{E}$, the set of conversion efficiency equations are written in terms of the $E$ unknowns as:

$$E(k_{out}) - F(k_{in}, k_{out})\, T(k_{in})\, E(k_{in}) = 0 \tag{5}$$

Equation (5) represents $N_c$ linear equations for $N_c$ simple processes, applicable for each time step.

## D. Storage-Distribution Nodes

Energy (or material) conservation equations must be written for the storage-distribution nodes, including those end-type nodes that represent either the energy supply reservoirs or building demands. Selecting a uniform state, uniform flow (USUF) control volume for the storage-distribution nodes, $N_{sd}$, will encompass internal flux changes during the time interval $\Delta t(m)$ from $t(m-10$ to $t(m)$. From Fig. 9, a summation over all process inlet and outlet links would give:

$$\underbrace{\Delta t(m) \sum_{k_{out}} \overline{E}(k_{out}, m)}_{\text{Sum over inlet links to a node (process outlet links where } p \text{ is positive)}} \underbrace{- \Delta t(m) \sum_{k_{in}} E(k_{in}, m)}_{\text{Sum over outlet links from a node (process inlet links where } p \text{ is negative)}} \underbrace{- SC(j,m) + SC[j,(m-1)]}_{\Delta \text{ Energy stored, within node from initial time } t(m-1) \text{ to final time } t(m)} = 0 \tag{6}$$

where $E$, $\overline{E}$ are rates computed at time $t(m)$, and $SC(j,m)$ is the internal energy (or material) content of storage node $j$ at time $t(m)$. The fluxes $E$ and $\overline{E}$ are assumed constant during the interval $\Delta t(m)$. If no nodal storage exists, the conservation equations for simple distribution nodes $N_d$ will be reduced to the case of a steady-state, steady-flow (SSSF) control volume where all $SC(j,m)$ are eliminated:

$$\sum_{k_{out}} \overline{E}(k_{out}, m) - \sum_{k_{in}} E(k_{in}, m) = 0 \tag{7}$$

Combining Eqs. (3) and (6) give for the conservation of the $j^{th}$ node out of $N_{sd}$ nodes:

$$\Delta t(m) \sum_{k_{out}} T(k_{out}, m)\, E(k_{out}, m) - \Delta t(m) \sum_{k_{in}} E(k_{in}, m) - SC(j,m) + SC(j,(m-1)) = 0 \tag{8}$$

Also, for $N_d$ distribution nodes, Eqs. (3) and (7) Give:

$$\sum_{k_{out}} T(k_{out}, m)\, E(k_{out}, m) - \sum_{k_{in}} E(k_{in}, m) = 0 \tag{9}$$

## E. Supply/Demand End Nodes

The end storage-distribution nodes representing an energy resource (supply) or a building load (demand) are treated similarly to Eq. (9), but with a minor modification. For an end node $j$ out of $N_{rd}$ nodes representing an energy resource distribution with intensity $S(j,m)$, Eq. (9) is changed to:

$$S(j,m)\, \overline{A}(j) - \sum_{k_{in}} E(k_{in}, m) = 0 \tag{10}$$

where $j = 1, 2, \cdots N_{rd}$.

Also, for a demand-distribution node $j$ out of $N_{dd}$ nodes representing each given type of building demand, $D(j,m)$, Eq. (9) is changed to:

$$\sum_{k_{out}} T(k_{out}, m)\, E(k_{out}, m) = D(j,m) \qquad (11)$$

where $j = 1, 2, \cdots N_{dd}$.

Note that either the product $[S(j,m)\ \overline{A}(j)]$ or the characteristic area $\overline{A}(j)$ will be the decision variable. For instance, in an energy system consisting of a field of solar collectors or a set of wind turbines, the energy intensities $S(j,m)$, will be known quantities and the decision variable in this case will be the characteristic projected area $\overline{A}(j)$ of either the solar collector field, or the blade area of wind turbines, respectively. If, however, energy is purchased from a supplier (such as a utility), the unknown product $[S(j,m)\ \overline{A}(j)]$, which is the flux, becomes the decision variable. Equations (8), (9), (10), and (11) represent a total of $N_o$ equations for $N_o$ nodes for each $m$ period.

## F. Storage Capacity Limit

There is a physical limitation in charging or discharging during any period $\Delta t(m)$ that constrains the changes in storage content to be always less than or equal to the design capacity $V(j)$. For additional cost saving, decisions are made to build or no-build storage media to match optimally the transient supply and demand curves without changing their daily sum of fluxes. Two decisions need to be made: the first is whether or not storage nodes of a new assembly need to be built and the second is to know which capacity should be designed for lowest cost. Since four time steps are assumed for each representative day of each month, it is appropriate to select the optimum storage capacity $V(j)$ that handles supply and demand matching only for a 24-hour time span. This means that at the beginning of the first day-period (where $m = 0, 4, 8, \cdots$, etc.), the storage content $SC(j,m)$ should be taken as zero (i.e., empty storage). The storage content must also be zero at the end of the fourth day period, which is also the beginning of the first day period of the next day representing the next month. Hence,

$$\left.\begin{aligned} SC(j,0) &= 0 \\ SC(j,4\nu) &= 0 \end{aligned}\right\} \qquad (12)$$

where $\nu$ is monthly time step, $(\nu = 1, 2, \cdots, 12\,N_y)$

This leaves us with the determination of the storage content only at the three intermediate times within a day, where $m = 4\nu - 3$, $4\nu - 2$ and $4\nu - 1$. Some storage nodes should be allowed, when fully charged, to "dump" the excess supply. This is the case, for instance, when a solar collector designed primarily for winter use is used for heating in a summer period. Another example is to intentionally waste available wind energy when both the supply exceeds the demand and the storage nodes are at full charged capacity. Excess energy (or material) flux should be treated in these cases as "slack" variables. During the study period, the storage content profile fluctuates from positive to negative regions as shown in Fig. 10. Although the storage node is designed to provide matching between supply and demand for one day only, different matching schemes over all the days of the study period need to be examined.

To allow for both negative and nonnegative values of the variable $SC$, it is commonly represented, for later use in the Simplex Algorithm, by the difference between two nonnegative variables:

$$\left.\begin{aligned} SC(j,m) &= SC'(j,m) - SC''(j,m) \\ SC'(j,m) &\geqslant 0 \\ SC''(j,m) &\geqslant 0 \end{aligned}\right\} \qquad (13)$$

Although Eq. (13) introduces an additional set of variables, it eliminates the need for nonnegativity constraints. In the final solution, the set $SC'(j,m)$ is reduced to $SC(j,m)$ if $SC(j,m) \geqslant 0$, thus making $SC''(j,m) = 0$. Also, the set $SC''(j,m)$ is reduced to $|SC(j,m)|$ if $SC(j,m) \leqslant 0$ thus making $SC'(j,m) = 0$. In other words,

$$SC'(j,m) = \begin{cases} SC(j,m) & \text{if } SC(j,m) \geqslant 0 \\ 0 & \text{otherwise} \end{cases} \qquad (14)$$

and

$$SC''(j,m) = \begin{cases} |SC(j,m) & \text{if } SC(j,m) \leqslant 0 \\ 0 & \text{otherwise} \end{cases} \qquad (15)$$

Further, the optimum capacity of the storage node $V(j)$ could be constrained to be either less than or equal to (i.e., undersized) or larger than or equal to (i.e., oversized) the maximum swing of the variable $SC$.

Defining the maxima of the sets $SC'(j,m)$ and $SC''(j,m)$ over all the $N_t$ time periods as $MSC'(j)$ and $MSC''(j)$, respectively, one may write the two possible constraint equations for $V(j)$ as either:

$$MSC'(j) + MSC''(j) - V(j) \geq 0 \qquad (16)$$

or

$$MSC'(j) + MSC''(j) - V(j) \leq 0 \qquad (17)$$

Equations (16) and (17) represent two different decisions: Eq. (16) constrains the storage capacity to be toward undersizing, and Eq. (17) toward oversizing. These two decisions differ in their impact on installation cost (due to the size of $V(j)$) and on the operation cost (due to the impact on the matching role between supply and demand fluxes). Only one out of the two constraints of Eqs. (16) and (17) must hold for any single node $j$. To solve this either/or constraints problem, two additional sets of dichotomous variables $\beta_1(j)$ and $\beta_2(j)$ should be introduced in conjunction with the "big $M$" method, (Ref. 2). $M$ is a very large number, arbitrarily selected to be larger than any feasible $MSC'$ or $MSC''$. Equations (16) and (17) are then rewritten as:

$$\left.\begin{aligned} MSC'(j) + MSC''(j) &\geq V(j) + M\beta_1(j) \\ MSC'(j) + MSC''(j) &\leq V(j) + M\beta_2(j) \\ \beta_1(j) + \beta_2(j) &= 1 \\ \beta_1(j), \beta_2(j) &\leq 1 \\ \beta_1(j), \beta_2(j) &\text{ integers} \end{aligned}\right\} \qquad (18)$$

Note that adding $M$ to the right-hand side of a constraint equation has the effect of eliminating it. The formulation of Eq. (18) guarantees that one out of the two original constraints must hold. In addition, the maxima $MSC'(j)$ and $MSC''(j)$ could be expressed by the following linear set of constraints:

$$SC'(j,q) - MSC'(j) \leq 0 \quad (q = 1, 2, \cdots 36N_y) \qquad (19)$$

and

$$SC''(j,q) - MSC''(j) \leq 0 \quad (q = 1, 2, \cdots 36N_y) \qquad (20)$$

where $q$ is the total number of intermediate times of day excluding the start and end times. For three intermediate times per day, 12 days per year, and $N_y$ years of the study period, $q$ ranges from 1 to $36N_y$.

## G. Beginning of Construction Time

For each assembly, $\ell$, there will be associated a set of conversion processes, storage-distribution nodes and links. For each construction time period $n$, there will be assigned a dichotomous decision variable $\lambda(\ell, n)$. The variable $\lambda$ is a binary integer that can be either 0 or 1. A zero value for $\lambda(\ell, n)$ means that assembly $\ell$ is not to be constructed at the beginning of construction-time interval $n$. A value of one assigned for $\lambda(\ell, n)$ means that for assembly $\ell$, construction starts at the beginning of construction time interval $n$. Construction of any assembly, no matter how large its size, is assumed to take place only once during the total period under study. The present model does not consider the possibility of adding capacity to any assembly sequentially, in multiple sizes and during multiple time steps. Therefore,

$$\sum_n \lambda(\ell, n) \leq 1, \quad \ell = 1, 2, \cdots N_a \qquad (21)$$

where all $\lambda(\ell, n)$ elements are subject to the nonnegativity and integer conditions:

$$\left.\begin{aligned} &\lambda(\ell, n) \geq 0 \\ &\lambda(\ell, n) \leq 1 \\ &\lambda(\ell, n) \text{ is an integer} \end{aligned}\right\} \qquad (22)$$

## H. Construction Cost of New Assemblies

The introduction of the dichotomous variables $\lambda(\ell, n)$ makes it feasible to reduce the construction and implementation cost of an assembly built at a decision time $\bar{n}(\ell)$ to the present worth at the base year. Let the construction cost of an element $(e)$ in an assembly $\ell(e \epsilon \ell)$ to be represented by a linear relationship with its size $P(e)$ in current dollars at the beginning of $n$

$$CC(e) = CCI(e) + CCS(e)\,P(e) \qquad (23)$$

where $CC(e)$ is the construction cost of an element $e$, $CCI(e)$ and $CCS(e)$ are the constant cost and the variable cost, respectively, for element $e$ $(e = 1, 2, \cdots N_{el})$. Discounting the current costs to the base year worth and summing over all elements of assembly gives present worth of totals construction cost (TCC):

$$TCC \equiv \left[\sum_{n=1}^{N_{ct}} \sum_e \lambda(\ell, n)\, CCI(e) + \sum_{n=1}^{N_{ct}} \sum_e P(e, n)\, CCS(e)\right] \Big/ [1 + i'(n)]^{n-1} \qquad (24)$$

The first term in the right-hand of Eq. (24) will only be nonzero when $\lambda(\ell, \overline{n}) = 1$, and is zero otherwise. The second term in the right-hand side is written in terms of "periodical ratings" of the $e^{th}$ element, $P(e, n)$, which are subject to additional constraints as described below, such that $P(e, n)$ are constrained to be only nonzero at a single $n = \overline{n}$ and are zero otherwise.

## I. Lower and Higher Bounds of Element Rating

If an element $e$, in an assembly $\ell(e \epsilon \ell)$ is to be constructed at the beginning of interval $\overline{n}$ (i.e., $\lambda(\ell, n) = 1$) then its rated capacity must be greater than or equal to some given lower limit $P_L(e)$ and below or equal to some given upper limit $P_H(e)$. On the other hand, if the rating is below the lower bound or above the higher bound, as illustrated in Fig. 11, then the element must not be constructed $[\lambda(\ell, n) = 0]$ and the element rating should be forced to equal zero, otherwise its construction cost is included in Eq. (24). Therefore, for the lower bound:

$$P(e, n) \geqslant \lambda(\ell, n) P_L(e) \quad \text{for all } e \epsilon \ell, \text{ all } n$$

or

$$P(e, n) - \lambda(\ell, n) P_L(e) \geqslant 0 \tag{25}$$

and for the upper bound:

$$P(e, n) \leqslant \lambda(\ell, n) P_H(e) \quad \text{for all } e \epsilon \ell, \text{ all } n$$

or

$$P(e, n) - \lambda(\ell, n) P_H(e) \leqslant 0 \tag{26}$$

Note that if an element is to be constructed, only one $\lambda(\ell, n)$ at $n = \overline{n}$ will be equal to 1, which means that the present worth of element construction cost from Eq. (24) will be nonzero only at the period $\overline{n}$, and zero otherwise.

Another way of forcing the periodical rating $P(e, n)$ to zero if the element is not to be constructed (i.e., if $\lambda(\ell, n) = 0$) is to use only a lower bound together with a "big $M$" control instead of the above lower and upper bounds. In this case, the upper bound constraint, Eq. (26), is replaced by:

$$P(e, n) \leqslant M \lambda(\ell, n)$$

or

$$P(e, n) - M \lambda(\ell, n) \leqslant 0 \tag{27}$$

where $M$ exceeds the maximum feasible value of any $P(e, n)$. Equation (27) ensures that when $\lambda(\ell, n)$ is zero, $P(e, n)$ must be zero, and when $P(e, n) > 0$, the variable $\lambda(\ell, n)$ will be 1 (and not zero). The optimal solution that minimizes the objective function will always choose $\lambda(\ell, n) = 0$ when $P(e, n) = 0$.

Equation (27) must be used if neither upper nor lower bounds are imposed on $P(e, n)$, or combined with Eq. (25) or (26) as appropriate.

## J. Periodical Flux in Links

Even if the start-of-construction decision is taken for assembly $\ell$ at an optimum time period $\overline{n}(\ell)$ where $\overline{n} = 1, 2, \cdots N_{ct}$, the energy (or material) interactions in all the assembly links, conversion processes and storage-distribution nodes will not start until construction is complete, i.e., after a given time delay $\theta(\ell)$ is elapsed. This construction time delay $\theta(\ell)$ could take, in general, any value. However, in this model it is rounded off to the nearest month, hence constrained to be a multiple of months. Therefore, the benefit of constructing a new assembly will commence or be accounted for at the beginning of the month immediately following the end of the construction period. Additional constraints are needed to ensure that the energy (or material) flux carried in the links of each new assembly must commence after the construction is complete at $t(\overline{m})$:

$$SC(j, m), E(k, m) \begin{cases} = 0 & 0 \leqslant t(m) \leqslant t(\overline{m}) \\ \geqslant 0 & t(N_t) \geqslant t(m) > t(\overline{m}) \end{cases} \tag{28}$$

where

$$\lambda(\ell, \overline{n}) = 1, \quad n = \overline{n}$$
$$\lambda(\ell, n) = 0, \quad n \neq \overline{n}$$
$$m = 1, 2, \cdots N_t$$

Two sets of dichotomous variables $\nu$ and $\eta$ are further derived from the original set $\lambda$ as illustrated in Fig. 12. The first set $\nu(\ell, m)$ are equivalent to the set $\lambda(\ell, n)$ but displaced $\theta(\ell)$ months due to the construction delay. Hence, for $n = 1, 2, \cdots [(N_t - 4\theta)/q' + 1]$:

$$\nu(\ell, m) = \begin{cases} 0 & 1 \leqslant m \leqslant 4\theta(\ell) \\ 0 & m \neq 4\theta + (n - 1) q', m > 4\theta \\ \lambda(\ell, n) & m = 4\theta + (n - 1) q' \end{cases} \tag{29}$$

where $q' = 4$, 12, 24, and 48 for monthly, quarterly, semi-annual, and annual construction decisions, respectively. The set $\eta(\ell, m)$, on the other hand, is the cumulative sum of $\nu(\ell, m)$ where:

$$\eta(\ell, m) = \sum_{r=1}^{m} \nu(\ell, r) \tag{30}$$

Equations (29) and (30) guarantee that the set $\eta$ will be zero before the time $t(\overline{m})$, and unity thereafter. Accordingly, the

energy (or material) flux in the various links will be subject to the following constraints:

$$E(k,m) \leqslant \eta(\ell,m)\, M$$

or

$$E(k,m) - \eta(\ell,m)\, M \leqslant 0 \tag{31}$$

Equation (31) forces $E(k,m)$ to be zero if $\eta(\ell,m)$ is zero, and eliminates the constraint equation as $\eta = 1$.

## K. Rating of Elements

The selection of the proper design capacity for an element ($e$) in an assembly ($\ell$) is subject to the peak flux value experienced during the element operation. Note that the daily time step, $\Delta t$, which is selected for this model to be in the order of 4 to 10 hours, makes the peak flux an "apparent peak" and not a "true peak." A "true peak" flux is commonly measured during a 15-minute sampling; therefore, two additional design allowances should be considered for a "safe" optimum design. The first design allowance, as shown in Fig. 13, is to make the element rated capacity $P(e)$ slightly larger than the "true peak" flux. The second allowance relates the "true peak" flux to the "apparent peak" flux. Assuming that the ratio between the element rating $P(e)$ to the "apparent peak" flux is 1.58 as determined from Fig. 13, then

$$1.58\, E(k,m) \leqslant P(e)$$

or

$$1.58\, E(k,m) - P(e) \leqslant 0 \quad (e,k) \epsilon \ell \tag{32}$$

Equation (32), which represents $m$ equations for each link $k$, constrains the ceiling value of the link flux to be less than the element rated capacity. Furthermore, to limit the periodical element rating $P(e,n)$ to be less than or equal to the element rating $P(e)$, we write for all $n = 1, 2, \cdots N_{ct}$ and all $e \epsilon \ell$

$$P(e,n) - P(e) \leqslant 0 \tag{33}$$

## L. Utility Costs

the costs of purchased energy (or material) will be summed over the study period on a monthly basis, where for each month the costs are discounted to the base year. The present worth of total operation cost (TOC) is:

$$TOC = \sum_m \sum_k E(k,m)\, \frac{1 + e'(v)}{[1 + i'(v)]^v}\, Cu[k,(m-1)]\, \Delta t(m) \tag{34}$$

where $Cu$ is the unit cost of purchased energy (or material) flux transmitted by link $k$, calculated at the beginning of the month corresponding to the period $m$. Unit cost and escalation rates are assumed to change only each month (i.e., every 4 periods).

## M. Maintenance and Sustaining Costs

Similar to the construction cost of an element $e$ of an assembly $\ell$ in Eq. (24), the maintenance and sustaining costs are assumed to be composed of two parts:

(1) A constant, uniformly recurring maintenance cost, $MCI(e,v)$, for the element $e$, at the end of the month $v$. Cost accrual starts only at the end of the month following the construction completion, and the start of assembly operation ($m > \overline{m}$). The dichotomous variable $\eta(\ell, 4v)$, given by Eq. (30), must be introduced.

(2) A variable maintenance cost, uniformly recurring, and proportional to the size or rated capacity of the constructed element. Also cost accrual should start only at the end of the month following the construction completion. The proportionality cost is $MCS(e)$.

Discount of both parts (a) and (b) should be made to a base-year worth, using a monthly maintenance escalation rate, $m'(v)$, which is counted above general inflation. Hence, present worth of total maintenance cost (TMC),

$$TMC = \frac{\sum_e \sum_{v=1}^{12\,N_y} [\eta(\ell, 4v)\, MCI(e,v) + MCS(e)\, \overline{P}(e,v)]}{[1 + m'(v)^v} \tag{35}$$

where $\overline{P}(e,n)$ are "modified periodical rating" of the element capacity, which are derived for maintenance cost calculations from $P(e,n)$. Note that $\overline{P}(e,v)$ are nonzero at $v \geqslant \overline{v}$ if an assembly $\ell$ is built (i.e., if $\lambda(\ell, \overline{n}) = 1$); otherwise, the rest of the periodical ratings at $v < \overline{v}$ will be zero. This is different from the variable construction cost $[P(e,n)\, CCS(e)]$, which is nonzero only once for the whole study period at $n = \overline{n}$.

The variable maintenance cost $[\overline{P}(e,n)\, MCS(e)]$, must be calculated for all time periods immediately after the construction is completed. The modified periodical ratings are also identical to the decision rating $P(e)$ for each month $v$, but must vanish at $v < \overline{v}$, and be nonzero as $v \geqslant \overline{v}$, hence,

$$\overline{P}(e,v) - M\eta(\ell, 4v) \leqslant 0 \tag{36}$$

$$\overline{P}(e,v) - P(e) \leqslant 0 \tag{37}$$

Equation (36) guarantees that $\bar{P}(e, v)$ will be zero as $\eta(\ell, 4v)$ is zero (i.e., at $v \leqslant \bar{v}$) and eliminates any constraint on $\bar{P}(e, v)$ as $\eta(\ell, 4v)$ is one. Equation (37), however, adds the additional constraint that each nonzero $\bar{P}(e, v)$ must be equal to $P(e)$.

## N. Total Life-Cycle Cost

The total life-cycle cost (TLCC) of the energy network will be the summation over all the elements and all the time periods of: (1) the total construction cost, *TCC*, from Eq. (24), (2) the total operation cost, *TOC*, from Eq. (34), and (3) the total maintenance and sustaining cost, *TMC*, from Eq. (35). The objective function *TLCC* is written as:

$$TLCC - (TCC + TOC + TMC) = 0 \tag{38}$$

where minimum *TLCC* is sought.

## O. Summary of Constraints

The constraints equations described above are listed as follows:

(1) Process efficiency:

$$E(k_{out}, m) - F^{(m)}(k_{in}, k_{out})\, T(k_{in}, m)\, E(k_{in}, m) = 0 \tag{5a}$$

This represents $N_c$ equations for each month.

(2) Conservation laws for nodes. Each is taking-in one of the four forms below.

for $N_{sd}$ storage-distribution nodes:

$$\Delta t(m) \sum_{k_{out}} T(k_{out}, m)\, E(k_{out}, m) - \Delta t(m) \sum_{k_{in}} - SC'(j, m) + SC''(j, m) + SC'[j, (m-1)] - SC''[j, (m-1)] = 0 \tag{8}$$

for $N_d$ distribution nodes:

$$\sum_{k_{out}} T(k_{out}, m)\, E(k_{out}, m) - \sum_{k_{in}} E(k_{in}, m) = 0 \tag{9}$$

for $N_{rd}$ resource-distribution nodes:

$$S(j, m)\, \bar{A}(j) - \sum_{k_{in}} E(k_{in}, m) = 0 \tag{10}$$

for $N_{dd}$ demand distribution nodes:

$$\sum_{k_{out}} T(k_{out}, m)\, E(k_{out}, m) = D(j, m) \tag{11}$$

Equations (8) through (11) represent $N_o$ equations for each $m$ period in $N_t$.

(3) Storage content at beginning and end of each day:

$$\left.\begin{aligned} &SC'(j, 0), SC''(j, 0) = 0 \\ &SC'(j, 4v), SC''(j, 4v) = 0 \quad (v = 1, 2, \cdots 12 N_y) \end{aligned}\right\} \tag{12}$$

Therefore, we need only to compute the intermediate values of $SC'(j,m)$ and $SC''(j,m)$ where $m = 4v - 3$, $4v - 2$, $4v - 1$ for $v = 1, 2, 3, \cdots 12N_y$, and $j = 1, 2, \cdots N_{sd}$.

(4) Maximum swing of storage content vs size:

$$\left.\begin{aligned} MSC'(j) + MSC''(j) - V(j) - M\beta_1(j) &\geqslant 0 \\ MSC'(j) + MSC''(j) - V(j) - M\beta_2(j) &\leqslant 0 \\ \beta_1(j) + \beta_2(j) &= 1 \\ \beta_1(j), \beta_1(j) &\leqslant 1 \\ \beta_1(j), \beta_2(j) &\text{ integers (0 or 1)} \end{aligned}\right\} \tag{18}$$

where $j = 1, 2, \cdots N_{sd}$. Equation (18) represents five equations for each of the $N_{sd}$ storage-distribution nodes.

(5) Maxima of storage content:

$$SC'(j, q) - MSC''(j) \leqslant 0 \tag{19}$$

$$SC''(j, q) - MSC''(j) \leqslant 0 \tag{20}$$

where $q = 4v - 3$, $4v - 2$, $4v - 1$, and $v = 1, 2, \cdots 12 N_y$, $j = 1, 2, \cdots N_{sd}$. Equation (19) or (20) represents $36 N_y N_{sd}$ equations.

(6) Construction dichotomous variable, $\lambda$:

$$\sum_{n} \lambda(\ell, n) \leqslant 1 \tag{21}$$

$$\left.\begin{aligned} &\lambda(\ell, n) \leqslant 1 \\ &\lambda(\ell, n) \text{ integer (0 or 1)} \end{aligned}\right\} \tag{22}$$

where $n = 1, 2, \cdots N_{ct}$, and $\ell = 1, 2, \cdots N_a$. Equation (21) represents $N_a$ equations for $N_a$ assemblies. Equation (22), however, represents $N_a$ equations for each period $n$.

(7) Lower and higher bounds of rated capacity:

For each $n$ $(= 1, 2, \cdots N_{ct})$ and $e \epsilon \ell$, where $e = 1, 2, \cdots N_{el}$, one may write

$$P(e,n) - \lambda(\ell,n)\, P_L(e) \geqslant 0 \qquad (25)$$

$$P(e,n) - \lambda(\ell,n)\, P_H(e) \leqslant 0 \qquad (26)$$

if no bound is imposed, Eq. (25), (26), or both could be replaced by:

$$P(e,n) - M\,\lambda(\ell,n) \leqslant 0 \qquad (27)$$

Equations (25), (26), or (27) represents $N_{e\ell}$ equations for each time period $n$.

(8) Flux dichotomous variables $(\nu, \eta)$:

$$\nu(\ell,m) = 0 \quad 0 \leqslant m < 4\theta(\ell)$$

$$\nu(\ell,m) - \lambda(\ell,n) = 0 \quad \begin{cases} m = 4\theta + (n-1)\,q' \\ n = 1, 2, \cdots, \left(\dfrac{N_t - 4\theta}{q'} + 1\right) \end{cases}$$

$$\nu(\ell,m) = 0 \quad \begin{cases} m > 4\theta \\ m \neq 4\theta + (n-1)\,q' \end{cases} \qquad (29)$$

Equation (29) represents $N_{el}$ equations for $N_{el}$ elements in $N_a$ assemblies. Also,

$$\eta(\ell,m) - \sum_{r=1}^{m} \nu(\ell,r) = 0 \qquad (30)$$

Equation (30) represents $(48\, N_y \cdot N_{el})$ equations.

(9) Flux in links after construction completion:

$$E(k,m) - \eta(\ell,m)\; M \leqslant 0 \qquad (31)$$

Equation (31) represents $N_{\ell}$ equations for $\ell$ links during each $m$ time step.

(10) Rated capacity of elements:

$$1.58\, E(k,m) - P(e) \leqslant 0 \qquad (32)$$

where Eq. (32) represents $N_{\ell}$ equations for each $m$ time step. Also,

$$P(e,n) - P(e) \leqslant 0 \qquad (33)$$

Equation (33) represents $N_{e\ell}$ equations for each $n$ time period.

(11) Modified periodical rating:

$$\bar{P}(e,\nu) - M\,\eta(\ell,4\nu) \leqslant 0 \qquad (36)$$

$$\bar{P}(e,\nu) - P(e) \leqslant 0 \qquad (37)$$

Equations (36) and (37) represent $N_{el}$ equations for each monthly period $\nu$.

## VII. Numerical Solution

Upon defining a block diagram of a "preferred" system configuration, the user will provide only some input data tabulated and grouped as listed in Appendix A. Currently, a computer program (OMEGA) for the *O*ptimization *M*odel of *E*nergy *Gen*era*t*ion and Distribution is being written using the objective function of Eq. (38) and the above-mentioned constraints. The data entered by the user are grouped as follows:

(1) *Configuration data* include the number of system elements (processes, storage-distribution nodes and links), the coupling between different elements and whether or not the capacity of a certain element is known or treated as an unknown decision variable, and the number of assemblies.

(2) *Conversion efficiency* could be assumed constant for the period under study or else given for some selected time periods to reflect changes due to aging, wear, partial loading, and increased maintenance.

(3) *Facility loads* include the electrical-connected loads, space heating load, space cooling load, domestic hot water heating loads, and process steam. The data are entered for each time step, and should account for changes in load profile due to the facility growth or decrease in activity.

(4) *Energy resource intensity S* of an available resource for the energy supply end nodes should be entered at the appropriate time steps. The data include: (1) the solar radiance for different collector orientations, tracking mechanisms, site latitude, and ground reflectivity, and

(2) the site's wind velocities, wind-turbine characteristics, and cut-in, rated, and cut-off velocities.

(5) *Time data* include the number of years for a total life-cycle cost study, $N_y$, delay time due to implementation $\theta$, on-peak, midpeak and off-peak schedule.

(6) *Economics data* include constant and variable construction costs of elements, constant and variable maintenance costs, escalation rates of purchased energy units, $e'$, maintenance costs, $m'$, money discount rates, $i'$, and reliability costs.

The printed output of the OMEGA program will include an echo of input data for cross checking, the minimum value of TLCC (objective function) and the following set of optimum decision variables:

(1) The energy (or material) flux $E(k,m)$ carried by each link $k$ during each time $t(m)$, whether the housing assembly is built or not.

(2) The construction decision variables $\lambda(\ell, n)$ giving the "build" or "no-build" decisions for new assemblies, and storage capacities. Also, the dichotomous variables $\nu$ and $\eta$ will be printed.

(3) The decision time ($\bar{n}$) at which construction should begin if construction is decided.

(4) The optimum storage capacity, $V(j)$ and the periodical storage content $SC(j,m)$ of storage-distribution nodes in an assembly if construction of the housing assembly and storage are decided.

(5) The optimum rating $P(e)$ of new conversion processes, and links in an assembly if construction of the assembly is decided.

(6) The optimum characteristic area $A(j)$ for solar-powered or wind-powered end nodes if construction of the housing assembly is decided.

(7) Dynamic representation of the optimum system energy (or material) flux and cost display, categorized by elements during every time step $\Delta t(m)$, and summed monthly, quarterly, semiannually or annually. This gives a detailed picture of the optimum path of the dynamic facility for future sensitivity analysis.

(8) Diagnostics and error messages: in addition, the programming will print out the problem title supplied, control parameters, problem size, and number of integer variables, bounds on the integer variables, the constraint types, and the matrix format type as part of the initial data. Error messages are printed for abnormal terminations, and suggest the reason and give the iteration number.

The user will assign consecutive numbers for conversion processes/components from 1 to $N_c$, for storage-distribution nodes from 1 to $N_o$, for connecting links from 1 to $N_\ell$, and for years of study period from 1 to $N_y$. Four time steps, not necessarily equal in time, for each representative day of the month will be assumed. This choice is made for better refinement of consumer loads and purchased energy costs during on-peak, midpeak, and off-peak periods. A preprocessor subroutine is envisioned to read a few input data elements, to configure both the objective function coefficients and the nonzero coefficients of the constraint equations, and put them in a "conventional" matrix form ready for execution.

Slack and artificial variables are introduced as appropriate to change inequality constraints to equality forms. Since the decision variables are divided into "continuous" variables (which can take in any value $\geqslant 0$) and integer variables (which are restricted in this model to binary values), the problem is a mixed-integer programming type. Both the objective function and the constraint equations are in linear form, which makes the problem very suitable to branch-and-bound mixed integer linear programming (Refs. 1 and 2). A dual, revised Simplex Algorithm will be imbedded in the program instead of the pivot-Simplex Algorithm for less memory storage, execution time, and cost. Due to the large number of variables that are expected to be dealt with, the diagonalized form of the constraint coefficients and the Dantzig-Wolfe Decomposition Principle (Refs. 1 and 2) will be employed. The set of unknowns or decision variables, $x$ are divided into two types: *Type a* are decision variables that are independent of time-of-day, such as $\bar{A}(j)$, $\beta_1(j)$, $\beta_2(j)$, $V(j)$, $MSC'(j)$ $MSC''(j)$ and $P(e)$; *type b* are decision variables that are time dependent, such as $E(k,m)$, $SC'(j,m)$, $SC''(j,m)$, $P(e,n)$, $\bar{P}(e,n)$, $\lambda(\ell,n)$, $\nu(\ell,m)$ and $\eta(\ell,m)$. Let the constraint equations be written in the matrix format

$$\mathbf{A}\,\mathbf{x} = \mathbf{b} \tag{39}$$

where the form of the vector $\mathbf{x}$ is composed of variables type (a) first, followed by type (b) variables. Further, if all type (b) variables are grouped such that month 1 and year 1 variables appear first, followed by month 2, and year 1 variables, and so on, the matrix $\mathbf{A}$ will be transformed to the preferred shape shown in Figs. 14 and 15. Monthly submatrices (or subproblems) will be placed diagonally together with a top "row" matrix, and a side "column" matrix. Standard Simplex methods (Refs. 1 and 2) using the decomposition principle and dual-Simplex programming could then be adopted. The rest of the matrix $\mathbf{A}$ is full of zeros. No attempt is made at this point to reduce further the monthly submatrix, though filled with many zeros, into smaller submatrices. The compact methods available for storing large matrices will be employed.

## VIII. Summary

In this first part of the optimization study, an analytical model is outlined using linear programming for the optimum generation and distribution of facility demands among competing resources at different design and economic criteria. The model will be used as an in-house engineering tool in the analysis of the Deep Space Network ground facilities and its energy systems and subsystems. The model will provide conceptual as well as design-oriented "optimum" decisions to complement the current practice of designing "working" systems. Some of the decisions included are: the optimum time to build an assembly of elements, the inclusion of a storage medium of some type, the size or capacity of the basic components that will minimize the total life-cycle cost over a given number of years. The broad-class model is structured in piecewise linear, multitime divisions to smooth out nonlinear load effects. Since the number of integer variables, noninteger variables, and constraint equations are found large, it becomes convenient to use the decomposition principle for large-size problems, the Branch-and-Bound Method in mixed-integer programming and the Revised Simplex technique for efficient and economic computer use.

Upon defining a block diagram of a "preferred" system configuration, the user will provide a few input data, partly described in Appendix A, which include the number and type of components, conversion efficiencies, consumer loads, resource intensities, time steps, and various cost and economics data.

A summary of all constraint equations is made in Subsection VI-N with a monthly submatrix form as shown in Fig. 15. For each month, the maximum number of decision variables will be $(6N_{sd} + 10N_{el} + 4N_{\ell} + N_a)$ and the number of constraint equations will be $(4N_c + 4N_o + N_a + 3N_{el} + 9N_{\ell})$. For the left-hand common column matrix, the number of variables is $(N_{rd} + 5N_{sd} + N_{el})$. Also, the number of constraint equations for the top common row matrix is $(72N_y N_{sd} + 5N_{sd} + N_a + 96N_y N_{el})$. Further reduction in the size of the $E(k)$ unknown where $k\ 1, 2, \cdots N_{\ell}$ is possible using the equalities in Eq. (5) such that fewer links are considered. Details of the computer program (OMEGA) that is being written, the testing and application of the model to several case studies, and the sensitivity of the results to changing design or economical factors will be addressed in a future progress report.

## Acknowledgment

The author acknowledges the assistance given by Kyle J. Voss in preparing the initial trial of the Simplex Algorithm. The valuable technical discussions and guidance by F. W. Stoller, R. Z. Toukdarian, and T. Charng have provided numerous insights that led to a broad model selection.

## References

1. Lasdon, L. S., *Optimization Theory for Large Systems*, The MacMillan Co., New York, New York, 1970.
2. Hillier, F. S., and Lieberman, G. J., *Operations Research*, Holden-Day Inc., 2nd Edition, 1974.
3. Stoecker, W. F., *Design of Thermal Systems*, McGraw-Hill Book Co., New York, 1971.
4. Kozlov, V. B., Rudyak, Y. B., "Optimization Model of a System Consisting of Solar Water Heater plus Accumulator plus User with Variable Heat Carrier Flow Rate," *Geliotechnica*, Vol. 14, No. 5, pp. 52-57, 1978.
5. Kydes, A. S., *The Brookhaven Energy System Optimization Model (BESOM) its Variants and Uses*, BNL 50873, UC-13. Brookhaven National Lab., New York, May 1978.
6. Kydes, A. S., and Rabinowitz, J., *The Time-Stepped Energy System Optimization Model (TESOM) Overview and Special Features*, BNL 51095, UC-13, Brookhaven National Lab., Publication, December 1979.
7. Kuester, J. L., and Mize, H. H., *Optimization Techniques with Fortran*, McGraw-Hill Book Co., New York, 1973.
8. Chapman, B. H., "Goldstone DSCC Energy Distribution Model," *The Deep Space Network Progress Report 42-43*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 132-151, February 1978.
9. Ruegg, R. T., *Life Cycle Costing Manual for the Federal Energy Management Programs*. National Bureau of Standards (NBS) Handbook 135, December 1980.
10. Nieves, L. A., "Energy Performance Standards: The Economic Issues," *ASHRAE Journal*, p. 49, January 1981.
11. McKenzie, M., *Life Cycle Cost Analysis*, Telecommunications and Data Acquisition (TDA) Standard Practice Document 810-23, Jet Propulsion Laboratory, Pasadena, Calif., September 1980. (JPL internal document.)
12. McKenzie, M., *Deep Space Network (DSN) Life Cycle Cost Analysis Handbook*, Telecommunications and Data Acquisition (TDA) Standard Practice Document 890-119, Revision A, Jet Propulsion Laboratory, Pasadena, Calif., January 1981. (JPL internal document.)
13. Shultz, G. P., *Discount Rates to be Used in Evaluating Time Distributed Costs and Benefits*, U.S. Office of Management and Budget (OMB), Circular No. A-94, March 27, 1972.

# List of Symbols

$A$ coefficient matrix of constraint equations

$\overline{A}$ characteristic surface area

$b$ constant coefficient vector

$C$ coefficient of the objective function

$Cu$ unit cost of purchased energy (or material)

$CCI$ construction cost intercept

$CCS$ construction cost slope

$D$ consumer energy (or material) demand rate

$e$ element index in an assembly $\ell$

$e'$ energy escalation rate above general inflation

$E$ energy (or material) flux at a sending end of a link

$\overline{E}$ energy (or material) flux at a receiving end of a link

$F$ process efficiency or yield

$i'$ money interest rate (or discount) above general inflation

$i$ conversion process index $(1, 2, \cdots N_c)$

$j$ storage-distribution node index

$k, k_{in}, k_{out}$ link index $(1, 2, \cdots N_\ell)$

$\ell$ assembly index $(1, 2, \cdots N_a)$

$m$ time index for flux calculations $(1, 2, \cdots N_t)$

$m'$ maintenance cost escalation rate above general inflation

$\overline{m}$ time at which energy (or material) flux is computed for a new assembly

$M$ very large number

$MCI$ maintenance cost intercept

$MCS$ maintenance cost slope

$MSC' MSC''$ maximum and minimum storage content of a node

$n$ construction-time index $(1, 2, \cdots N_{ct})$

$\overline{n}$ construction time for a new assembly

$N$ number of constituents

$p$ process-node coupling identifier

$P, \overline{P}$ power or rating of an element

$PW$ present worth

$q, q'$ parameter

$S$ energy resource rate per unit area

$SC$ storage content of node

$t$ time

$T$ transmission efficiency of a link

$TCC$ total construction cost

$TLCC$ total life-cycle cost

$TMC$ total maintenance cost

$TOC$ total operation cost

$V$ storage capacity of a node

$\nu$ monthly index

$X$ unknown (decision) variable

$\alpha$ rate of energy (or material) losses across a link

$\beta_1, \beta_2$ dichotomous variables for storage size decisions

$\lambda, \nu, \eta$ dichotomous variables for construction decisions

$\theta$ months to complete construction

Subscripts

*a* assembly

*c* conversion process

*ct* construction time step

*d* distribution node (no storage)

*dd* demand-distribution end node

*el* elements in an assembly

*H* higher bound

*g* constraint equations

*ℓ* link

*L* lower bound

*o* all nodes

*rd* resource-distribution end node

*sd* storage-distribution node

*t* time-element index

*u* unknown (decision) variables

*y* total years of study

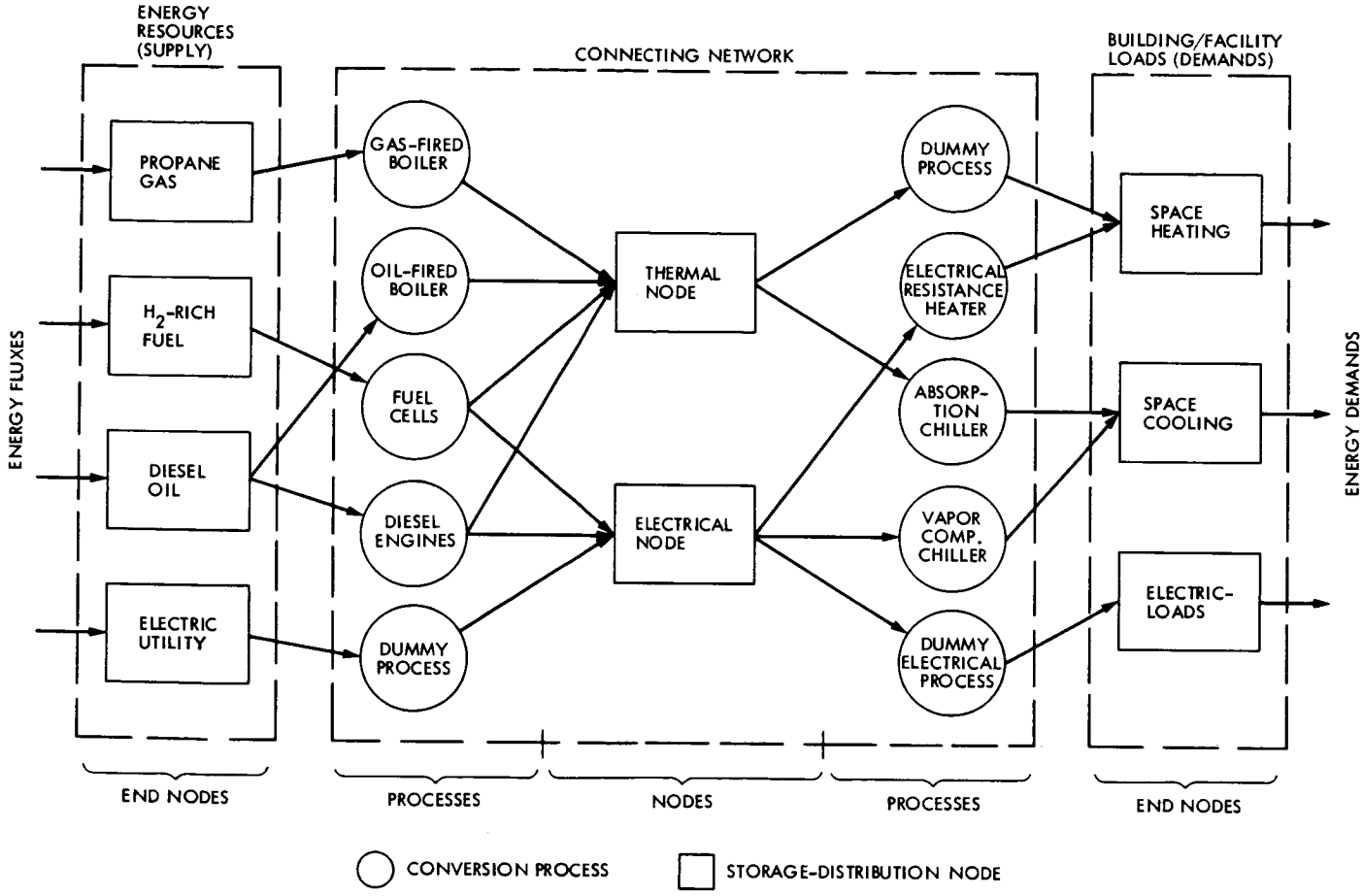



**Fig. 1. Combined fuel cells and oil-fired diesel engines as prime mover with waste heat recovery**

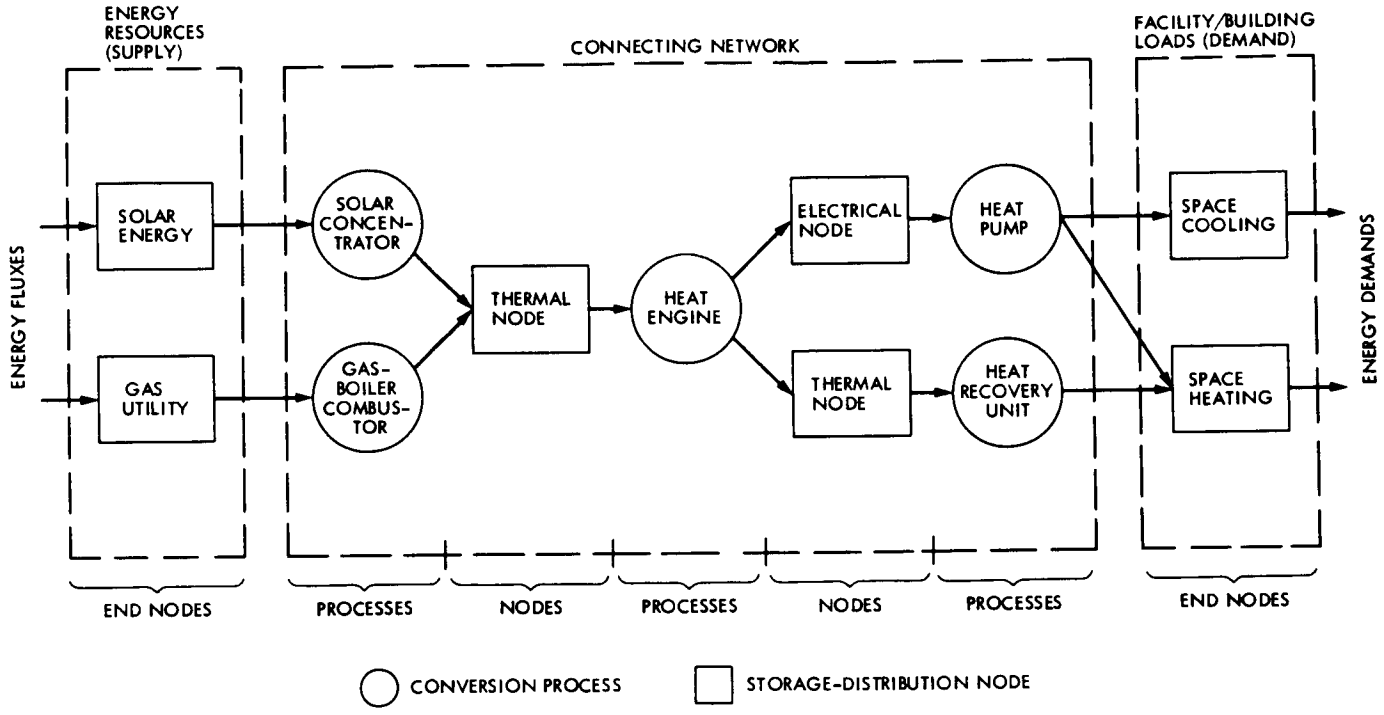



**Fig. 2. Solar-assisted gas-fired heat-pump system**

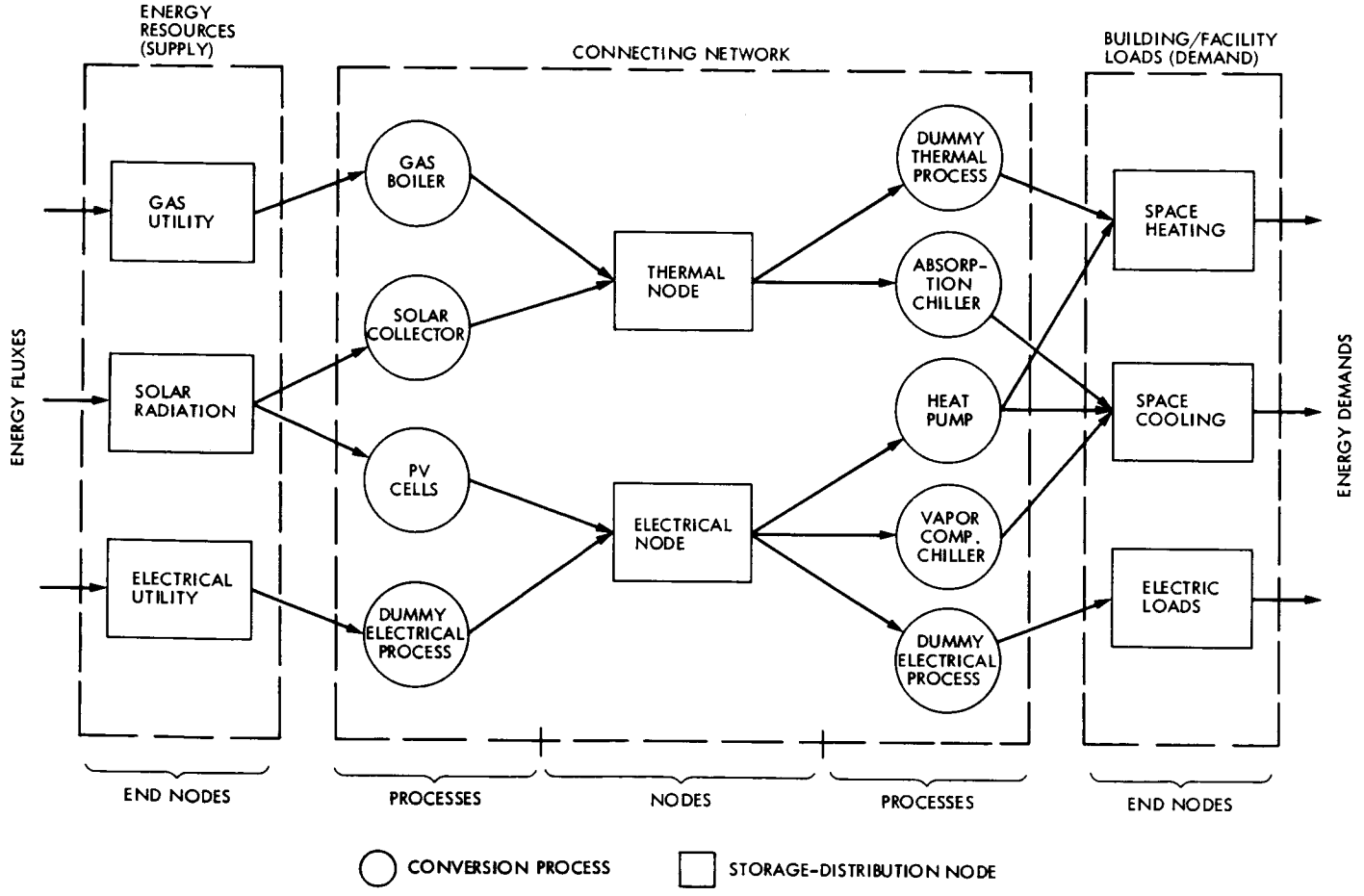



**Fig. 3. Superimposed solar-PV and solar-thermal elements on a conventional gas-electricity building**

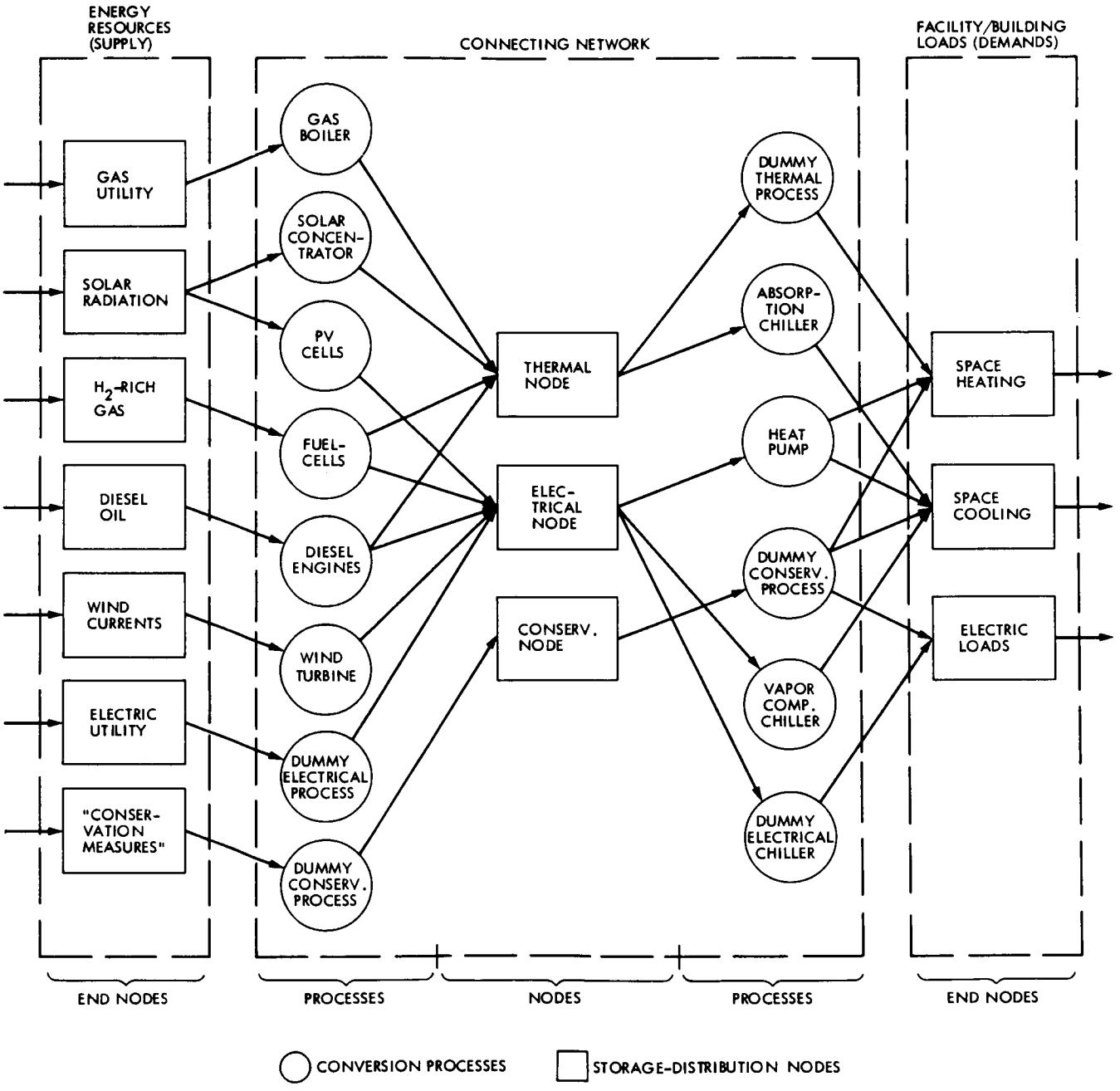



**Fig. 4. A large network for energy generation and distribution including "conservation measures"**

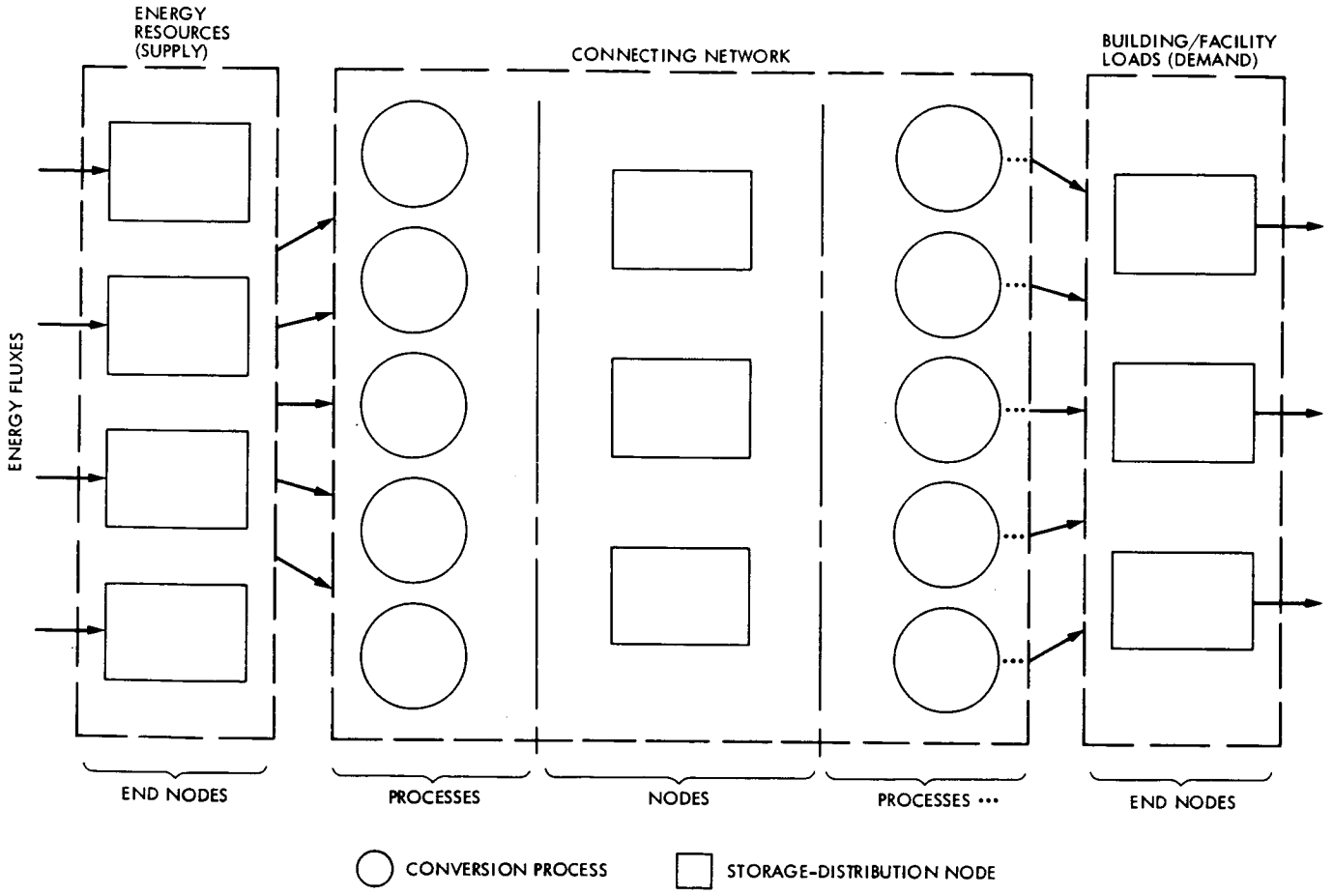



**Fig. 5. Block diagram for general energy generation and distribution system**

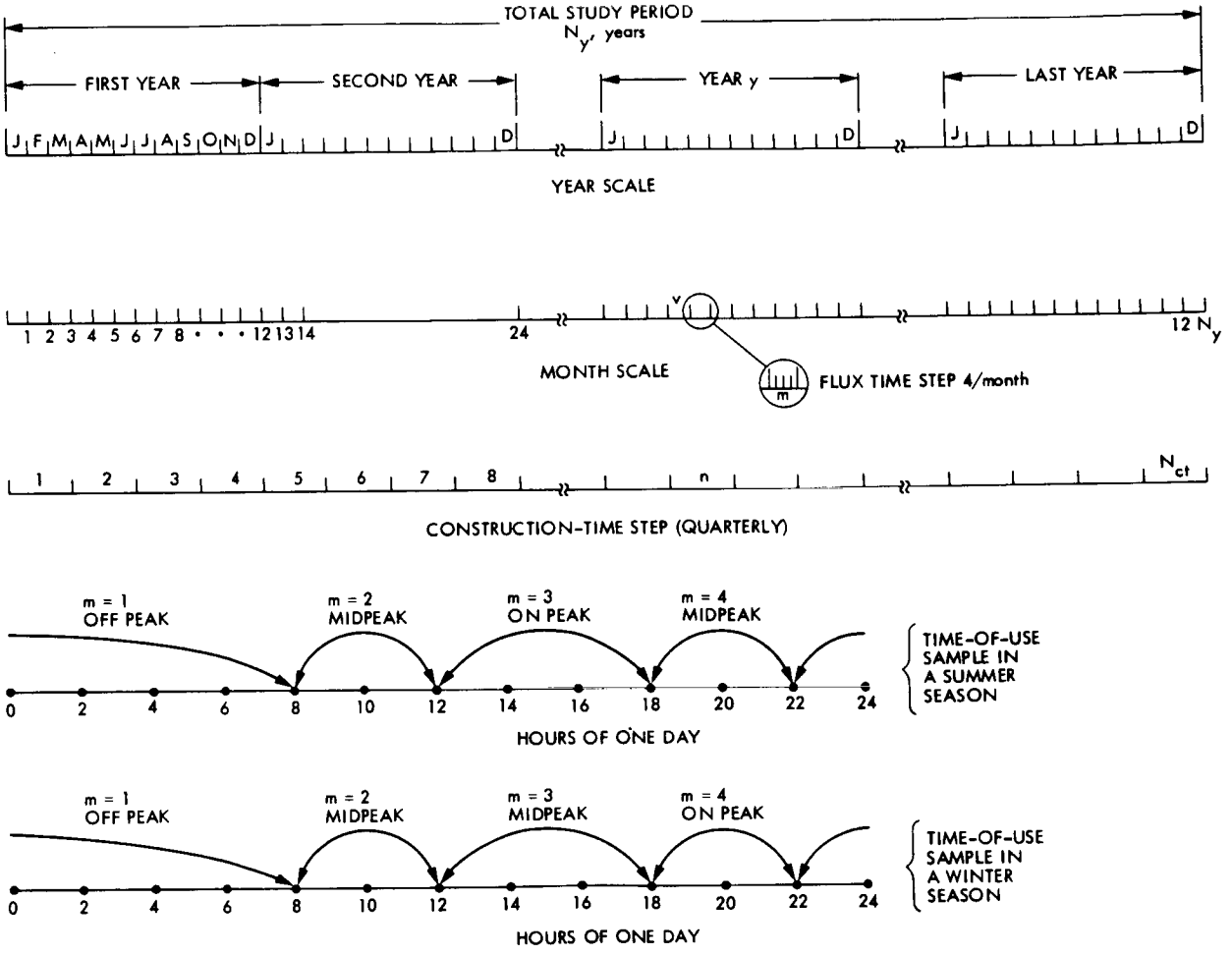



**Fig. 6. Time steps**

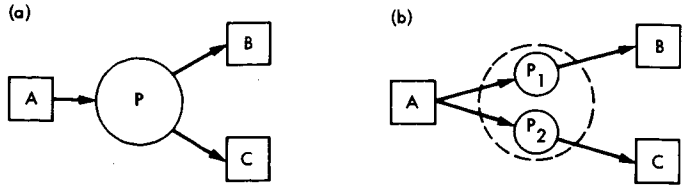



**Fig. 7. Decomposition of a compound process into simple processes: (a) compound process $P$; (b) equivalent simple processes $P_1, P_2$**

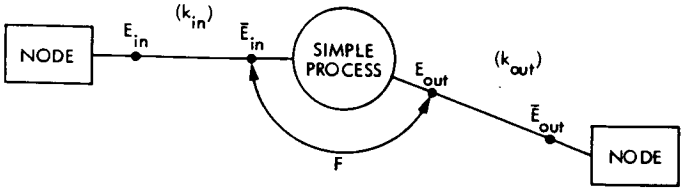



**Fig. 8. Conversion efficiency of a single process**

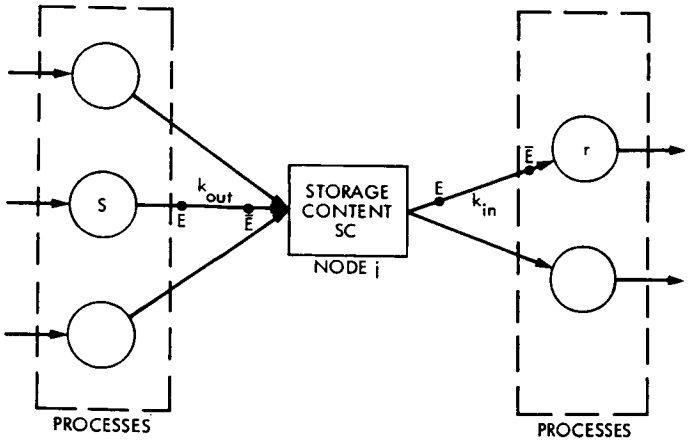



**Fig. 9. Balancing fluxes around a storage-distribution node**

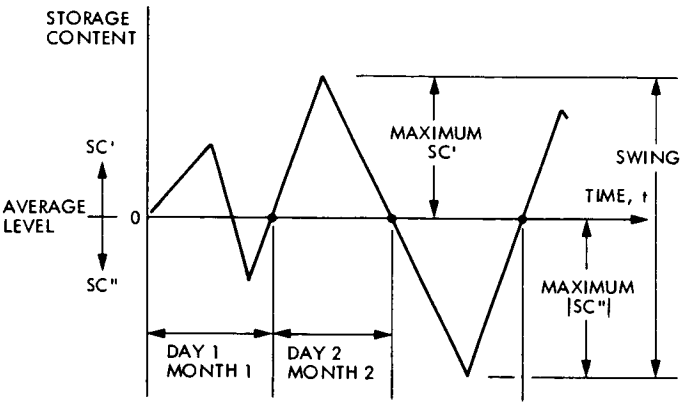



**Fig. 10. Storage content at different time steps**

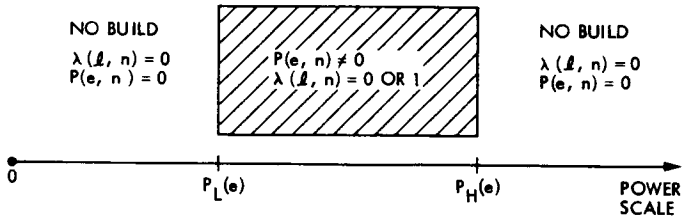



**Fig. 11. Lower and upper bounds for an element $e$ in an assembly at construction time $n$**

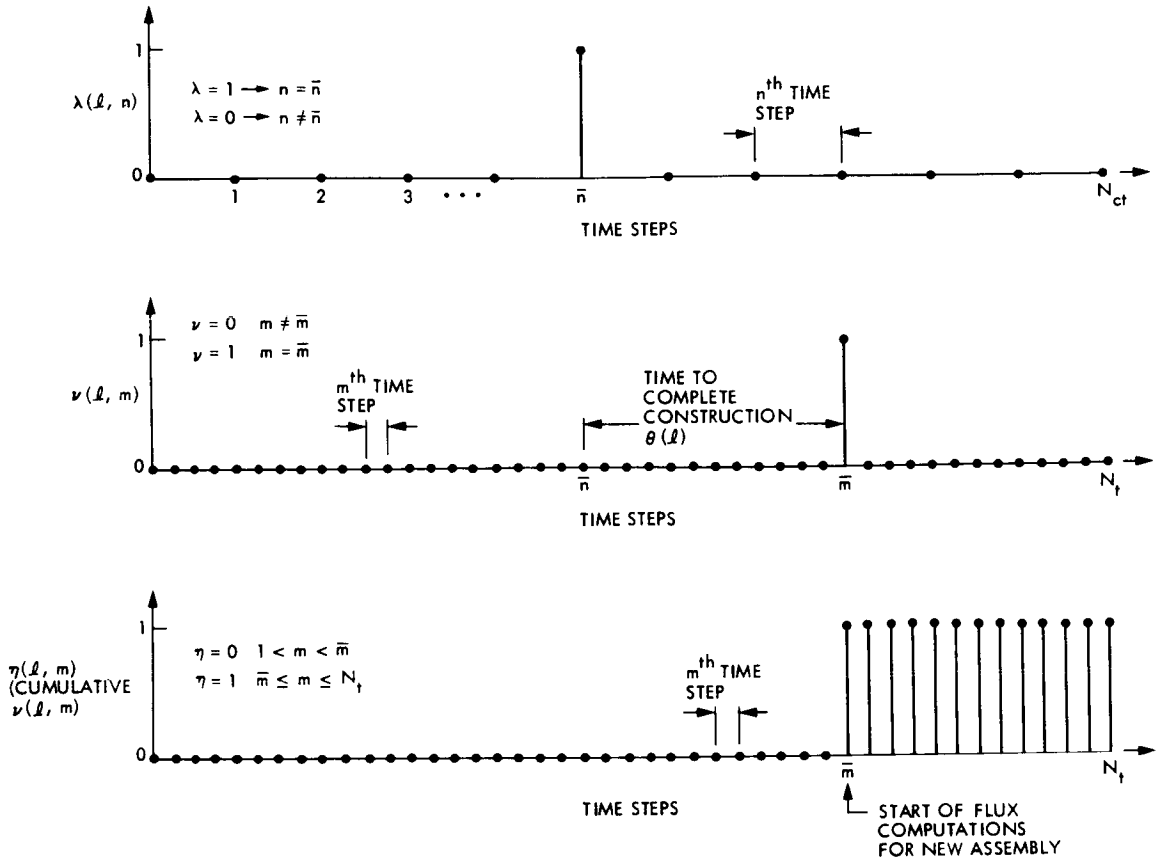



**Fig. 12. Relationships between dichotomous variables λ, ν, η**

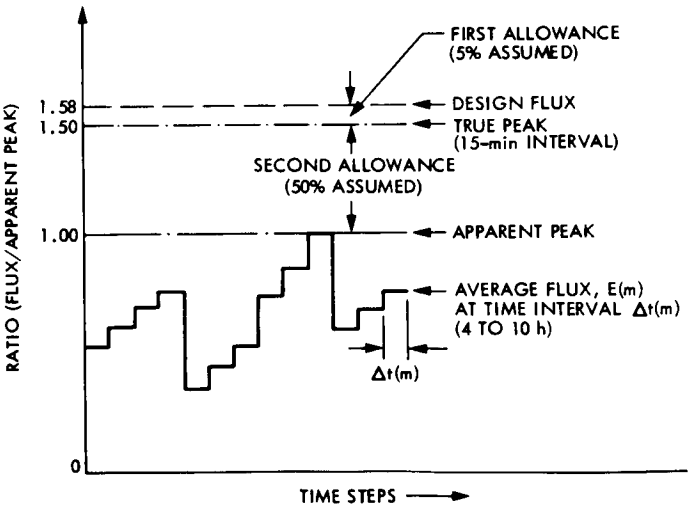



**Fig. 13. Relationship between flux average, apparent peak, true peak, and design capacity**

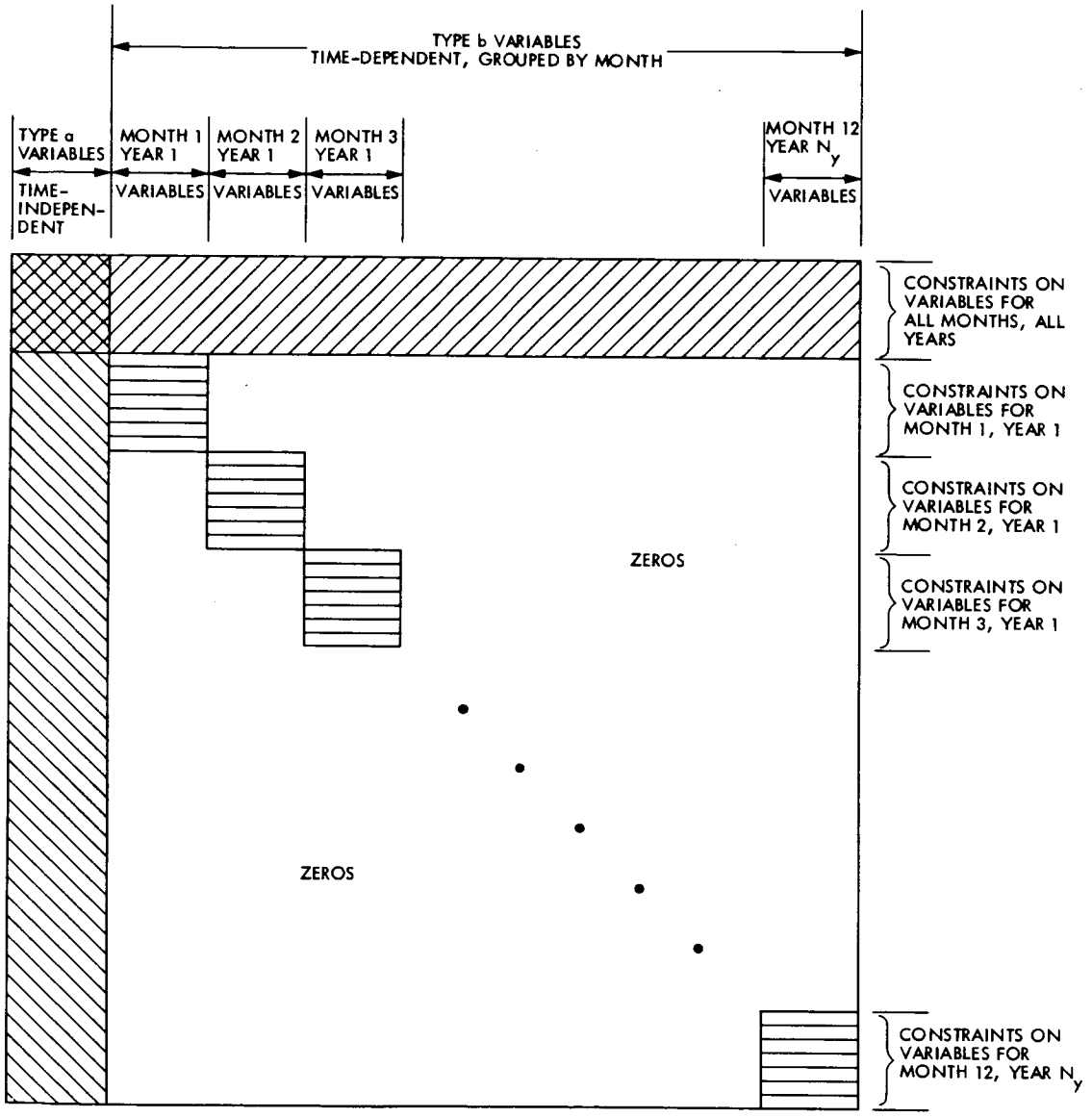
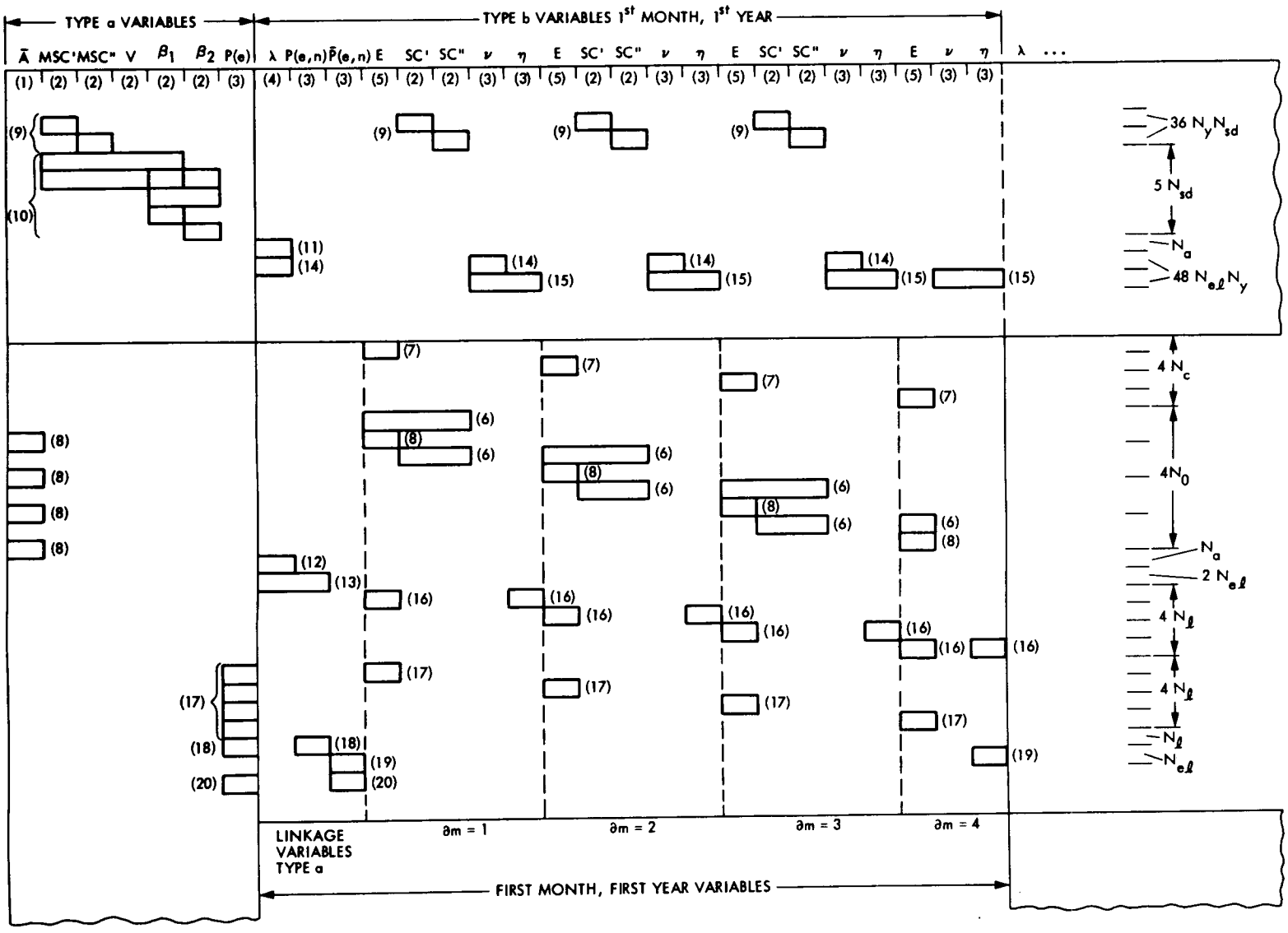



**Fig. 14. Structure of the nonzero elements of the coefficient matrix A**




(1) $N_{rd}$ ELEMENTS
(2) $N_{sd}$ ELEMENTS
(3) $N_{e\ell}$ ELEMENTS
(4) $N_a$ ELEMENTS
(5) $N_\ell$ ELEMENTS

(6) Eqs. (8), (12)
(7) Eq. (5)
(8) Eqs. (9), (10), (11)
(9) Eqs. (19), (20)
(10) Eq. (18)

(11) Eq. (21)
(12) Eq. (22)
(13) Eqs. (25, (26), (27)
(14) Eq. (29)
(15) Eq. (30)

(16) Eq. (31)
(17) Eq. (32)
(18) Eq. (33)
(19) Eq. (36)
(20) Eq. (37)

**Fig. 15. Details of the submatrices of A**

# Appendix A
# Selected Input Data

Tables A-1 through A-10 list the requirements to be considered by the user in entering some selected variables. A preprocessor routine will organize the user information and generate the necessary coefficients of both the constraint equations and the objective function. The user must supply the proper index number wherever a dot appears in the table.

**Table A-1. Description matrix for conversion processes, $DC(N_c, 2)$**

| Process No. | Description | Housing Assembly No. |
|---|---|---|
| 1 | • | • |
| 2 | • | • |
| $i$ | • ← 0, 1, 2 | • |
| $N_c$ | • | • |

All conversion processes will be numbered including dummy processes, consecutively starting from 1, in steps of 1 up to the maximum number of processes $N_c$. The elements of description column will take in 0, 1, or 2 where

0 = a new process under optimization decisions

1 = a dummy process

2 = an existing process

**Table A-2. Description matrix for storage-distribution nodes, $DS(N_s, 2)$**

| Node No. | Description | Housing Assembly No. |
|---|---|---|
| 1 | • | • |
| 2 | • ← 0, 1 | • |
| $j$ | • ← 2, 3 | • |
| $N_o$ | • ← 4 | • |

All storage-distribution nodes will be numbered, including dummy nodes, consecutively starting from 1 in steps of 1, up to the maximum number of nodes $N_0$. The elements of the description column will take 0, 1, 2, 3 or 4 where

0 = a new node under optimization decisions

1 = a dummy node or an existing node without storage

2 = an existing node with storage

3 = an end node for energy resource (supply)

4 = an end node for facility loads (demand)

**Table A-3. Description of assemblies, $DA(N_a, 2)$**

| Assembly No. | Description | Construction time $\theta$, months |
|---|---|---|
| 1 | • | • |
| 2 | • | • |
| $\ell$ | • ← 0, 1 | • |
| $N_a$ | • | • |

All conversion processes, storage-distribution nodes, and the associated linkages will be grouped into assemblies that could have some overlap elements. Assemblies are numbered consecutively starting from 1, in steps of 1 up to the maximum number of assemblies $N_a$. The elements of the description column will take in 0 or 1 where

0 = a new assembly under optimization decision

1 = existing assembly

**Table A-4. Coupling matrix between conversion processes and storage-distribution nodes, $P(N_c, N_o)$**

| Process No. | 1 | 2 | $j$ | $N_o$ |
|---|---|---|---|---|
| 1 | • | • | • | • |
| 2 | • | • | • | • |
| $i$ | • | • | $k$ | • |
| $N_c$ | • | • | • | • |

The coupling matrix, $p(i,j)$ is an asymmetric matrix that relates the $i^{th}$ process with the $j^{th}$ node and the $k^{th}$ link connecting them. $p(i,j)$ will take in

0 if no link exists between process $i$ and node $j$

$k$ if flux in link $k$ is from process $i$ into node $j$

$-k$ if flux in link $k$ is from node $j$ into process $i$

The elements of the $i^{th}$ row of the matrix $p$ identify all the nodes connected to process $i$ through inlet and outlet links. The elements of the $j^{th}$ column of the matrix $p$ identifies all the processes connected to the $j^{th}$ node.

**Table A-5. Conversion efficiency, F**

| Links at process inlet | 1 | 2 | $k_{out}$ | $N_\ell$ |
|---|---|---|---|---|
| 1 | • | • | • | • |
| 2 | • | • | • | • |
| $k_{in}$ | • | • | $F$ | • |
| $N_\ell$ | • | • | • | • |

The conversion $F$ is an asymmetric matrix that gives the ratio of output to input fluxes $E(k_{out})/\bar{E}(k_{in})$ associated with simple processes, at a given time period. Several matrices for $F$ could be specified for different time periods to allow inefficient operation due to equipment aging and partial loading. Elements of one row represent the products efficiency in converting the input link flux. Elements of one column, however, represent the relative contributions by all inlet links to the selected outlet link flux.

**Table A-6. Facility loads for demand-end nodes**

| | Time step | | | |
|---|---|---|---|---|
| Demand node No. | 1 | 2 | $m$ | $N_t$ |
| 1 | • | • | • | • |
| 2 | • | • | • | • |
| $j$ | • | • | $D(j,m)$ | • |
| $N_{dd}$ | • | • | • | • |

Each node represents one type of facility load such as electrical, space heating, space cooling, domestic water heating, and process steam. Data are entered for each time step during the study period; hence allowance is made for changes due to facility growth or decreases in activity by time.

**Table A-7. Energy resource intensities for supply-end nodes**

| | Time step | | | |
|---|---|---|---|---|
| Resource node No. | 1 | 2 | $m$ | $N_t$ |
| 1 | • | • | • | • |
| 2 | • | • | • | • |
| $j$ | • | • | $S(j,m)$ | • |
| $N_{rd}$ | • | • | • | • |

Each node represents one type of nondepletable energy resource (such as solar and wind) whose intensity $S$ is given for the appropriate time periods. A preprocessor should be included to obtain the data of Table A-7 from other sources of information.

For the following three tables (A-8, A-9, and A-10), the NBS rules for LCC computations (Ref. 9) are summarized as follows:

(1) All investment costs, nonfuel operation and maintenance costs, repair and replacement costs, salvage values, and energy costs should be accounted for.

(2) All future dollar amounts must be estimated in "constant dollars" (i.e., excluding the effects of general price inflation) and not in "current dollars."

(3) A real interest (or discount) rate, $i'$, also excluding inflation, must be used to adjust all dollar values to a present worth in the base year.

(4) Energy prices and price growth projections provided by The Department of Energy (DOE) may be used unless the actual prices to the facility are higher. Projections are provided (Ref. 9) by region, consuming sector, fuel type, and time period.

(5) The study period should be the lesser of 25 years or: (a) the expected life of the system for a building system retrofit, (b) the period of intended use for a new building design, (c) the effective remaining term of the lease for a leased building, or (d) an equivalent study period where choices are mutually exclusive.[1]

**Table A-8. Construction and maintenance costs**

| Type of element | Description | Construction cost | | Maintenance cost | |
|---|---|---|---|---|---|
| | | Fixed CCI | Variable CCS | Fixed MCI | Variable MCS |
| Conversion processes | Solar collector | • | • | • | • |
| | PV cells | • | • | • | • |
| | Wind turbine | • | • | • | • |
| | Heat pump | • | • | • | • |
| | Boiler | • | • | • | • |
| | Electrical resistance | • | • | • | • |
| Storage–distribution nodes | Cold water tank | • | • | • | • |
| | Hot water tank | • | • | • | • |
| | Battery | • | • | • | • |
| Links | Piping | • | • | • | • |
| | Electrical line | • | • | • | • |

[1] Mutually exclusive means that choosing one alternative precludes choosing the other. Nonmutually exclusive choices means making one choice does not necessarily preclude making the other.

The common study period for evaluating mutually exclusive choices may be either: (a) the estimated life of the choice having the longest life, or (b) the lowest common multiple of the estimated lives of the alternatives not to exceed 25 years.

**Table A-9. Purchased Energy (or material) costs at beginning of base year**

| Type of resource node | Unit cost | | | | | |
|---|---|---|---|---|---|---|
| | Summer | | | Winter | | |
| | Off peak | Mid-peak | On peak | Off peak | Mid-peak | On peak |
| Electricity | • | • | • | • | • | • |
| Natural gas | • | • | • | • | • | • |
| Diesel oil | • | • | • | • | • | • |
| LPG | • | • | • | • | • | • |

**Table A-10. Escalation rates above general inflation**

| Escalation rate | | Month | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | $v$ | $12N_y$ |
| Money $i'$ | | • | • | • | • |
| Energy $e'$ | Electricity | • | • | • | • |
| | Gas | • | • | • | • |
| | Oil | • | • | • | • |
| Maintenance $m'$ | | • | • | • | • |

Monthly escalation rates are computed from yearly rates by simple (not compound) division. No escalation is considered for time steps less than one month. Costs are incurred at the *end* of each month.

TECHNICAL REPORT STANDARD TITLE PAG

| 1. Report No. TDA PR 42-65 | 2. Government Accession No. | 3. Recipient's Catalog No. |
|---|---|---|
| 4. Title and Subtitle<br>The Telecommunications and Data Acquisition Progress Report 42-65 | | 5. Report Date<br>10-15-81<br>6. Performing Organization Code |
| 7. Author(s)<br>N. A. Renzetti (ed.) | | 8. Performing Organization Report No |
| 9. Performing Organization Name and Address<br>JET PROPULSION LABORATORY<br>California Institute of Technology<br>4800 Oak Grove Drive<br>Pasadena, California 91103 | | 10. Work Unit No.<br>11. Contract or Grant No.<br>NAS 7-100 |
| 12. Sponsoring Agency Name and Address<br>NATIONAL AERONAUTICS AND SPACE ADMINISTRATION<br>Washington, D.C. 20546 | | 13. Type of Report and Period Covered<br>TDA Progress Report<br>July-August 1981<br>14. Sponsoring Agency Code<br>RD142 G-311-03-31-10-83 |

15. Supplementary Notes

16. Abstract

This publication, one of a series formerly titled The Deep Space Network Progress Report, documents DSN progress in flight project support, tracking and data acquisition research and technology, network engineering, hardware and software implementation, and operations. In addition, developments in Earth-based radio technology as applied to geodynamics, astrophysics, and the radio search for extraterrestrial intelligence are reported.

| 17. Key Words (Selected by Author(s)) | 18. Distribution Statement |
|---|---|
| Astronautics<br>Communications<br>Mathematical and Computer Sciences<br>Space Sciences | Unclassified -- Unlimited |

| 19. Security Classif. (of this report) | 20. Security Classif. (of this page) | 21. No. of Pages | 22. Price |
|---|---|---|---|
| Unclassified | Unclassified | 244 | |